Introductory Course on Non-smooth Optimisation

Lecture 05 - Peaceman–Rachford, Douglas–Rachford splitting

Jingwei Liang

Department of Applied Mathematics and Theoretical Physics

Table of contents

Sum of two operators

Problem

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

Assumptions

- $A, B : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ are maximal monotone.
- the resolvents of A, B are simple, *i.e.* easy to compute.
- $\mathrm{zer}(A + B) \neq \emptyset$.

Outline

Peaceman–Rachford splitting

Let $z_0 \in \mathbb{R}^n$, $\gamma > 0$:

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

- dates back to 1950s for solving numerical PDEs.
- the resolvents of A, B are evaluated separately.

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

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

- apply the resolvent

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

- equivalent formulation

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

- fixed-point iteration

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

Fixed-point formulation Recall reflection operator $\mathcal{R}_{\gamma A} = 2\mathcal{J}_{\gamma A} - \mathrm{Id}$.

- $y_k = \mathcal{J}_{\gamma A}(2x_k - z_k) = \mathcal{J}_{\gamma A} \circ (2\mathcal{J}_{\gamma B} - \mathrm{Id})(z_k)$.
- For z_k,

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

Property

- $\mathcal{R}_{\gamma A} = 2\mathcal{J}_{\gamma A} - \mathrm{Id}, \mathcal{R}_{\gamma B} = 2\mathcal{J}_{\gamma B} - \mathrm{Id}$ are non-expansive.
- $\mathcal{T}_{\mathrm{PR}} = \mathcal{R}_{\gamma A} \circ \mathcal{R}_{\gamma B}$ is non-expansive.

NB: Cannot guarantee convergence in general.

- Uniform monotonicity: $\phi : \mathbb{R}_+ \to [0, +\infty]$ is increasing and vanishes only at 0

$$\langle u - v,\, x - y\rangle \geq \phi(\|x - y\|),\ \ (x, u), (y, v) \in \mathrm{gra}(B).$$

- If B is uniformly monotone, then $\mathrm{zer}(A + B) = \{x^\star\}$ and $\mathrm{fix}(\mathcal{T}_{\mathrm{PR}}) \neq \emptyset$. Moreover

$$\langle x - y,\, \mathcal{J}_{\gamma B}(x) - \mathcal{J}_{\gamma B}(y)\rangle \geq \|\mathcal{J}_{\gamma B}(x) - \mathcal{J}_{\gamma B}(y)\|^2 + \gamma\phi(\|\mathcal{J}_{\gamma B}(x) - \mathcal{J}_{\gamma B}(y)\|).$$

- Let $z^\star \in \mathrm{fix}(\mathcal{T}_{\mathrm{PR}})$, then $x^\star = \mathcal{J}_{\gamma A}(z^\star)$, and

$$\begin{aligned}\|z_{k+1} - z^\star\|^2 &= \|\mathcal{R}_{\gamma A}\mathcal{R}_{\gamma B}(z_k) - \mathcal{R}_{\gamma A}\mathcal{R}_{\gamma B}(z^\star)\|^2 \\ &\leq \|(2\mathcal{J}_{\gamma B} - \mathrm{Id})(z_k) - (2\mathcal{J}_{\gamma B} - \mathrm{Id})(z^\star)\|^2 \\ &= \|z_k - z^\star\|^2 - 4\langle z_k - z^\star,\, \mathcal{J}_{\gamma B}(z_k) - \mathcal{J}_{\gamma B}(z^\star)\rangle + 4\|\mathcal{J}_{\gamma B}(z_k) - \mathcal{J}_{\gamma B}(z^\star)\|^2 \\ &\leq \|z_k - z^\star\|^2 - 4\gamma\phi(\|\mathcal{J}_{\gamma B}(z_k) - \mathcal{J}_{\gamma B}(z^\star)\|).\end{aligned}$$

- $\phi(\|z_k - z^\star\|) \to 0$ and $\|z_k - z^\star\| \to 0$.

Outline

To overcome the drawback of Peaceman–Rachford splitting.

Douglas–Rachford splitting

Let $z_0 \in \mathbb{R}^n$, $\gamma > 0$, $\lambda \in]0, 2[$:

$$\begin{aligned} x_k &= \mathcal{J}_{\gamma B}(z_k), \\ y_k &= \mathcal{J}_{\gamma A}(2x_k - z_k), \\ z_{k+1} &= z_k + \lambda(y_k - x_k). \end{aligned}$$

How to derive

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

$$\begin{cases} z^\star - x^\star \in \gamma A(x^\star) \\ x^\star - z^\star \in \gamma B(x^\star) \end{cases} \Longrightarrow \begin{cases} \quad\quad z^\star \in x^\star + \gamma A(x^\star), \\ 2x^\star - z^\star \in x^\star + \gamma B(x^\star). \end{cases}$$

- apply the resolvent

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

- equivalent formulation

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

- fixed-point iteration

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

Fixed-point formulation Same as PR, $y_k = \mathcal{J}_{\gamma A} \circ \mathcal{R}_{\gamma B}(z_k)$

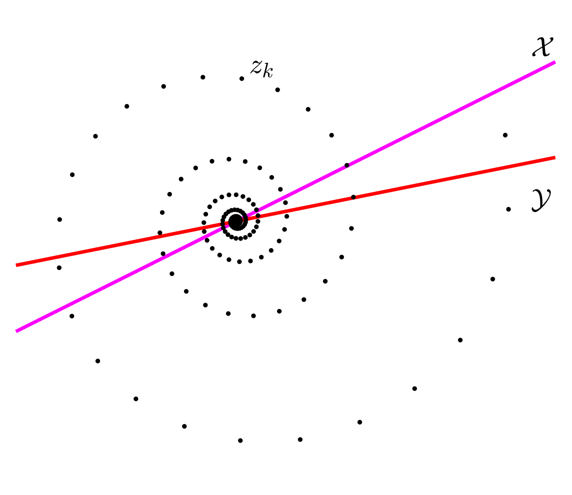
$$\begin{aligned} z_{k+1} &= (1-\lambda)z_k + \lambda\big(z_k + (y_k - x_k)\big) \\ &= (1-\lambda)z_k + \lambda\big(\tfrac{1}{2}z_k + \tfrac{1}{2}(z_k + 2(y_k - x_k))\big) \\ &= (1-\lambda)z_k + \lambda\tfrac{1}{2}(\mathrm{Id} + \mathcal{R}_{\gamma A} \circ \mathcal{R}_{\gamma B})(z_k). \end{aligned}$$

Property

- $\mathcal{T}_{\mathrm{DR}} = \frac{1}{2}(\mathrm{Id} + \mathcal{R}_{\gamma A} \circ \mathcal{R}_{\gamma B})$ is firmly non-expansive.
- $\mathcal{T}_{\mathrm{DR}}^{\lambda} = (1-\lambda)\mathrm{Id} + \lambda\mathcal{T}_{\mathrm{DR}}$ is $\frac{\lambda}{2}$-averaged non-expansive.
- Peaceman–Rachford is the limiting case of Douglas–Rachford, $\lambda = 2$.

NB: guaranteed convergence if $\lambda(2-\lambda) > 0$.

Convergence rate



Convergence rate

- Let $\mathcal{X}, \mathcal{Y}$ be two subspaces
$$\mathcal{X} = \{x : ax = 0\}, \;\; \mathcal{Y} = \{x : bx = 0\}$$
and assume
$$1 \leq p \stackrel{\text{def}}{=} \mathrm{dim}(\mathcal{X}) \leq q \stackrel{\text{def}}{=} \mathrm{dim}(\mathcal{Y}) \leq n - 1.$$

- Projection onto subspace
$$\mathcal{P}_{\mathcal{X}}(x) = x - a^T (aa^T)^{-1} ax.$$

- Define diagonal matrices
$$c = \mathrm{diag}\big(\cos(\theta_1), \cdots, \cos(\theta_p)\big),$$
$$s = \mathrm{diag}\big(\sin(\theta_1), \cdots, \sin(\theta_p)\big).$$

Convergence rate

- Suppose $p+q<n$, then there exists orthogonal matrix U such that

$$\mathcal{P}_{\mathcal{X}} = U \left[\begin{array}{cc|cc} \mathrm{Id}_p & 0 & 0 & 0 \\ 0 & 0_p & 0 & 0 \\ \hline 0 & 0 & 0_{q-p} & 0 \\ 0 & 0 & 0 & 0_{n-p-q} \end{array}\right] U^*$$

and

$$\mathcal{P}_{\mathcal{Y}} = U \left[\begin{array}{cc|cc} c^2 & cs & 0 & 0 \\ cs & c^2 & 0 & 0 \\ \hline 0 & 0 & \mathrm{Id}_{q-p} & 0 \\ 0 & 0 & 0 & 0_{n-p-q} \end{array}\right] U^*.$$

Convergence rate

- For the composition

$$\mathcal{P}_{\mathcal{X}} \circ \mathcal{P}_{\mathcal{Y}} = U \left[\begin{array}{cc|cc} c^2 & cs & 0 & 0 \\ 0 & 0_p & 0 & 0 \\ \hline 0 & 0 & 0_{q-p} & 0 \\ 0 & 0 & 0 & 0_{n-p-q} \end{array}\right] U^*$$

  and

$$\mathcal{P}_{\mathcal{X}^\perp} \circ \mathcal{P}_{\mathcal{Y}^\perp} = U \left[\begin{array}{cc|cc} 0_p & 0 & 0 & 0 \\ -cs & c^2 & 0 & 0 \\ \hline 0 & 0 & 0_{q-p} & 0 \\ 0 & 0 & 0 & \mathrm{Id}_{n-p-q} \end{array}\right] U^*.$$

- Fixed-point operator

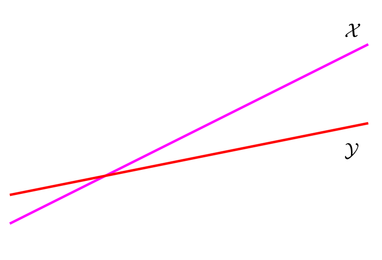
$$\begin{aligned}\mathcal{T}_{\mathrm{DR}} &= \mathcal{P}_{\mathcal{X}} \circ \mathcal{P}_{\mathcal{Y}} + \mathcal{P}_{\mathcal{X}^\perp} \circ \mathcal{P}_{\mathcal{Y}^\perp} \\ &= U \left[\begin{array}{cc|cc} c^2 & cs & 0 & 0 \\ -cs & c^2 & 0 & 0 \\ \hline 0 & 0 & 0_{q-p} & 0 \\ 0 & 0 & 0 & \mathrm{Id}_{n-p-q} \end{array}\right] U^*.\end{aligned}$$

- Consider relaxation

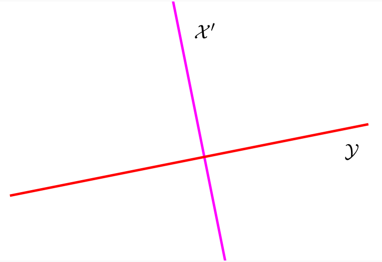
$$\begin{aligned}\mathcal{T}_{\mathrm{DR}}^{\lambda} &= (1-\lambda)\mathrm{Id} + \lambda \mathcal{T}_{\mathrm{DR}} \\ &= U \left[\begin{array}{cc|cc} \mathrm{Id}_p - \lambda s^2 & \lambda cs & 0 & 0 \\ -\lambda cs & \mathrm{Id}_p - \lambda s^2 & 0 & 0 \\ \hline 0 & 0 & (1-\lambda)\mathrm{Id}_{q-p} & 0 \\ 0 & 0 & 0 & \mathrm{Id}_{n-p-q} \end{array}\right] U^*.\end{aligned}$$

Convergence rate

- Eigenvalues
$$\sigma(\mathcal{T}_{\mathrm{DR}}^{\lambda}) = \begin{cases} \{1-\lambda\sin^2(\theta_i) \pm \mathrm{i}\lambda\cos(\theta_i)\sin(\theta_i) | i=1,...,p\} \cup \{1\} : q = p, \\ \{1-\lambda\sin^2(\theta_i) \pm \mathrm{i}\lambda\cos(\theta_i)\sin(\theta_i) | i=1,...,p\} \cup \{1\} \cup \{1-\lambda\} : q > p. \end{cases}$$
- Complex eigenvalues
$$|1-\lambda\sin^2(\theta_i) \pm \mathrm{i}\lambda\cos(\theta_i)\sin(\theta_i)| = \sqrt{\lambda(2-\lambda)\cos^2(\theta_i) + (1-\lambda)^2}$$
and
$$1 \geq \sqrt{\lambda(2-\lambda)\cos^2(\theta_i) + (1-\lambda)^2} \geq |1-\lambda|.$$
- $\lim_{k\to+\infty} \mathcal{T}_{\mathrm{DR}}^{k} = \mathcal{T}_{\mathrm{DR}}^{\infty}$ and $z_k - z^\star = (\mathcal{T}_{\mathrm{DR}} - \mathcal{T}_{\mathrm{DR}}^{\infty})(z_{k-1} - z^\star)$.
- Spectral radius, minimises at $\lambda = 1$
$$\rho(\mathcal{T}_{\mathrm{DR}} - \mathcal{T}_{\mathrm{DR}}^{\infty}) = \sqrt{\lambda(2-\lambda)\cos^2(\theta_i) + (1-\lambda)^2}.$$
- $\widetilde{\mathcal{T}_{\mathrm{DR}}} = \mathcal{T}_{\mathrm{DR}} - \mathcal{T}_{\mathrm{DR}}^{\infty}$
$$\begin{aligned} \|z_k - z^\star\| &= \|\widetilde{\mathcal{T}_{\mathrm{DR}}} z_{k-1} - \widetilde{\mathcal{T}_{\mathrm{DR}}} z^\star\| = \ldots = \|\widetilde{\mathcal{T}_{\mathrm{DR}}}^{k}(z_0 - z^\star)\| \\ &\leq C\big(\rho(\widetilde{\mathcal{T}_{\mathrm{DR}}})\big)^k \|z_0 - z^\star\|. \end{aligned}$$



$\mathcal{X}$ and $\mathcal{Y}$



$\mathcal{X}'$ and $\mathcal{Y}$

Optimal metric A invertable operation which makes the Friedrichs angle between $\mathcal{X}'$ and $\mathcal{Y}$ the largest, *e.g.* $\frac{\pi}{2}$...

Problem $s \in \mathbb{N}_+$ and $s \geq 2$

$$\text{Find } \ x \in \mathbb{R}^n \ \text{ such that } \ 0 \in \textstyle\sum_i A_i(x).$$

Assumptions

- for each $i = 1, ..., s$, $A_i : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is maximal monotone.
- $\mathrm{zer}(\sum_i A_i) \neq \emptyset$.

Product space

- Let $\mathcal{H} = \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{s \text{ times}}$ endowed with the scalar inner-product and norm

$$\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{H},\ \langle \boldsymbol{x}, \boldsymbol{y} \rangle = \textstyle\sum_{i=1}^{s} \langle x_i,\ y_i \rangle,\ \|\boldsymbol{x}\| = \sqrt{\textstyle\sum_{i=1}^{s} \|x_i\|^2}.$$

- Let

$$\boldsymbol{\mathcal{S}} = \{ \boldsymbol{x} = (x_i)_i \in \mathcal{H} : x_1 = \cdots = x_s \}$$

and its orthogonal complement

$$\boldsymbol{\mathcal{S}}^{\perp} = \{ \boldsymbol{x} = (x_i)_i \in \mathcal{H} : \textstyle\sum_{i=1}^{s} x_i = 0 \}.$$

Define $\boldsymbol{A}$ by

$$\boldsymbol{A}(\boldsymbol{x}) : \boldsymbol{x} \in \boldsymbol{\mathcal{H}} \to A_1(x_1) \times \cdots \times A_s(x_s).$$

Lifted problem

$$\text{Find } \ \boldsymbol{x} \in \boldsymbol{\mathcal{H}} \ \text{ such that } \ 0 \in \boldsymbol{A}(\boldsymbol{x}) + \mathcal{N}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{x}).$$

- the resolvent of $\boldsymbol{A}$ is seperable, *i.e.* $\mathcal{J}_{\gamma \boldsymbol{A}} = (\mathcal{J}_{\gamma A_i})_i$.
- define the canonical isometry,

$$\boldsymbol{C} : \mathbb{R}^n \to \boldsymbol{\mathcal{S}},\ x \mapsto (x, \cdots, x),$$

then $\mathcal{P}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{z}) = \boldsymbol{C}(\frac{1}{s}\sum_{i=1}^{s} z_i)$.

Outline

DR in product space for $\boldsymbol{x}^\star \in \boldsymbol{\mathcal{S}}$, $\exists - \boldsymbol{v} \in \boldsymbol{\mathcal{S}}$ such that

$$-\boldsymbol{v} \in \boldsymbol{\mathcal{S}}^\perp = \mathcal{N}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{x}^\star) \quad \text{and} \quad \boldsymbol{v} \in \boldsymbol{A}(\boldsymbol{x}^\star).$$

Problem V is a close subspace

$$\text{Find } \ x \in V \text{ and } v \in V^\perp \ \text{ such that } \ v \in A(x).$$

Assumptions

- $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is maximal monotone.
- admits at least one solution.

Partial inverse

Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be set-valued and $V \subseteq \mathbb{R}^n$ be a closed subspace. The partial inverse of A respect to V is the operator $A_V : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ define by

$$\mathrm{gra}(A_V) = \big\{ \big(\mathcal{P}_V(x) + \mathcal{P}_{V^\perp}(u), \mathcal{P}_{V^\perp}(x) + \mathcal{P}_V(u)\big) : (x, u) \in \mathrm{gra}(A) \big\}.$$

Example Let $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, then $A_{\mathbb{R}^n} = A$ and $A_{\{0\}} = A^{-1}$.

An application of Proximal Point Algorithm.

Spingarn

Let $x_0 \in V,\ u_0 \in V^\perp$:

$$\begin{aligned} y_k &= \mathcal{J}_A(x_k + u_k), \\ v_k &= x_k + u_k - y_k, \\ (x_{k+1}, u_{k+1}) &= \big(\mathcal{P}_V(y_k), \mathcal{P}_{V^\perp}(v_k)\big). \end{aligned}$$

Fixed-point characterisation

- define mapping

$$L : \mathbb{R}^n \oplus \mathbb{R}^n \to \mathbb{R}^n \oplus \mathbb{R}^n : (x, u) \to \big(\mathcal{P}_V(x) + \mathcal{P}_{V^\perp}(u), \mathcal{P}_{V^\perp}(x) + \mathcal{P}_V(u)\big).$$

- $$\begin{aligned} p = \mathcal{J}_{A_V}(x) &\iff (p, x-p) \in \mathrm{gra}(A_V) \\ &\iff L(p, x-p) \in L\big(\mathrm{gra}(A_V)\big) = \mathrm{gra}(A) \\ &\iff \big(\mathcal{P}_V(p) + \mathcal{P}_{V^\perp}(x-p), \mathcal{P}_V(x-p) + \mathcal{P}_{V^\perp}(p)\big) \in \mathrm{gra}(A). \end{aligned}$$

- let $q = \mathcal{P}_V(p) + \mathcal{P}_{V^\perp}(x-p)$

$$\begin{aligned} p = \mathcal{J}_{A_V}(x) &\iff x - q = \mathcal{P}_V(x-p) + \mathcal{P}_{V^\perp}p \in A(q) \\ &\iff q = \mathcal{J}_A(x). \end{aligned}$$

- let $z_k = x_k + u_k$, since $x_k \in V$ and $u_k \in V^\perp$

$$\begin{aligned} \mathcal{P}_V(z_{k+1}) + \mathcal{P}_{V^\perp}(z_k - z_{k+1}) &= x_{k+1} + \mathcal{P}_{V^\perp}(u_k) - u_{k+1} \\ &= \mathcal{P}_V(y_k) + \mathcal{P}_{V^\perp}(v_k - x_k + y_k) - \mathcal{P}_{V^\perp}(v_k) \\ &= \mathcal{P}_V(y_k) + \mathcal{P}_{V^\perp}(v_k) + \mathcal{P}_{V^\perp}(y_k) - \mathcal{P}_{V^\perp}(v_k). \end{aligned}$$

- $z_{k+1} = \mathcal{J}_A(z_k)$.

Outline

An inertial DR splitting

Initial : $x_0 \in \mathbb{R}^n, x_{-1} = x_0$ and $\gamma > 0$;

$$
\begin{aligned}
y_k &= z_k + a_{0,k}(z_k - z_{k-1}) + a_{1,k}(z_{k-1} - z_{k-2}) + \cdots ,\\
z_{k+1} &= \mathcal{T}_{\mathrm{DR}}(y_k)
\end{aligned}
$$

- relaxation can be applied.

Outline

Basis pursuit

$$\min_{x \in \mathbb{R}^n} \quad \|x\|_1 \qquad \text{such that} \quad Ax = b,$$

- $A : \mathbb{R}^n \to \mathbb{R}^m$ with $m << n$.
- $b \in \mathrm{Img}(A)$.

Image inpainting

$$\min_{X \in \mathbb{R}^{n\times n}} \quad \|WX\|_1$$
$$\text{such that} \quad \mathcal{P}_\Omega(X) = \bar{X},$$

- W: total variation, orthonomal basis, redundant wavelet frame.
- Observation constraint

$$\big(\mathcal{P}_\Omega(X)\big)_{i,j} = \begin{cases} \bar{X}_{i,j} : (i,j) \in \Omega, \\ \quad 0 : (i,j) \notin \Omega. \end{cases}$$

- Painting reconstruction in museum.

Matrix completion

$$\min_{X\in\mathbb{R}^{n\times n}} \quad \|X\|_*$$
$$\text{such that} \quad \mathcal{P}_\Omega(X) = \bar{X},$$

- Observation constraint

$$\big(\mathcal{P}_\Omega(X)\big)_{i,j} = \begin{cases} \bar{X}_{i,j} : (i,j) \in \Omega, \\ \quad 0 : (i,j) \notin \Omega. \end{cases}$$

- Netflix prize, recommendation system.

Variation ineuality

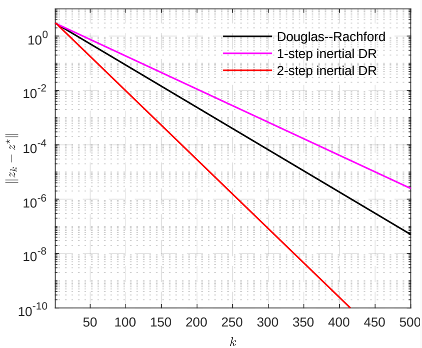
$$\text{Find } x \in \mathbb{R}^n \text{ such that } \quad \exists u \in A(x), \forall y \in \mathbb{R}^n : \ \langle x - y, \, u \rangle + R(x) \leq R(y).$$

- $R \in \Gamma_0$.
- $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is maximal monotone.

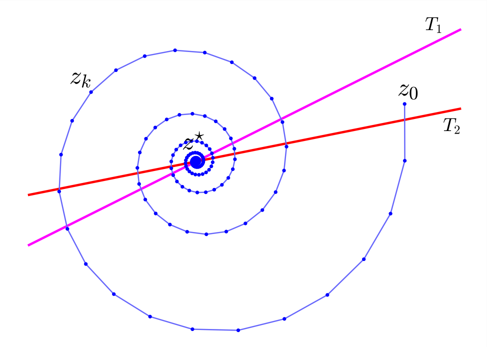
Example Let $R, J \in \Gamma_0$, and $x^\star \in \mathrm{Argmin}(R + J)$, then $\exists u \in \partial J(x^\star)$ s.t. $-u \in \partial R(x^\star)$ and

$$\begin{aligned} & \langle y - x^\star, \, -u \rangle + R(x^\star) \leq R(y) \\ \iff \ & \langle x^\star - y, \, u \rangle + R(x^\star) \leq R(y). \end{aligned}$$

Numerical experiment



Comparison



Tracjectory

- H. H. Bauschke, J. Y. Bello Cruz, T. T. A. Nghia, H. M. Pha, and X. Wang. “Optimal rates of linear convergence of relaxed alternating projections and generalized Douglas–Rachford methods for two subspaces”. Numerical Algorithms, 73(1):33–76, 2016.
- H. Bauschke and P. L. Combettes. “Convex Analysis and Monotone Operator Theory in Hilbert Spaces”. Springer, 2011.
- J. Liang. “Convergence rates of first-order operator splitting methods”. Diss. Normandie Université; GREYC CNRS UMR 6072, 2016.