On the Douglas–Rachford operator in the (possibly) inconsistent case and related progress

Speaker: Walaa M. Moursi

University of British Columbia, University of Calgary, Canada
& Simons Institute for Theory and Computing, UC Berkeley, USA

Joint work with

Heinz H. Bauschke
University of British Columbia

Minh N. Dao
University of Newcastle

Dedicated to the memory of Jonathan Borwein

Workshop on Splitting Algorithms, Modern Operator Theory and Applications
September 17–22, 2017, Oaxaca, Mexico

Monotone operators

Throughout this talk

X is a real Hilbert space

with inner product $\langle \cdot, \cdot \rangle$, and induced norm $\|\cdot\|$.
Recall that an operator $A\colon X \rightrightarrows X$ is monotone if

$$(x, u), (y, v) \in \operatorname{gr} A \Rightarrow \langle x - y, u - v \rangle \geq 0.$$

Recall also that a monotone operator A is maximally monotone if A cannot be properly extended without destroying monotonicity.
In the following we assume that

A and B are maximally monotone operators on X.

The problem:
Find $x \in X$ such that

$$x \in \operatorname{zer}(A + B) := (A + B)^{-1}(0).$$

Connection to optimization

The problem:
Find $x \in X$ such that

$$x \in \operatorname{zer}(A + B) := (A + B)^{-1}(0).$$

When A and B are subdifferential operators we recover the setting of many optimization problems,

---

We shall use ι_U and N_U to denote the indicator function and the normal cone operator of a nonempty closed convex subset U of X.

Connection to optimization

The problem:
Find $x \in X$ such that

$$x \in \operatorname{zer}(A+B) := (A+B)^{-1}(0).$$

When A and B are subdifferential operators we recover the setting of many optimization problems, for instance:

- Choosing $A = \partial f$ and $B = \partial \iota_C = N_C$, the sum problem reduces to solving the constrained convex optimization:

$$\left.\begin{array}{l}\text{minimize } f(x)\\ \text{subject to } x \in C\end{array}\right\} \longrightarrow \text{find } x \in X \text{ such that } 0 \in (\partial f + N_C)x.$$

We shall use ι_U and N_U to denote the indicator function and the normal cone operator of a nonempty closed convex subset U of X.

Connection to optimization

The problem:
Find $x \in X$ such that

$$x \in \operatorname{zer}(A+B) := (A+B)^{-1}(0).$$

When A and B are subdifferential operators we recover the setting of many optimization problems, for instance:

- Choosing $A = \partial f$ and $B = \partial \iota_C = N_C$, the sum problem reduces to solving the constrained convex optimization:
  $$\left.\begin{array}{l}\text{minimize } f(x)\\ \text{subject to } x \in C\end{array}\right\} \longrightarrow \text{find } x \in X \text{ such that } 0 \in (\partial f + N_C)x.$$
- Choosing $A = \partial \iota_U = N_U$ and $B = \partial \iota_V = N_V$, the sum problem reduces to solving the convex feasibility problems:
  $$\text{find } x \text{ such } x \in U \cap V \longrightarrow \text{find } x \in X \text{ such that } 0 \in (N_U + N_V)x.$$

We shall use ι_U and N_U to denote the indicator function and the normal cone operator of a nonempty closed convex subset U of X.

Firmly nonexpansive operators and resolvents

Definition (resolvent and reflected resolvent)

The resolvent and the reflected resolvent of A are the operators

$$J_A := (\mathrm{Id} + A)^{-1}, \quad R_A := 2J_A - \mathrm{Id}.$$

Let $T : X \to X$. Then T is nonexpansive if $\|Tx - Ty\| \leq \|x - y\|$.
T is firmly nonexpansive if $\|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \leq \|x - y\|^2$.

Firmly nonexpansive operators and resolvents

Definition (resolvent and reflected resolvent)

The resolvent and the reflected resolvent of A are the operators

$$J_A := (\mathrm{Id} + A)^{-1}, \quad R_A := 2J_A - \mathrm{Id}\,.$$

Example

- Let $f : X \to \left]-\infty, +\infty\right]$ be proper lower semicontinuous convex function. Let $A := \partial f \Rightarrow J_A = (\mathrm{Id} + \partial f)^{-1} = \mathrm{Prox}_f$, where Prox_f is the Moreau prox operator of the function f.

Let $T : X \to X$. Then T is nonexpansive if $\|Tx - Ty\| \leq \|x - y\|$.
T is firmly nonexpansive if $\|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \leq \|x - y\|^2$.

Firmly nonexpansive operators and resolvents

Definition (resolvent and reflected resolvent)

The resolvent and the reflected resolvent of A are the operators

$$J_A := (\mathrm{Id} + A)^{-1}, \quad R_A := 2J_A - \mathrm{Id}\,.$$

Example

- Let $f : X \to \left]-\infty, +\infty\right]$ be proper lower semicontinuous convex function. Let $A := \partial f \Rightarrow J_A = (\mathrm{Id} + \partial f)^{-1} = \mathrm{Prox}_f$, where Prox_f is the Moreau prox operator of the function f.
- Suppose that U is a nonempty closed convex subset of X. Let $A := N_U \Rightarrow J_A = (\mathrm{Id} + N_U)^{-1} = \mathrm{Prox}_{\iota_U} = P_U$.

Let $T : X \to X$. Then T is nonexpansive if $\|Tx - Ty\| \leq \|x - y\|$.
T is firmly nonexpansive if $\|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \leq \|x - y\|^2$.

Firmly nonexpansive operators and resolvents

Definition (resolvent and reflected resolvent)

The resolvent and the reflected resolvent of A are the operators

$$J_A := (\mathrm{Id} + A)^{-1}, \quad R_A := 2J_A - \mathrm{Id}\,.$$

Example

- Let $f : X \to \left]-\infty, +\infty\right]$ be proper lower semicontinuous convex function. Let $A := \partial f \Rightarrow J_A = (\mathrm{Id} + \partial f)^{-1} = \mathrm{Prox}_f$, where Prox_f is the Moreau prox operator of the function f.
- Suppose that U is a nonempty closed convex subset of X. Let $A := N_U \Rightarrow J_A = (\mathrm{Id} + N_U)^{-1} = \mathrm{Prox}_{\iota_U} = P_U$.

Fact

J_A is firmly nonexpansive and R_A is nonexpansive.

Let $T : X \to X$. Then T is nonexpansive if $\|Tx - Ty\| \leq \|x - y\|$.
T is firmly nonexpansive if $\|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \leq \|x - y\|^2$.

The Douglas–Rachford splitting operator

The Douglas–Rachford splitting operator associated with the ordered pair (A, B) is

$$T := T_{A,B} := \mathrm{Id} - J_A + J_B R_A = \tfrac{1}{2}(\mathrm{Id} + R_B R_A).$$

The resolvent and the reflected resolvent of A are the operators $J_A := (\mathrm{Id} + A)^{-1}$ and $R_A := 2J_A - \mathrm{Id}$, respectively.
$\mathrm{zer}(A + B) = (A + B)^{-1}(0)$.

The Douglas–Rachford splitting operator

The Douglas–Rachford splitting operator associated with the ordered pair (A, B) is

$$T := T_{A,B} := \mathrm{Id} - J_A + J_B R_A = \frac{1}{2}(\mathrm{Id} + R_B R_A).$$

- T is firmly nonexpansive.

The resolvent and the reflected resolvent of A are the operators $J_A := (\mathrm{Id} + A)^{-1}$ and $R_A := 2J_A - \mathrm{Id}$, respectively.
$\mathrm{zer}(A + B) = (A + B)^{-1}(0)$.

The Douglas–Rachford splitting operator

The Douglas–Rachford splitting operator associated with the ordered pair (A, B) is

$$T := T_{A,B} := \operatorname{Id} - J_A + J_B R_A = \tfrac{1}{2}(\operatorname{Id} + R_B R_A).$$

- T is firmly nonexpansive.
- Thanks to Combettes, we know

$$J_A(\operatorname{Fix} T) = \operatorname{zer}(A + B).$$

The resolvent and the reflected resolvent of A are the operators $J_A := (\operatorname{Id} + A)^{-1}$ and $R_A := 2J_A - \operatorname{Id}$, respectively.
$\operatorname{zer}(A + B) = (A + B)^{-1}(0)$.

Known results

Suppose that

$$\operatorname{zer}(A+B) := \{x \in X \mid 0 \in Ax + Bx\} \neq \varnothing.$$

Known results

Suppose that

$$\operatorname{zer}(A+B) := \{x \in X \mid 0 \in Ax + Bx\} \neq \varnothing.$$

- Combettes (2004) $J_A(\operatorname{Fix} T) = \operatorname{zer}(A+B)$. Consequently,

$$\operatorname{Fix} T \neq \varnothing \Leftrightarrow \operatorname{zer}(A+B) \neq \varnothing.$$

Known results

Suppose that

$$\operatorname{zer}(A+B) := \{x \in X \mid 0 \in Ax + Bx\} \neq \varnothing.$$

- Combettes (2004) $J_A(\operatorname{Fix} T) = \operatorname{zer}(A+B)$. Consequently,

$$\operatorname{Fix} T \neq \varnothing \Leftrightarrow \operatorname{zer}(A+B) \neq \varnothing.$$

- Krasnosel'skiĭ–Mann (1950s)

$$T^n x \xrightarrow[\textit{weakly}]{} \text{some point in } \operatorname{Fix} T \neq \operatorname{zer}(A+B).$$

Known results

Suppose that

$$\operatorname{zer}(A+B) := \{x \in X \mid 0 \in Ax + Bx\} \neq \varnothing.$$

- Combettes (2004) $J_A(\operatorname{Fix} T) = \operatorname{zer}(A+B)$. Consequently,

$$\operatorname{Fix} T \neq \varnothing \Leftrightarrow \operatorname{zer}(A+B) \neq \varnothing.$$

- Krasnosel'skiĭ–Mann (1950s)

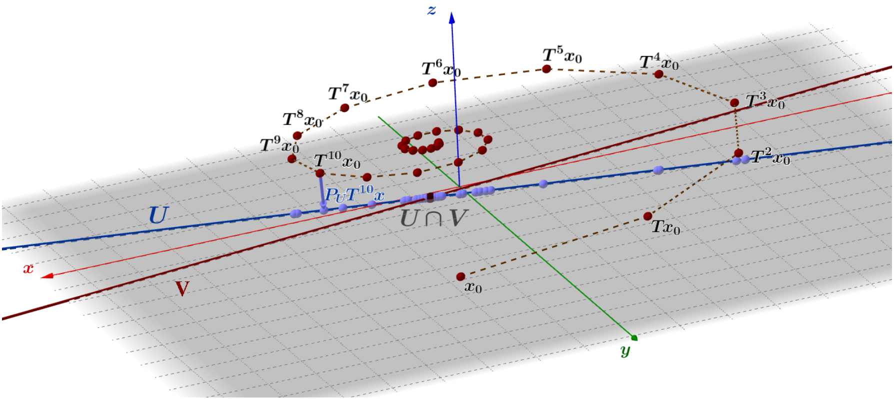
$$T^n x \xrightarrow[\text{weakly}]{} \text{some point in } \operatorname{Fix} T \neq \operatorname{zer}(A+B).$$

- Lions–Mercier (1979) and Svaiter (2011)

$$J_A T^n x \xrightarrow[\text{weakly}]{} \text{some point in } \operatorname{zer}(A+B).$$

DR for two lines in $\mathbb{R}^3$

$$A = N_U,\ B = N_V \text{ and } T = \mathrm{Id} - P_U + P_V(2P_U - \mathrm{Id}).$$



U = the blue line,
V = the red line,
$(T^n x_0)_{n\in\mathbb{N}}$ = the red sequence,
$(P_U T^n x_0)_{n\in\mathbb{N}}$ = the blue sequence.

Motivation

Recall that when

$$\operatorname{zer}(A+B) \neq \varnothing$$

we have:

- Combettes (2004) $J_A(\operatorname{Fix} T) = \operatorname{zer}(A+B)$. Consequently,

$$\operatorname{Fix} T \neq \varnothing \Leftrightarrow \operatorname{zer}(A+B) \neq \varnothing.$$

- Krasnosel'skiĭ–Mann (1950s)

$$T^n x \xrightarrow[\text{weakly}]{} \text{some point in } \operatorname{Fix} T \neq \operatorname{zer}(A+B).$$

- Lions–Mercier (1979) and Svaiter (2011)

$$J_A T^n x \xrightarrow[\text{weakly}]{} \text{some point in } \operatorname{zer}(A+B).$$

Motivation

Recall that when

$$\operatorname{zer}(A+B) \neq \varnothing$$

we have:

- Combettes (2004) $J_A(\operatorname{Fix} T) = \operatorname{zer}(A+B)$. Consequently,

$$\operatorname{Fix} T \neq \varnothing \Leftrightarrow \operatorname{zer}(A+B) \neq \varnothing.$$

- Krasnosel'skiĭ–Mann (1950s)

$$T^n x \xrightarrow[\text{weakly}]{} \text{some point in } \operatorname{Fix} T \neq \operatorname{zer}(A+B).$$

- Lions–Mercier (1979) and Svaiter (2011)

$$J_A T^n x \xrightarrow[\text{weakly}]{} \text{some point in } \operatorname{zer}(A+B).$$

- Question: What happens when $\operatorname{zer}(A+B) = \varnothing$?

Inconsistent feasibility problem

Not every sum problem admits a solution:

- Suppose that U and V are nonempty closed convex subsets of X such that $U \cap V = \varnothing$.
- Set $A := N_U$ and $B := N_V$.
- Then $\operatorname{zer}(A+B) = (A+B)^{-1}(0) = U \cap V = \varnothing$.
- By an earlier fact[1] we have $\operatorname{zer}(A+B) = \varnothing \Leftrightarrow \operatorname{Fix} T = \varnothing$.

[1] Fact (Combettes (2004)): $J_A(\operatorname{Fix} T_{A,B}) = \operatorname{zer}(A+B)$.

The w-perturbed problem

Let $w \in X$ and $x \in X$. The corresponding inner and outer perturbations of A are

$$A_w x := A(x - w) \text{ and } {}_wAx := Ax - w.$$

The w-perturbed problem

Let $w \in X$ and $x \in X$. The corresponding inner and outer perturbations of A are

$$A_w x := A(x - w) \text{ and } {}_w A x := Ax - w.$$

The w-perturbed problem associated with (A, B) is to find a point in the set of zeros

$$\begin{aligned} Z_w &:= \operatorname{zer}\left({}_w A, B_w\right) = \left({}_w A + B_w\right)^{-1}(0) \\ &= \left\{x \in X \;\middle|\; w \in Ax + B(x - w)\right\}. \end{aligned}$$

The w-perturbed problem

Let $w \in X$ and $x \in X$. The corresponding inner and outer perturbations of A are

$$A_w x := A(x - w) \text{ and } {}_w A x := Ax - w.$$

The w-perturbed problem associated with (A, B) is to find a point in the set of zeros

$$\begin{aligned} Z_w &:= \operatorname{zer}\left({}_w A, B_w\right) = \left({}_w A + B_w\right)^{-1}(0) \\ &= \left\{x \in X \;\middle|\; w \in Ax + B(x - w)\right\}. \end{aligned}$$

Proposition

$$Z_w \neq \varnothing \Leftrightarrow w \in \operatorname{ran}(\operatorname{Id} - T).$$

The w-perturbed problem

Let $w \in X$ and $x \in X$. The corresponding inner and outer perturbations of A are

$$A_w x := A(x - w) \text{ and } {}_wAx := Ax - w.$$

The w-perturbed problem associated with (A, B) is to find a point in the set of zeros

$$\begin{aligned} Z_w &:= \operatorname{zer}({}_wA, B_w) = \left({}_wA + B_w\right)^{-1}(0) \\ &= \left\{x \in X \;\middle|\; w \in Ax + B(x - w)\right\}. \end{aligned}$$

Proposition

$$Z_w \neq \varnothing \Leftrightarrow w \in \operatorname{ran}(\mathrm{Id} - T).$$

Corollary

$$\left\{w \in X \;\middle|\; Z_w \neq \varnothing\right\} = \operatorname{ran}(\mathrm{Id} - T).$$

The normal problem: Definition

The normal problem associated with (A, B) is to find a point in the set of zeros

$$Z_v := \operatorname{zer}\left({}_vA, B_v\right) = \left({}_vA + B_v\right)^{-1}(0) = \left\{x \in X \mid v \in Ax + B(x - v)\right\}.$$

where

$$v := v_{(A,B)} := P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}(0)$$

is the minimal displacement vector of (A, B) and the set of normal solutions is $Z_v = Z_{v_{(A,B)}}$.

$T_v = T(\cdot - v)$.

The normal problem: Definition

The normal problem associated with (A, B) is to find a point in the set of zeros

$$Z_v := \operatorname{zer}\left({}_vA, B_v\right) = \left({}_vA + B_v\right)^{-1}(0) = \left\{x \in X \;\middle|\; v \in Ax + B(x - v)\right\}.$$

where

$$v := v_{(A,B)} := P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}(0)$$

is the minimal displacement vector of (A, B) and the set of normal solutions is $Z_v = Z_{v_{(A,B)}}$.

- The normal problem is well defined: Indeed, $\operatorname{Id} - T$ is maximally monotone and consequently $\overline{\operatorname{ran}}(\operatorname{Id} - T)$ (Fact) is closed and convex.

$T_v = T(\cdot - v)$.

The normal problem: Definition

The normal problem associated with (A, B) is to find a point in the set of zeros

$$Z_v := \operatorname{zer}\left({}_vA, B_v\right) = \left({}_vA + B_v\right)^{-1}(0) = \left\{x \in X \;\middle|\; v \in Ax + B(x - v)\right\}.$$

where

$$v := v_{(A,B)} := P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}(0)$$

is the minimal displacement vector of (A, B) and the set of normal solutions is $Z_v = Z_{v_{(A,B)}}$.

- The normal problem is well defined: Indeed, $\operatorname{Id} - T$ is maximally monotone and consequently $\overline{\operatorname{ran}}(\operatorname{Id} - T)$ (Fact) is closed and convex.
- $T_{({}_vA, B_v)} = T_{-v} = T(\cdot + v)$.

$T_v = T(\cdot - v)$.

The normal problem: Definition

The normal problem associated with (A, B) is to find a point in the set of zeros

$$Z_v := \operatorname{zer}({}_vA, B_v) = ({}_vA + B_v)^{-1}(0) = \{x \in X \mid v \in Ax + B(x - v)\}.$$

where

$$v := v_{(A,B)} := P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}(0)$$

is the minimal displacement vector of (A, B) and the set of normal solutions is $Z_v = Z_{v_{(A,B)}}$.

- The normal problem is well defined: Indeed, $\operatorname{Id} - T$ is maximally monotone and consequently $\overline{\operatorname{ran}}(\operatorname{Id} - T)$ (Fact) is closed and convex.
- $T_{({}_vA, B_v)} = T_{-v} = T(\cdot + v)$.
- If $(A, B) = (\partial\iota_U, \partial\iota_V) = (N_U, N_V)$ then

$$v = P_{\overline{U-V}}(0) \text{ and } Z_v = U \cap (v + V).$$

$T_v = T(\cdot - v)$.

Motivation

Recall that $U \cap V$ could be possibly empty. We recall also that

$$v := P_{\overline{U-V}}(0) = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}(0).$$

In the following we assume that

$$v \in \operatorname{ran}(\operatorname{Id} - T).$$

$T_{-v} = T(\cdot + v)$.

Motivation

Recall that $U \cap V$ could be possibly empty. We recall also that

$$v := P_{\overline{U-V}}(0) = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}(0).$$

In the following we assume that

$$v \in \operatorname{ran}(\operatorname{Id} - T).$$

So far we have:

- $(\forall x \in X)\ T^n x - T^{n+1} x \to v$. (Fact)

$T_{-v} = T(\cdot + v)$.

Motivation

Recall that $U \cap V$ could be possibly empty. We recall also that

$$v := P_{\overline{U-V}}(0) = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}(0).$$

In the following we assume that

$$v \in \operatorname{ran}(\operatorname{Id} - T).$$

So far we have:

- $(\forall x \in X)\ T^n x - T^{n+1} x \to v$. (Fact)
- $(\forall x \in X)\ P_U((T_{-v})^n x + v) \to$ a best approximation solution. (Fact)

$T_{-v} = T(\cdot + v)$.

Motivation

Recall that $U \cap V$ could be possibly empty. We recall also that

$$v := P_{\overline{U-V}}(0) = P_{\overline{\text{ran}}(\text{Id} - T)}(0).$$

In the following we assume that

$$v \in \text{ran}(\text{Id} - T).$$

So far we have:

- $(\forall x \in X)\ T^n x - T^{n+1} x \to v$. (Fact)
- $(\forall x \in X)\ P_U((T_{-v})^n x + v) \to$ a best approximation solution. (Fact)

Question: Can we come up with one algorithm that finds a best approximation solution and the gap vector (or even just a best approximation solution)?

$T_{-v} = T(\cdot + v)$.

Motivation

Fact (Pazy (1970))

Suppose that $T : X \to X$ *is nonexpansive such that* $\operatorname{Fix} T = \varnothing$. *Then* $(\forall x \in X)\ \|T^n x\| \to +\infty$.

Let $T : X \to X$. Then T is nonexpansive if $\|Tx - Ty\| \leq \|x - y\|$.

Motivation

Fact (Pazy (1970))

Suppose that $T : X \to X$ *is nonexpansive such that* $\operatorname{Fix} T = \varnothing$*. Then* $(\forall x \in X)\ \|T^n x\| \to +\infty$*.*

Fact (Bauschke-Combettes-Luke (2004))

Suppose that U *and* V *are nonempty closed convex subsets of* X *such that* $U \cap V = \varnothing$*. Then* $(\forall x \in X)$ *the shadow sequence* $(P_U T^n x)_{n\in\mathbb{N}}$ *is bounded and its weak cluster points lie in* $U \cap (v + V)$*, hence are best approximation solutions.*

Let $T : X \to X$. Then T is nonexpansive if $\|Tx - Ty\| \le \|x - y\|$.

The case of infeasible affine subspaces: Example

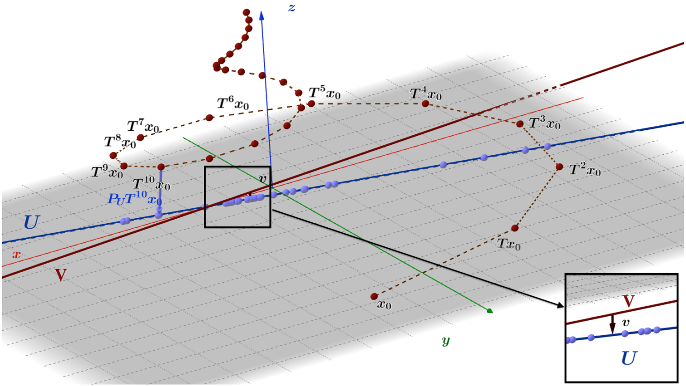


Figure: A GeoGebra snapshot. Two nonintersecting affine subspaces U (blue line) and V (red line) in $\mathbb{R}^3$. Shown are also the first few iterates of $(T^n x_0)_{n\in\mathbb{N}}$ (red points) and $(P_U T^n x_0)_{n\in\mathbb{N}}$ (blue points).

New useful identities

Let $(a, b, z) \in X^3$. Then

$$\|z\|^2 = \|z - a + b\|^2 + \|a - b\|^2 + 2\langle a, z - a\rangle + 2\langle b, 2a - z - b\rangle.$$

New useful identities

Let $(a, b, z) \in X^3$. Then

$$\|z\|^2 = \|z - a + b\|^2 + \|a - b\|^2 + 2\langle a, z - a\rangle + 2\langle b, 2a - z - b\rangle.$$

Theorem

Let $x \in X$ and let $y \in X$. Then

$$\begin{aligned}\|x - y\|^2 = \|Tx - Ty\|^2 &+ \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \\ &+ 2\langle J_A x - J_A y, J_{A^{-1}} x - J_{A^{-1}} y\rangle \\ &+ 2\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle.\end{aligned}$$

Proof.

Apply the above identity with (a, b, z) replaced by $(J_A x - J_A y, J_B R_A x - J_B R_A y, x - y)$ and use that $T = \mathrm{Id} - J_A + J_B R_A$. □

New useful identities

Theorem

Let $x \in X$ and let $y \in X$. Then

$$\begin{aligned}\|x-y\|^2 = \|Tx - Ty\|^2 + \|(\mathrm{Id}-T)x - (\mathrm{Id}-T)y\|^2 \\ + 2\,\langle J_A x - J_A y, J_{A^{-1}} x - J_{A^{-1}} y\rangle \\ \\ + 2\,\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle\end{aligned}$$

New useful identities

Theorem

Let $x \in X$ and let $y \in X$. Then

$$\|x - y\|^2 = \|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 + 2\left\langle J_A x - J_A y, J_{A^{-1}} x - J_{A^{-1}} y \right\rangle$$

$$+ 2\left\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y \right\rangle$$

New useful identities

Theorem

Let $x \in X$ and let $y \in X$. Then

$$\|x-y\|^2 = \|Tx - Ty\|^2 + \|(\mathrm{Id}-T)x - (\mathrm{Id}-T)y\|^2 + 2\underbrace{\langle J_A x - J_A y, J_{A^{-1}} x - J_{A^{-1}} y\rangle}_{\geq 0} + 2\underbrace{\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle}_{\geq 0}.$$

New useful identities

Theorem

Let $x \in X$ and let $y \in X$. Then

$$\|x - y\|^2 = \|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 + 2\underbrace{\langle J_A x - J_A y, J_{A^{-1}} x - J_{A^{-1}} y\rangle}_{\geq 0} + 2\underbrace{\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle}_{\geq 0}.$$

Corollary

Let $x \in X$ and let $y \in X$. Then the following hold:

$$(\mathrm{Id} - T)T^n x - (\mathrm{Id} - T)T^n y \to 0,$$
$$\langle J_A T^n x - J_A T^n y, J_{A^{-1}} T^n x - J_{A^{-1}} T^n y\rangle \to 0,$$
$$\langle J_B R_A T^n x - J_B R_A T^n y, J_{B^{-1}} R_A T^n x - J_{B^{-1}} R_A T^n y\rangle \to 0.$$

Proof.

This follows from the above theorem by telescoping. □

New Fejér monotonicity principle

Lemma

Suppose that

- E *is a nonempty closed convex subset of* X*,*

New Fejér monotonicity principle

Lemma

Suppose that

- E *is a nonempty closed convex subset of* X,
- $(x_n)_{n\in\mathbb{N}}$ *is a sequence in* X *that is Fejér monotone with respect to* E, *i.e.,*

$$(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \leq \|x_n - e\|,$$

- $(u_n)_{n\in\mathbb{N}}$ *is a bounded sequence in* X *such that its weak cluster points lie in* E,

New Fejér monotonicity principle

Lemma

Suppose that

- E *is a nonempty closed convex subset of* X,
- $(x_n)_{n\in\mathbb{N}}$ *is a sequence in* X *that is Fejér monotone with respect to* E*, i.e.,*

$$(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \leq \|x_n - e\|,$$

- $(u_n)_{n\in\mathbb{N}}$ *is a bounded sequence in* X *such that its weak cluster points lie in* E,
- *and*

$$(\forall e \in E) \ \langle u_n - e, u_n - x_n \rangle \to 0.$$

New Fejér monotonicity principle

Lemma

Suppose that

- E *is a nonempty closed convex subset of* X,
- $(x_n)_{n\in\mathbb{N}}$ *is a sequence in* X *that is Fejér monotone with respect to* E, *i.e.,*
$$(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \leq \|x_n - e\|,$$
- $(u_n)_{n\in\mathbb{N}}$ *is a bounded sequence in* X *such that its weak cluster points lie in* E,
- *and*
$$(\forall e \in E) \ \ \langle u_n - e, u_n - x_n \rangle \to 0.$$

Then $(u_n)_{n\in\mathbb{N}}$ *converges weakly to some point in* E.

New Fejér monotonicity principle

Lemma

Suppose that

- E *is a nonempty closed convex subset of* X,
- $(x_n)_{n\in\mathbb{N}}$ *is a sequence in* X *that is Fejér monotone with respect to* E, *i.e.,*

$$(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \leq \|x_n - e\|,$$

- $(u_n)_{n\in\mathbb{N}}$ *is a bounded sequence in* X *such that its weak cluster points lie in* E,
- *and*

$$(\forall e \in E) \ \langle u_n - e, u_n - x_n \rangle \to 0.$$

Then $(u_n)_{n\in\mathbb{N}}$ *converges weakly to some point in* E.

Remark

$(x_n)_{n\in\mathbb{N}} = (u_n)_{n\in\mathbb{N}} \Rightarrow$ we recover the classical Fejér monotonicity principle!

New Fejér monotonicity principle: proof

- Step 1: $(\forall (e_1, e_2) \in E \times E) \quad \langle e_2 - e_1, u_n - x_n \rangle = \langle u_n - e_1, u_n - x_n \rangle - \langle u_n - e_2, u_n - x_n \rangle \to 0$.

Lemma: Suppose that E is a nonempty closed convex subset of X, that $(x_n)_{n \in \mathbb{N}}$ is a sequence in X that is *Fejér monotone with respect to E*, i.e., $(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \leq \|x_n - e\|$, that $(u_n)_{n \in \mathbb{N}}$ is a bounded sequence in X such that its weak cluster points lie in E, and that $(\forall e \in E) \ \langle u_n - e, x_n - u_n \rangle \to 0$. Then $(u_n)_{n \in \mathbb{N}}$ converges weakly to some point in E.

New Fejér monotonicity principle: proof

- Step 1: $(\forall (e_1, e_2) \in E \times E) \quad \langle e_2 - e_1, u_n - x_n \rangle = \langle u_n - e_1, u_n - x_n \rangle - \langle u_n - e_2, u_n - x_n \rangle \to 0$.
- Step 2: Obtain four subsequences $(x_{k_n})_{n\in\mathbb{N}}$, $(x_{l_n})_{n\in\mathbb{N}}$, $(u_{k_n})_{n\in\mathbb{N}}$ and $(u_{l_n})_{n\in\mathbb{N}}$ such that $x_{k_n} \rightharpoonup \bar{x}_1$, $x_{l_n} \rightharpoonup \bar{x}_2$, $u_{k_n} \rightharpoonup e_1$ and $u_{l_n} \rightharpoonup e_2$.

Lemma: Suppose that E is a nonempty closed convex subset of X, that $(x_n)_{n\in\mathbb{N}}$ is a sequence in X that is *Fejér monotone with respect to E*, i.e., $(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \leq \|x_n - e\|$, that $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in X such that its weak cluster points lie in E, and that $(\forall e \in E)\ \langle u_n - e, x_n - u_n \rangle \to 0$. Then $(u_n)_{n\in\mathbb{N}}$ converges weakly to some point in E.

New Fejér monotonicity principle: proof

- Step 1: $(\forall (e_1, e_2) \in E \times E) \quad \langle e_2 - e_1, u_n - x_n \rangle = \langle u_n - e_1, u_n - x_n \rangle - \langle u_n - e_2, u_n - x_n \rangle \to 0$.
- Step 2: Obtain four subsequences $(x_{k_n})_{n\in\mathbb{N}}$, $(x_{l_n})_{n\in\mathbb{N}}$, $(u_{k_n})_{n\in\mathbb{N}}$ and $(u_{l_n})_{n\in\mathbb{N}}$ such that $x_{k_n} \rightharpoonup \bar{x}_1$, $x_{l_n} \rightharpoonup \bar{x}_2$, $u_{k_n} \rightharpoonup e_1$ and $u_{l_n} \rightharpoonup e_2$. Taking the limit in view of Step 1 along these subsequences we have $\langle e_2 - e_1, e_1 - \bar{x}_1 \rangle = 0 = \langle e_2 - e_1, e_2 - \bar{x}_2 \rangle$

Lemma: Suppose that E is a nonempty closed convex subset of X, that $(x_n)_{n\in\mathbb{N}}$ is a sequence in X that is *Fejér monotone with respect to E*, i.e., $(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \leq \|x_n - e\|$, that $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in X such that its weak cluster points lie in E, and that $(\forall e \in E)\ \langle u_n - e, x_n - u_n \rangle \to 0$. Then $(u_n)_{n\in\mathbb{N}}$ converges weakly to some point in E.

New Fejér monotonicity principle: proof

- Step 1: $(\forall (e_1, e_2) \in E \times E) \quad \langle e_2 - e_1, u_n - x_n \rangle = \langle u_n - e_1, u_n - x_n \rangle - \langle u_n - e_2, u_n - x_n \rangle \to 0$.
- Step 2: Obtain four subsequences $(x_{k_n})_{n\in\mathbb{N}}$, $(x_{l_n})_{n\in\mathbb{N}}$, $(u_{k_n})_{n\in\mathbb{N}}$ and $(u_{l_n})_{n\in\mathbb{N}}$ such that $x_{k_n} \rightharpoonup \bar{x}_1$, $x_{l_n} \rightharpoonup \bar{x}_2$, $u_{k_n} \rightharpoonup e_1$ and $u_{l_n} \rightharpoonup e_2$. Taking the limit in view of Step 1 along these subsequences we have $\langle e_2 - e_1, e_1 - \bar{x}_1 \rangle = 0 = \langle e_2 - e_1, e_2 - \bar{x}_2 \rangle$ hence $\|e_2 - e_1\|^2 = \langle e_2 - e_1, \bar{x}_2 - \bar{x}_1 \rangle$.

Lemma: Suppose that E is a nonempty closed convex subset of X, that $(x_n)_{n\in\mathbb{N}}$ is a sequence in X that is *Fejér monotone with respect to E*, i.e., $(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \leq \|x_n - e\|$, that $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in X such that its weak cluster points lie in E, and that $(\forall e \in E)\ \langle u_n - e, x_n - u_n \rangle \to 0$. Then $(u_n)_{n\in\mathbb{N}}$ converges weakly to some point in E.

New Fejér monotonicity principle: proof

- Step 1: $(\forall (e_1, e_2) \in E \times E) \quad \langle e_2 - e_1, u_n - x_n \rangle = \langle u_n - e_1, u_n - x_n \rangle - \langle u_n - e_2, u_n - x_n \rangle \to 0$.
- Step 2: Obtain four subsequences $(x_{k_n})_{n\in\mathbb{N}}$, $(x_{l_n})_{n\in\mathbb{N}}$, $(u_{k_n})_{n\in\mathbb{N}}$ and $(u_{l_n})_{n\in\mathbb{N}}$ such that $x_{k_n} \rightharpoonup \bar{x}_1$, $x_{l_n} \rightharpoonup \bar{x}_2$, $u_{k_n} \rightharpoonup e_1$ and $u_{l_n} \rightharpoonup e_2$. Taking the limit in view of Step 1 along these subsequences we have $\langle e_2 - e_1, e_1 - \bar{x}_1 \rangle = 0 = \langle e_2 - e_1, e_2 - \bar{x}_2 \rangle$ hence $\|e_2 - e_1\|^2 = \langle e_2 - e_1, \bar{x}_2 - \bar{x}_1 \rangle$.
- We recall the following fact: Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X that is Fejér monotone with respect to a nonempty closed convex subset C of X. Let w_1 and w_2 be weak cluster points of $(x_n)_{n\in\mathbb{N}}$. Then $w_1 - w_2 \in (C - C)^\perp$.

Lemma: Suppose that E is a nonempty closed convex subset of X, that $(x_n)_{n\in\mathbb{N}}$ is a sequence in X that is *Fejér monotone with respect to* E, i.e., $(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \le \|x_n - e\|$, that $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in X such that its weak cluster points lie in E, and that $(\forall e \in E) \ \langle u_n - e, x_n - u_n \rangle \to 0$. Then $(u_n)_{n\in\mathbb{N}}$ converges weakly to some point in E.

New Fejér monotonicity principle: proof

- Step 1: $(\forall (e_1, e_2) \in E \times E) \quad \langle e_2 - e_1, u_n - x_n \rangle = \langle u_n - e_1, u_n - x_n \rangle - \langle u_n - e_2, u_n - x_n \rangle \to 0$.
- Step 2: Obtain four subsequences $(x_{k_n})_{n\in\mathbb{N}}$, $(x_{l_n})_{n\in\mathbb{N}}$, $(u_{k_n})_{n\in\mathbb{N}}$ and $(u_{l_n})_{n\in\mathbb{N}}$ such that $x_{k_n} \rightharpoonup \bar{x}_1$, $x_{l_n} \rightharpoonup \bar{x}_2$, $u_{k_n} \rightharpoonup e_1$ and $u_{l_n} \rightharpoonup e_2$. Taking the limit in view of Step 1 along these subsequences we have $\langle e_2 - e_1, e_1 - \bar{x}_1 \rangle = 0 = \langle e_2 - e_1, e_2 - \bar{x}_2 \rangle$ hence $\|e_2 - e_1\|^2 = \langle e_2 - e_1, \bar{x}_2 - \bar{x}_1 \rangle$.
- We recall the following fact: Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X that is Fejér monotone with respect to a nonempty closed convex subset C of X. Let w_1 and w_2 be weak cluster points of $(x_n)_{n\in\mathbb{N}}$. Then $w_1 - w_2 \in (C - C)^\perp$.
- Step 3: Since $\{e_1, e_2\} \subseteq E$, applying the previous fact with (w_1, w_2, C) replaced by $(\bar{x}_1, \bar{x}_2, E)$ we conclude that $\langle e_2 - e_1, \bar{x}_2 - \bar{x}_1 \rangle = 0$.

Lemma: Suppose that E is a nonempty closed convex subset of X, that $(x_n)_{n\in\mathbb{N}}$ is a sequence in X that is *Fejér monotone with respect to E*, i.e., $(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \leq \|x_n - e\|$, that $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in X such that its weak cluster points lie in E, and that $(\forall e \in E) \ \langle u_n - e, x_n - u_n \rangle \to 0$. Then $(u_n)_{n\in\mathbb{N}}$ converges weakly to some point in E.

New Fejér monotonicity principle: proof

- Step 1: $(\forall (e_1, e_2) \in E \times E) \quad \langle e_2 - e_1, u_n - x_n \rangle = \langle u_n - e_1, u_n - x_n \rangle - \langle u_n - e_2, u_n - x_n \rangle \to 0$.
- Step 2: Obtain four subsequences $(x_{k_n})_{n\in\mathbb{N}}$, $(x_{l_n})_{n\in\mathbb{N}}$, $(u_{k_n})_{n\in\mathbb{N}}$ and $(u_{l_n})_{n\in\mathbb{N}}$ such that $x_{k_n} \rightharpoonup \bar{x}_1$, $x_{l_n} \rightharpoonup \bar{x}_2$, $u_{k_n} \rightharpoonup e_1$ and $u_{l_n} \rightharpoonup e_2$. Taking the limit in view of Step 1 along these subsequences we have $\langle e_2 - e_1, e_1 - \bar{x}_1 \rangle = 0 = \langle e_2 - e_1, e_2 - \bar{x}_2 \rangle$ hence $\|e_2 - e_1\|^2 = \langle e_2 - e_1, \bar{x}_2 - \bar{x}_1 \rangle$.
- We recall the following fact: Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X that is Fejér monotone with respect to a nonempty closed convex subset C of X. Let w_1 and w_2 be weak cluster points of $(x_n)_{n\in\mathbb{N}}$. Then $w_1 - w_2 \in (C - C)^{\perp}$.
- Step 3: Since $\{e_1, e_2\} \subseteq E$, applying the previous fact with (w_1, w_2, C) replaced by $(\bar{x}_1, \bar{x}_2, E)$ we conclude that $\langle e_2 - e_1, \bar{x}_2 - \bar{x}_1 \rangle = 0$. By Step2, $e_1 = e_2$.

Lemma: Suppose that E is a nonempty closed convex subset of X, that $(x_n)_{n\in\mathbb{N}}$ is a sequence in X that is *Fejér monotone with respect to* E, i.e., $(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \le \|x_n - e\|$, that $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in X such that its weak cluster points lie in E, and that $(\forall e \in E) \ \langle u_n - e, x_n - u_n \rangle \to 0$. Then $(u_n)_{n\in\mathbb{N}}$ converges weakly to some point in E.

Application to the convex feasibility problem

Theorem

Suppose that U and V are nonempty closed convex subsets of X, and that $U \cap (v + V) \neq \varnothing$. Let $x \in X$. Then $(P_U T^n x)_{n \in \mathbb{N}}$ converges weakly to some point in $U \cap (v + V)$.

When $A = N_U$ and $B = N_V$ we have $T = T_{A,B} = \mathrm{Id} - P_U + P_V(2P_U - \mathrm{Id})$.

Application to the convex feasibility problem

Theorem

Suppose that U and V are nonempty closed convex subsets of X, and that $U \cap (v + V) \neq \varnothing$. Let $x \in X$. Then $(P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in $U \cap (v + V)$.

Proof.

Let $y \in U \cap (v + V) \subseteq \mathrm{Fix}(v + T)$ (Fact).

When $A = N_U$ and $B = N_V$ we have $T = T_{A,B} = \mathrm{Id} - P_U + P_V(2P_U - \mathrm{Id})$.

Application to the convex feasibility problem

Theorem

Suppose that U and V are nonempty closed convex subsets of X, and that $U \cap (v + V) \neq \varnothing$. Let $x \in X$. Then $(P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in $U \cap (v + V)$.

Proof.

Let $y \in U \cap (v + V) \subseteq \mathrm{Fix}(v + T)$ (Fact). Then $P_U T^n y = y$ and $T^n y = y - nv$ (Fact).

When $A = N_U$ and $B = N_V$ we have $T = T_{A,B} = \mathrm{Id} - P_U + P_V(2P_U - \mathrm{Id})$.

Application to the convex feasibility problem

Theorem

Suppose that U and V are nonempty closed convex subsets of X, and that $U \cap (v + V) \neq \varnothing$. Let $x \in X$. Then $(P_U T^n x)_{n \in \mathbb{N}}$ converges weakly to some point in $U \cap (v + V)$.

Proof.

Let $y \in U \cap (v + V) \subseteq \mathrm{Fix}(v + T)$ (Fact). Then $P_U T^n y = y$ and $T^n y = y - nv$ (Fact). We have

$$\langle P_U T^n x - y, T^n x + nv - P_U T^n x \rangle$$
$$= \langle P_U T^n x - y, T^n x - P_U T^n x - (y - nv - y) \rangle$$

□

When $A = N_U$ and $B = N_V$ we have $T = T_{A,B} = \mathrm{Id} - P_U + P_V(2P_U - \mathrm{Id})$.

Application to the convex feasibility problem

Theorem

Suppose that U and V are nonempty closed convex subsets of X, and that $U \cap (v + V) \neq \varnothing$. Let $x \in X$. Then $(P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in $U \cap (v + V)$.

Proof.

Let $y \in U \cap (v + V) \subseteq \mathrm{Fix}(v + T)$ (Fact). Then $P_U T^n y = y$ and $T^n y = y - nv$ (Fact). We have

$$\begin{aligned}
&\langle P_U T^n x - y, T^n x + nv - P_U T^n x\rangle \\
&= \langle P_U T^n x - y, T^n x - P_U T^n x - (y - nv - y)\rangle \\
&= \langle P_U T^n x - P_U T^n y, T^n x - P_U T^n x - (T^n y - P_U T^n y)\rangle
\end{aligned}$$

□

When $A = N_U$ and $B = N_V$ we have $T = T_{A,B} = \mathrm{Id} - P_U + P_V(2P_U - \mathrm{Id})$.

Application to the convex feasibility problem

Theorem

Suppose that U and V are nonempty closed convex subsets of X, and that $U \cap (v + V) \neq \varnothing$. Let $x \in X$. Then $(P_U T^n x)_{n \in \mathbb{N}}$ converges weakly to some point in $U \cap (v + V)$.

Proof.

Let $y \in U \cap (v + V) \subseteq \operatorname{Fix}(v + T)$ (Fact). Then $P_U T^n y = y$ and $T^n y = y - nv$ (Fact). We have

$$\begin{aligned}
&\langle P_U T^n x - y, T^n x + nv - P_U T^n x \rangle \\
&= \langle P_U T^n x - y, T^n x - P_U T^n x - (y - nv - y) \rangle \\
&= \langle P_U T^n x - P_U T^n y, T^n x - P_U T^n x - (T^n y - P_U T^n y) \rangle \\
&= \langle P_U T^n x - P_U T^n y, (\operatorname{Id} - P_U) T^n x - (\operatorname{Id} - P_U) T^n y \rangle \\
&\to 0,
\end{aligned}$$

where the limit follows from an earlier result. □

When $A = N_U$ and $B = N_V$ we have $T = T_{A,B} = \operatorname{Id} - P_U + P_V(2P_U - \operatorname{Id})$.

Application to the convex feasibility problem

Theorem

Suppose that U and V are nonempty closed convex subsets of X, and that $U \cap (v + V) \neq \varnothing$. Let $x \in X$. Then $(P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in $U \cap (v + V)$.

Proof.

Let $y \in U \cap (v + V) \subseteq \operatorname{Fix}(v + T)$ (Fact). Then $P_U T^n y = y$ and $T^n y = y - nv$ (Fact). We have

$$
\begin{aligned}
&\langle P_U T^n x - y, T^n x + nv - P_U T^n x\rangle \\
&= \langle P_U T^n x - y, T^n x - P_U T^n x - (y - nv - y)\rangle \\
&= \langle P_U T^n x - P_U T^n y, T^n x - P_U T^n x - (T^n y - P_U T^n y)\rangle \\
&= \langle P_U T^n x - P_U T^n y, (\operatorname{Id} - P_U) T^n x - (\operatorname{Id} - P_U) T^n y\rangle \\
&\to 0,
\end{aligned}
$$

where the limit follows from an earlier result. □

When $A = N_U$ and $B = N_V$ we have $T = T_{A,B} = \operatorname{Id} - P_U + P_V(2P_U - \operatorname{Id})$.

Application to the convex feasibility problem

Proof continued.

- Fact: $(T^n x + nv)_{n\in\mathbb{N}}$ is Fejér monotone w.r.t. $U \cap (v + V)$.

Lemma: Suppose that E is a nonempty closed convex subset of X, that $(x_n)_{n\in\mathbb{N}}$ is a sequence in X that is *Fejér monotone with respect to E*, i.e., $(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \leq \|x_n - e\|$, that $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in X such that its weak cluster points lie in E, and that $(\forall e \in E)\ \langle u_n - e, x_n - u_n\rangle \to 0$. Then $(u_n)_{n\in\mathbb{N}}$ converges weakly to some point in E.

Application to the convex feasibility problem

Proof continued.

- Fact: $(T^n x + nv)_{n\in\mathbb{N}}$ is Fejér monotone w.r.t. $U \cap (v + V)$.
- Fact: $(P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points lie in $U \cap (v + V)$.

Lemma: Suppose that E is a nonempty closed convex subset of X, that $(x_n)_{n\in\mathbb{N}}$ is a sequence in X that is *Fejér monotone with respect to E*, i.e., $(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \leq \|x_n - e\|$, that $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in X such that its weak cluster points lie in E, and that $(\forall e \in E)\ \langle u_n - e, x_n - u_n\rangle \to 0$. Then $(u_n)_{n\in\mathbb{N}}$ converges weakly to some point in E.

Application to the convex feasibility problem

Proof continued.

- Fact: $(T^n x + nv)_{n\in\mathbb{N}}$ is Fejér monotone w.r.t. $U \cap (v+V)$.
- Fact: $(P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points lie in $U \cap (v+V)$.
- We just proved

$$\langle \underset{\substack{\parallel \\ u_n}}{P_U T^n x} - \underset{\substack{\cap \\ U\cap(v+V)}}{y}, \underset{\substack{\parallel \\ x_n}}{(T^n x + nv)} - \underset{\substack{\parallel \\ u_n}}{P_U T^n x} \rangle \to 0.$$

Lemma: Suppose that E is a nonempty closed convex subset of X, that $(x_n)_{n\in\mathbb{N}}$ is a sequence in X that is *Fejér monotone with respect to E*, i.e., $(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \le \|x_n - e\|$, that $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in X such that its weak cluster points lie in E, and that $(\forall e \in E)\ \langle u_n - e, x_n - u_n\rangle \to 0$. Then $(u_n)_{n\in\mathbb{N}}$ converges weakly to some point in E.

Application to the convex feasibility problem

Proof continued.

- Fact: $(T^n x + nv)_{n\in\mathbb{N}}$ is Fejér monotone w.r.t. $U \cap (v+V)$.
- Fact: $(P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points lie in $U \cap (v+V)$.
- We just proved

$$\langle \underbrace{P_U T^n x}_{= u_n} - \underbrace{y}_{\in U\cap(v+V)}, \underbrace{(T^n x + nv)}_{= x_n} - \underbrace{P_U T^n x}_{= u_n} \rangle \to 0.$$

- Apply the Fejér monotonicity lemma with $(E, (u_n)_{n\in\mathbb{N}}, (x_n)_{n\in\mathbb{N}})$ replaced by $(U\cap(v+V), (P_U T^n x)_{n\in\mathbb{N}}, (T^n x + nv)_{n\in\mathbb{N}})$.

□

Lemma: Suppose that E is a nonempty closed convex subset of X, that $(x_n)_{n\in\mathbb{N}}$ is a sequence in X that is *Fejér monotone with respect to E*, i.e., $(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \le \|x_n - e\|$, that $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in X such that its weak cluster points lie in E, and that $(\forall e \in E)\ \langle u_n - e, x_n - u_n\rangle \to 0$. Then $(u_n)_{n\in\mathbb{N}}$ converges weakly to some point in E.

Example

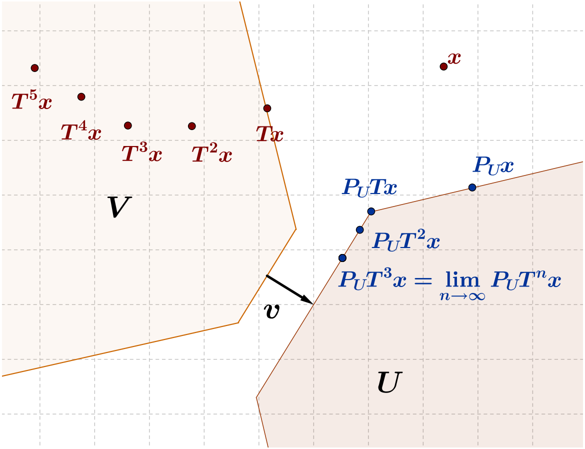


Figure: A `GeoGebra` snapshot. U and V are two nonintersecting sets in $\mathbb{R}^2$. Also, the first few iterates of the governing sequence $(T^n x)_{n\in\mathbb{N}}$ (red points) and the shadow sequence $(P_U T^n x)_{n\in\mathbb{N}}$ (blue points) are shown.

The Douglas–Rachford operator for two affine subspaces

In the following we set

$$T_{U,V} := T_{N_U, N_V},$$

where U and V are nonempty closed convex subsets of X.

Proposition

Suppose that U and V are affine subspaces of X. Set $A := N_U$, $B := N_V$ and $T := T_{U,V}$. Let $x \in X$. Then the following hold.

(i) $v \in (\operatorname{par} U)^{\perp} \cap (\operatorname{par} V)^{\perp}$.

Let U be an affine subspace of X. Then $\operatorname{par} U = U - U$.

The Douglas–Rachford operator for two affine subspaces

In the following we set

$$T_{U,V} := T_{N_U,N_V},$$

where U and V are nonempty closed convex subsets of X.

Proposition

Suppose that U and V are affine subspaces of X. Set $A := N_U$, $B := N_V$ and $T := T_{U,V}$. Let $x \in X$. Then the following hold.

(i) $v \in (\text{par } U)^{\perp} \cap (\text{par } V)^{\perp}$.

(ii) $(\forall \alpha \in \mathbb{R})\ P_U x = P_U(x + \alpha v)$.

Let U be an affine subspace of X. Then $\text{par } U = U - U$.

The Douglas–Rachford operator for two affine subspaces

In the following we set

$$T_{U,V} := T_{N_U,N_V},$$

where U and V are nonempty closed convex subsets of X.

Proposition

Suppose that U and V are affine subspaces of X. Set $A := N_U$, $B := N_V$ and $T := T_{U,V}$. Let $x \in X$. Then the following hold.

(i) $v \in (\operatorname{par} U)^{\perp} \cap (\operatorname{par} V)^{\perp}$.

(ii) $(\forall \alpha \in \mathbb{R})\ P_U x = P_U(x + \alpha v)$.

(iii) $(\forall n \in \mathbb{N})\ T^n x + nv = T^n_{U,v+V} x$.

Let U be an affine subspace of X. Then $\operatorname{par} U = U - U$.

Convergence of the shadows

Theorem

Let $x \in X$. Then the following hold.

(i) $(\forall n \in \mathbb{N})\ P_U T^n x = P_U T^n_{U,v+V} x$.

(ii) $P_U T^n x \to P_{U\cap(v+V)} x$. *If* $\operatorname{par} U + \operatorname{par} V$ *is closed then the convergence is linear with rate* $c_F(\operatorname{par} U, \operatorname{par} V) < 1$.

The cosine of the Friedrichs angle between $\operatorname{par} U$ and $\operatorname{par} V$ is $c_F :=$ $\sup\left\{|\langle u, v\rangle| \mid u \in \operatorname{par} U \cap (\operatorname{par} U \cap \operatorname{par} V)^\perp, v \in \operatorname{par} V \cap (\operatorname{par} U \cap \operatorname{par} V)^\perp, \|u\| \leq 1, \|v\| \leq 1\right\}$

Convergence of the shadows

Theorem

Let $x \in X$. Then the following hold.

(i) $(\forall n \in \mathbb{N})\ P_U T^n x = P_U T^n_{U,v+V} x$.

(ii) $P_U T^n x \to P_{U\cap(v+V)} x$. *If* $\operatorname{par} U + \operatorname{par} V$ *is closed then the convergence is linear with rate* $c_F(\operatorname{par} U, \operatorname{par} V) < 1$.

Proof

(i): We want to show that $(\forall n \in \mathbb{N})\ P_U T^n x = P_U T^n_{U,v+V} x$.

The cosine of the Friedrichs angle between $\operatorname{par} U$ and $\operatorname{par} V$ is $c_F :=$ $\sup\left\{|\langle u, v\rangle| \mid u \in \operatorname{par} U \cap (\operatorname{par} U \cap \operatorname{par} V)^{\perp}, v \in \operatorname{par} V \cap (\operatorname{par} U \cap \operatorname{par} V)^{\perp}, \|u\| \leq 1, \|v\| \leq 1\right\}$

Convergence of the shadows

Theorem

Let $x \in X$. Then the following hold.

(i) $(\forall n \in \mathbb{N})\ P_U T^n x = P_U T^n_{U,v+V} x$.

(ii) $P_U T^n x \to P_{U\cap(v+V)} x$. *If* $\operatorname{par} U + \operatorname{par} V$ *is closed then the convergence is linear with rate* $c_F(\operatorname{par} U, \operatorname{par} V) < 1$.

Proof

(i): We want to show that $(\forall n \in \mathbb{N})\ P_U T^n x = P_U T^n_{U,v+V} x$.
Let $n \in \mathbb{N}$ and recall that $v \in (\operatorname{par} U)^\perp$.
$P_U T^n x = P_U(T^n x + nv)$

The cosine of the Friedrichs angle between $\operatorname{par} U$ and $\operatorname{par} V$ is $c_F :=$ $\sup\{|\langle u, v\rangle| \mid u \in \operatorname{par} U \cap (\operatorname{par} U \cap \operatorname{par} V)^\perp, v \in \operatorname{par} V \cap (\operatorname{par} U \cap \operatorname{par} V)^\perp, \|u\| \le 1, \|v\| \le 1\}$

Convergence of the shadows

Theorem

Let $x \in X$. Then the following hold.

(i) $(\forall n \in \mathbb{N})\ P_U T^n x = P_U T^n_{U,v+V} x$.

(ii) $P_U T^n x \to P_{U\cap(v+V)} x$. *If* $\operatorname{par} U + \operatorname{par} V$ *is closed then the convergence is linear with rate* $c_F(\operatorname{par} U, \operatorname{par} V) < 1$.

Proof

(i): We want to show that $(\forall n \in \mathbb{N})\ P_U T^n x = P_U T^n_{U,v+V} x$.
Let $n \in \mathbb{N}$ and recall that $v \in (\operatorname{par} U)^{\perp}$.
$P_U T^n x = P_U(T^n x + nv) = P_U T^n_{U,v+V} x$.

The cosine of the Friedrichs angle between par U and par V is $c_F :=$
$\sup\{|\langle u, v\rangle| \mid u \in \operatorname{par} U \cap (\operatorname{par} U \cap \operatorname{par} V)^{\perp}, v \in \operatorname{par} V \cap (\operatorname{par} U \cap \operatorname{par} V)^{\perp}, \|u\| \le 1, \|v\| \le 1\}$

Sketch of the proof (continued)

(ii): We want to show that $P_U T^n x \to P_{U \cap (v+V)} x$.

Fact (Bauschke, Cruz, Nghia, Phan, Wang (2014))

Suppose that U and V are closed affine subspace of X such that $U \cap V \neq \varnothing$. Then $T^n x \to P_{\mathrm{Fix}\, T} x$, $P_U T^n x \to P_{U \cap V} x$, and $P_V T^n x \to P_{U \cap V} x$. If $\mathrm{par}\, U + \mathrm{par}\, V$ is closed then the convergence is linear with rate $c_F(\mathrm{par}\, U, \mathrm{par}\, V) < 1$.

$c_F :=$
$\sup \{ |\langle u, v \rangle| \mid u \in \mathrm{par}\, U \cap (\mathrm{par}\, U \cap \mathrm{par}\, V)^{\perp}, v \in \mathrm{par}\, V \cap (\mathrm{par}\, U \cap \mathrm{par}\, V)^{\perp}, \|u\| \leq 1, \|v\| \leq 1 \}$

Sketch of the proof (continued)

(ii): We want to show that $P_U T^n x \to P_{U\cap(v+V)} x$.

Fact (Bauschke, Cruz, Nghia, Phan, Wang (2014))

Suppose that U and V are closed affine subspace of X such that $U \cap V \neq \varnothing$. Then $T^n x \to P_{\mathrm{Fix}\, T} x$, $P_U T^n x \to P_{U\cap V} x$, and $P_V T^n x \to P_{U\cap V} x$. If $\mathrm{par}\, U + \mathrm{par}\, V$ is closed then the convergence is linear with rate $c_F(\mathrm{par}\, U, \mathrm{par}\, V) < 1$.

Step 1: Since U and $v + V$ are closed affine subspace of X and $U \cap (v + V) \neq \varnothing$, we can apply the above fact to the sets U and $v + V$ to get $P_U T^n_{U,v+V} x \to P_{U\cap(v+V)} x$.

$c_F :=$
$\sup\left\{|\langle u, v\rangle| \mid u \in \mathrm{par}\, U \cap (\mathrm{par}\, U \cap \mathrm{par}\, V)^\perp, v \in \mathrm{par}\, V \cap (\mathrm{par}\, U \cap \mathrm{par}\, V)^\perp, \|u\| \leq 1, \|v\| \leq 1\right\}$

Sketch of the proof (continued)

(ii): We want to show that $P_U T^n x \to P_{U\cap(v+V)}x$.

Fact (Bauschke, Cruz, Nghia, Phan, Wang (2014))

Suppose that U and V are closed affine subspace of X such that $U \cap V \neq \varnothing$. Then $T^n x \to P_{\operatorname{Fix} T}x$, $P_U T^n x \to P_{U\cap V}x$, and $P_V T^n x \to P_{U\cap V}x$. If $\operatorname{par} U + \operatorname{par} V$ is closed then the convergence is linear with rate $c_F(\operatorname{par} U, \operatorname{par} V) < 1$.

Step 1: Since U and $v + V$ are closed affine subspace of X and $U \cap (v + V) \neq \varnothing$, we can apply the above fact to the sets U and $v + V$ to get $P_U T^n_{U,v+V}x \to P_{U\cap(v+V)}x$.

Step 2: Using (i) we have $P_U T^n x = P_U T^n_{U,v+V}x$, which when combined with step 1 proves the claim.

$c_F :=$
$\sup\left\{|\langle u, v\rangle| \mid u \in \operatorname{par} U \cap (\operatorname{par} U \cap \operatorname{par} V)^\perp, v \in \operatorname{par} V \cap (\operatorname{par} U \cap \operatorname{par} V)^\perp, \|u\| \le 1, \|v\| \le 1\right\}$

Sketch of the proof (continued)

(ii): We want to show that $P_U T^n x \to P_{U\cap(v+V)}x$.

Fact (Bauschke, Cruz, Nghia, Phan, Wang (2014))

Suppose that U and V are closed affine subspace of X such that $U \cap V \neq \varnothing$. Then $T^n x \to P_{\mathrm{Fix}\, T}x$, $P_U T^n x \to P_{U\cap V}x$, and $P_V T^n x \to P_{U\cap V}x$. If $\mathrm{par}\, U + \mathrm{par}\, V$ is closed then the convergence is linear with rate $c_F(\mathrm{par}\, U, \mathrm{par}\, V) < 1$.

Step 1: Since U and $v + V$ are closed affine subspace of X and $U \cap (v + V) \neq \varnothing$, we can apply the above fact to the sets U and $v + V$ to get $P_U T^n_{U,v+V}x \to P_{U\cap(v+V)}x$.

Step 2: Using (i) we have $P_U T^n x = P_U T^n_{U,v+V}x$, which when combined with step 1 proves the claim.

Step 3: Finally notice that $\mathrm{par}(v + V) = \mathrm{par}\, V$, hence if $\mathrm{par}\, U + \mathrm{par}\, V$ is closed then the convergence is linear with rate $c_F(\mathrm{par}\, U, \mathrm{par}\, V) < 1$, where c_F is the cosine of the Friedrichs angle between U and V.

$c_F :=$
$\sup\left\{|\langle u, v\rangle| \mid u \in \mathrm{par}\, U \cap (\mathrm{par}\, U \cap \mathrm{par}\, V)^\perp, v \in \mathrm{par}\, V \cap (\mathrm{par}\, U \cap \mathrm{par}\, V)^\perp, \|u\| \le 1, \|v\| \le 1\right\}$

When one set is an affine subspace

Recall that

$$v := P_{\overline{U-V}}(0) \in \operatorname{ran}(\mathrm{Id} - T).$$

Theorem (convergence of DRA when U is a closed affine subspace)

Suppose that U is a closed affine subspace of X and that V is a nonempty closed convex subset of X. Let $x \in X$. Then

(i) *The shadow sequence $(P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in $U \cap (V + v)$.*

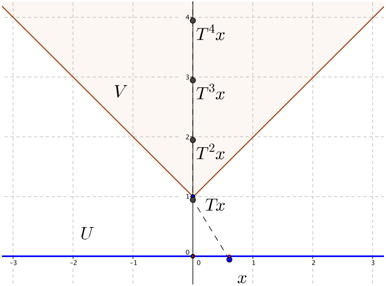
(ii) *No general conclusion can be drawn about the sequence $(P_V T^n x)_{n\in\mathbb{N}}$.*

Example

To prove: No general conclusion can be drawn about the sequence $(P_V T^n x)_{n\in\mathbb{N}}$. Recall that we proved the weak convergence of $(P_U T^n x)_{n\in\mathbb{N}}$ to a best approximation solution.

Example

Suppose that $X = \mathbb{R}^2$, that $U = \mathbb{R} \times \{0\}$ and that $V = \operatorname{epi}(|\cdot| + 1)$. Then $U \cap V = \varnothing$ and for the starting point $x \in [-1, 1] \times \{0\}$ we have $(\forall n \in \{1, 2, \ldots\})$ $T^n x = (0, n) \in V$ and therefore $\|P_V T^n x\| = \|T^n x\| = n \to +\infty$.



Application to the convex feasibility problems for more than two sets

Theorem

Suppose that $V_1, \ldots, V_M$ *are closed convex subsets of* X. *Set* $\mathbf{X} = X^M$, $\mathbf{U} = \big\{(x, \ldots, x) \in \mathbf{X} \;\big|\; x \in X\big\}$ *and* $\mathbf{V} = V_1 \times \cdots \times V_M$. *Let* $\mathbf{T} = \mathrm{Id} - P_{\mathbf{U}} + P_{\mathbf{V}}(2P_{\mathbf{U}} - \mathrm{Id})$, *let* $\mathbf{x} \in \mathbf{X}$ *and suppose that* $\mathbf{v} = (v_1, \ldots, v_M) := P_{\overline{\mathbf{U}-\mathbf{V}}}0 \in \mathbf{U} - \mathbf{V}$. *Then the shadow sequence* $(P_{\mathbf{U}}\mathbf{T}^n\mathbf{x})_{n\in\mathbb{N}}$ *converges to* $\bar{\mathbf{x}} = (\bar{x}, \ldots, \bar{x}) \in \mathbf{U} \cap (\mathbf{v} + \mathbf{V})$, *where* $\bar{x} \in \bigcap_{i=1}^{M}(v_i + V_i)$ *and* $\bar{x}$ *is a least-squares solution of*

$$\textit{find a minimizer of } \sum_{i=1}^{M} d_{V_i}^2.$$

Application to the convex feasibility problems for more than two sets

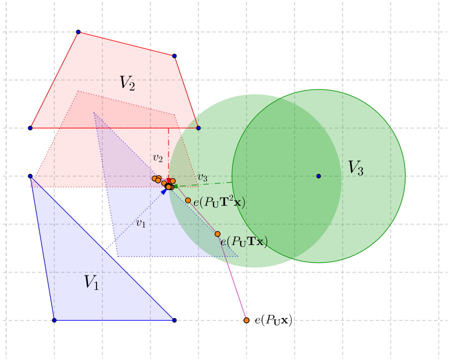


Figure: A GeoGebra snapshot. The DRA finds a point in the generalized intersection. Shown are the original sets as well the translated sets that forms this intersection.

And beyond feasibility!

Theorem

Suppose that

- *U is a closed affine subspace of X,*
- *$A = N_U$, that B is rectangular,*
- *$v = P_{\overline{\mathrm{ran}}(\mathrm{Id}-T)}0 \in \mathrm{ran}(\mathrm{Id}-T)$,*
- *$\mathrm{zer}({}_vA) \cap \mathrm{zer}(B_v) \neq \varnothing$ and*
- *all weak cluster points of $(J_A T^n x)_{n\in\mathbb{N}} = (P_U T^n x)_{n\in\mathbb{N}}$ lie in Z_v.*

Let $x \in X$. Then $(J_A T^n x)_{n\in\mathbb{N}} = (P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in Z_v.

Let $C\colon X \rightrightarrows X$. Then C rectangular (this is also known as paramonotone) if A is monotone and we have the implication

$$\left.\begin{array}{c}(x,u)\in \mathrm{gr}\,C\\(y,v)\in \mathrm{gr}\,C\\\langle x-y,u-v\rangle = 0\end{array}\right\} \quad\Rightarrow\quad \{(x,v),(y,u)\} \subseteq \mathrm{gr}\,C.$$

How far could the results be generalized?

- **Known:** U and V are (possibly nonintersecting) nonempty closed convex subsets $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points are normal solutions.

How far could the results be generalized?

- **Known:** U and V are (possibly nonintersecting) nonempty closed convex subsets $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points are normal solutions.
- **New results:** U and V are nonempty closed convex subset (possibly nonintersecting) $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to a best approximation solution.

How far could the results be generalized?

- **Known:** U and V are (possibly nonintersecting) nonempty closed convex subsets $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points are normal solutions.
- **New results:** U and V are nonempty closed convex subset (possibly nonintersecting) $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to a best approximation solution.
- **Still open:** What happens in the case of general subdifferential operators, i.e., when A and B are subdifferential operators of convex functions that are not necessarily indicator functions?

How far could the results be generalized?

- **Known:** U and V are (possibly nonintersecting) nonempty closed convex subsets $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points are normal solutions.
- **New results:** U and V are nonempty closed convex subset (possibly nonintersecting) $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to a best approximation solution.
- **Still open:** What happens in the case of general subdifferential operators, i.e., when A and B are subdifferential operators of convex functions that are not necessarily indicator functions?

Example

Suppose that U is a closed linear subspace of X, that $A := N_U$, that $b \in U \smallsetminus \{0\}$ and that $B\colon X \to X\colon x \mapsto b$.

How far could the results be generalized?

- **Known:** U and V are (possibly nonintersecting) nonempty closed convex subsets $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points are normal solutions.
- **New results:** U and V are nonempty closed convex subset (possibly nonintersecting) $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to a best approximation solution.
- **Still open:** What happens in the case of general subdifferential operators, i.e., when A and B are subdifferential operators of convex functions that are not necessarily indicator functions?

Example

Suppose that U is a closed linear subspace of X, that $A := N_U$, that $b \in U \smallsetminus \{0\}$ and that $B\colon X \to X\colon x \mapsto b$.

- Recall that $v = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}0 = P_{\overline{\operatorname{ran}}(b+U^\perp)}0 = b - P_{U^\perp} b = b \neq 0$,

How far could the results be generalized?

- **Known:** U and V are (possibly nonintersecting) nonempty closed convex subsets $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points are normal solutions.
- **New results:** U and V are nonempty closed convex subset (possibly nonintersecting) $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to a best approximation solution.
- **Still open:** What happens in the case of general subdifferential operators, i.e., when A and B are subdifferential operators of convex functions that are not necessarily indicator functions?

Example

Suppose that U is a closed linear subspace of X, that $A := N_U$, that $b \in U \smallsetminus \{0\}$ and that $B\colon X \to X\colon x \mapsto b$.

- Recall that $v = P_{\overline{\operatorname{ran}}(\mathrm{Id}-T)}0 = P_{\overline{\operatorname{ran}}(b+U^\perp)}0 = b - P_{U^\perp} b = b \neq 0$, i.e., the sum problem (find $x \in X$ such that $0 \in Ax + Bx$) is inconsistent. However, $v = b \in b + U^\perp = \operatorname{ran}(\mathrm{Id}-T)$, hence $Z_v \neq \varnothing$.

How far could the results be generalized?

- **Known:** U and V are (possibly nonintersecting) nonempty closed convex subsets $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points are normal solutions.
- **New results:** U and V are nonempty closed convex subset (possibly nonintersecting) $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to a best approximation solution.
- **Still open:** What happens in the case of general subdifferential operators, i.e., when A and B are subdifferential operators of convex functions that are not necessarily indicator functions?

Example

Suppose that U is a closed linear subspace of X, that $A := N_U$, that $b \in U \smallsetminus \{0\}$ and that $B\colon X \to X\colon x \mapsto b$.

- Recall that $v = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}0 = P_{\overline{\operatorname{ran}}(b+U^\perp)}0 = b - P_{U^\perp} b = b \neq 0$, i.e., the sum problem (find $x \in X$ such that $0 \in Ax + Bx$) is inconsistent. However, $v = b \in b + U^\perp = \operatorname{ran}(\operatorname{Id} - T)$, hence $Z_v \neq \varnothing$.
- One can show that $(\forall n \in \mathbb{N})\ T^n x = P_U x - nb$, hence $\|T^n x\| \to +\infty$.

How far could the results be generalized?

- **Known:** U and V are (possibly nonintersecting) nonempty closed convex subsets $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points are normal solutions.
- **New results:** U and V are nonempty closed convex subset (possibly nonintersecting) $\Rightarrow (P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to a best approximation solution.
- **Still open:** What happens in the case of general subdifferential operators, i.e., when A and B are subdifferential operators of convex functions that are not necessarily indicator functions?

Example

Suppose that U is a closed linear subspace of X, that $A := N_U$, that $b \in U \smallsetminus \{0\}$ and that $B\colon X \to X\colon x \mapsto b$.

- Recall that $v = P_{\overline{\operatorname{ran}}(\operatorname{Id}-T)}0 = P_{\overline{\operatorname{ran}}(b+U^\perp)}0 = b - P_{U^\perp}b = b \neq 0$, i.e., the sum problem (find $x \in X$ such that $0 \in Ax + Bx$) is inconsistent. However, $v = b \in b + U^\perp = \operatorname{ran}(\operatorname{Id}-T)$, hence $Z_v \neq \varnothing$.
- One can show that $(\forall n \in \mathbb{N})\ T^n x = P_U x - nb$, hence $\|T^n x\| \to +\infty$.
- Consequently, $P_U T^n x = T^n x$, hence $\|P_U T^n x\| \to +\infty$ (unbounded!).

Convergence of shadows: Brief literature review

- Krasnosel'skiĭ–Mann (1950s)

$$T^n x \xrightarrow[\textit{weakly}]{} \text{some point in } \mathrm{Fix}\, T \neq \mathrm{zer}(A+B).$$

Convergence of shadows: Brief literature review

- Krasnosel'skiĭ–Mann (1950s)

$$T^n x \xrightarrow[\textit{weakly}]{} \text{some point in } \mathrm{Fix}\, T \neq \mathrm{zer}(A+B).$$

- Combettes (2004) $J_A(\mathrm{Fix}\, T) = \mathrm{zer}(A+B)$.

Convergence of shadows: Brief literature review

- Krasnosel'skiĭ–Mann (1950s)

$$T^n x \xrightarrow[\text{weakly}]{} \text{some point in } \mathrm{Fix}\, T \neq \mathrm{zer}(A+B).$$

- Combettes (2004) $J_A(\mathrm{Fix}\, T) = \mathrm{zer}(A+B)$.
- Lions–Mercier (1979) The bounded sequence $(J_A T^n x)_{n\in\mathbb{N}}$ has its weak cluster points lie in $\mathrm{zer}(A+B)$ provided that $A+B$ is maximally monotone.

Convergence of shadows: Brief literature review

- Krasnosel'skiĭ–Mann (1950s)

$$T^n x \xrightarrow[\text{weakly}]{} \text{some point in } \mathrm{Fix}\, T \neq \mathrm{zer}(A+B).$$

- Combettes (2004) $J_A(\mathrm{Fix}\, T) = \mathrm{zer}(A+B)$.
- Lions–Mercier (1979) The bounded sequence $(J_A T^n x)_{n\in\mathbb{N}}$ has its weak cluster points lie in $\mathrm{zer}(A+B)$ provided that $A+B$ is maximally monotone.
- Svaiter (2011)

$$J_A T^n x \xrightarrow[\text{weakly}]{} \text{some point in } \mathrm{zer}(A+B).$$

Attouch–Théra duality and the Douglas–Rachford operator

The (Attouch–Théra) dual problem for the ordered pair (A, B) is to find a zero of $A^{-1} + B^{-\varovee}$, where $B := (-\operatorname{Id}) \circ B \circ (-\operatorname{Id})$. The primal (respectively dual) solutions are the solutions to the primal (respectively dual) problem given by

$$Z := \operatorname{zer}(A + B) \quad \text{and} \quad K := \operatorname{zer}(A^{-1} + B^{-\varovee}).$$

Attouch–Théra duality and the Douglas–Rachford operator

The (Attouch–Théra) dual problem for the ordered pair (A, B) is to find a zero of $A^{-1} + B^{-\ovee}$, where $B := (-\operatorname{Id}) \circ B \circ (-\operatorname{Id})$. The primal (respectively dual) solutions are the solutions to the primal (respectively dual) problem given by

$$Z := \operatorname{zer}(A + B) \quad \text{and} \quad K := \operatorname{zer}(A^{-1} + B^{-\ovee}).$$

Fact (Eckstein (1989))

$$T_{(A,B)} = T_{(A^{-1},B^{-\ovee})}.$$

Attouch–Théra duality and the Douglas–Rachford operator

The (Attouch–Théra) dual problem for the ordered pair (A, B) is to find a zero of $A^{-1} + B^{-\circledvee}$, where $B := (-\operatorname{Id}) \circ B \circ (-\operatorname{Id})$. The primal (respectively dual) solutions are the solutions to the primal (respectively dual) problem given by

$$Z := \operatorname{zer}(A + B) \quad \text{and} \quad K := \operatorname{zer}(A^{-1} + B^{-\circledvee}).$$

Fact (Eckstein (1989))

$$T_{(A,B)} = T_{(A^{-1}, B^{-\circledvee})} := T.$$

Corollary

$$Z \times K = J_A(\operatorname{Fix} T) \times J_{A^{-1}}(\operatorname{Fix} T).$$

Proof.

Combine Combettes's result ($Z = J_A(\operatorname{Fix} T)$), applied to the primal and the dual problems, with Eckstein's above result.

□

Shadows' convergence: Useful identities

Recall that we proved earlier the useful identity:

$$\begin{aligned}\|x-y\|^2 = \|Tx-Ty\|^2 &+ \|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2\\ &+ 2\langle J_A x - J_A y, J_{A^{-1}}x - J_{A^{-1}}y\rangle\\ &+ 2\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle.\end{aligned}$$

Shadows' convergence: Useful identities

Recall that we proved earlier the useful identity:

$$\begin{aligned}\|x-y\|^2 = \|Tx - Ty\|^2 + \|(\mathrm{Id}-T)x - (\mathrm{Id}-T)y\|^2 \\ + 2\langle J_A x - J_A y, J_{A^{-1}} x - J_{A^{-1}} y\rangle \\ + 2\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle.\end{aligned}$$

Using the inverse resolvent identity $J_A + J_{A^{-1}} = \mathrm{Id}$, write:

$$\begin{aligned}\|x-y\|^2 &= \|J_A x - J_A y + J_{A^{-1}} x - J_{A^{-1}} y\|^2 \\ &= \|J_A x - J_A y\|^2 + \|J_{A^{-1}} x - J_{A^{-1}} y\|^2 + 2\langle J_A x - J_A y, J_{A^{-1}} x - J_{A^{-1}} y\rangle.\end{aligned}$$

Shadows' convergence: Useful identities

Recall that we proved earlier the useful identity:

$$\begin{aligned}\|x-y\|^2 = \|Tx-Ty\|^2 &+ \|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2 \\ &+ 2\langle J_A x - J_A y, J_{A^{-1}}x - J_{A^{-1}}y\rangle \\ &+ 2\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle.\end{aligned}$$

Using the inverse resolvent identity $J_A + J_{A^{-1}} = \mathrm{Id}$, write:

$$\begin{aligned}\|x-y\|^2 &= \|J_A x - J_A y + J_{A^{-1}}x - J_{A^{-1}}y\|^2 \\ &= \|J_A x - J_A y\|^2 + \|J_{A^{-1}}x - J_{A^{-1}}y\|^2 + 2\langle J_A x - J_A y, J_{A^{-1}}x - J_{A^{-1}}y\rangle.\end{aligned}$$

and

$$\begin{aligned}\|Tx-Ty\|^2 = \|J_A Tx - J_A Ty\|^2 &+ \|J_{A^{-1}}Tx - J_{A^{-1}}Ty\|^2 \\ &+ 2\langle J_A Tx - J_A Ty, J_{A^{-1}}Tx - J_{A^{-1}}Ty\rangle.\end{aligned}$$

Shadows' convergence: Useful identities

Recall that we proved earlier the useful identity:

$$\begin{aligned}\|x-y\|^2 = \|Tx - Ty\|^2 &+ \|(\mathrm{Id}-T)x - (\mathrm{Id}-T)y\|^2 \\ &+ 2\langle J_A x - J_A y, J_{A^{-1}}x - J_{A^{-1}}y\rangle \\ &+ 2\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle.\end{aligned}$$

Using the inverse resolvent identity $J_A + J_{A^{-1}} = \mathrm{Id}$, write:

$$\begin{aligned}\|x-y\|^2 &= \|J_A x - J_A y + J_{A^{-1}}x - J_{A^{-1}}y\|^2 \\ &= \|J_A x - J_A y\|^2 + \|J_{A^{-1}}x - J_{A^{-1}}y\|^2 + 2\langle J_A x - J_A y, J_{A^{-1}}x - J_{A^{-1}}y\rangle.\end{aligned}$$

and

$$\begin{aligned}\|Tx - Ty\|^2 = \|J_A Tx - J_A Ty\|^2 &+ \|J_{A^{-1}}Tx - J_{A^{-1}}Ty\|^2 \\ &+ 2\langle J_A Tx - J_A Ty, J_{A^{-1}}Tx - J_{A^{-1}}Ty\rangle.\end{aligned}$$

Substituting in the first identity and simplifying yields:

$$\begin{aligned}&\|J_A x - J_A y\|^2 + \|J_{A^{-1}}x - J_{A^{-1}}y\|^2 - \|J_A Tx - J_A Ty\|^2 - \|J_{A^{-1}}Tx - J_{A^{-1}}Ty\|^2 \\ &= \|(\mathrm{Id}-T)x - (\mathrm{Id}-T)y\|^2 + 2\langle J_A Tx - J_A Ty, J_{A^{-1}}Tx - J_{A^{-1}}Ty\rangle \\ &\quad + 2\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle.\end{aligned}$$

Shadows' convergence: Useful identities

We now have

$$\|J_A x - J_A y\|^2 + \|J_{A^{-1}} x - J_{A^{-1}} y\|^2 - \|J_A Tx - J_A Ty\|^2 - \|J_{A^{-1}} Tx - J_{A^{-1}} Ty\|^2$$
$$= \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 + \underbrace{2\langle J_A Tx - J_A Ty, J_{A^{-1}} Tx - J_{A^{-1}} Ty\rangle}_{\geq 0}$$
$$+ \underbrace{2\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle}_{\geq 0}.$$

Shadows' convergence: Useful identities

We now have

$$\|J_A x - J_A y\|^2 + \|J_{A^{-1}} x - J_{A^{-1}} y\|^2 - \|J_A Tx - J_A Ty\|^2 - \|J_{A^{-1}} Tx - J_{A^{-1}} Ty\|^2$$
$$= \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 + \underbrace{2\langle J_A Tx - J_A Ty, J_{A^{-1}} Tx - J_{A^{-1}} Ty\rangle}_{\geq 0}$$
$$+ \underbrace{2\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle}_{\geq 0}.$$

Hence we conclude that

$$\|J_A Tx - J_A Ty\|^2 + \|J_{A^{-1}} Tx - J_{A^{-1}} Ty\|^2 \leq \|J_A x - J_A y\|^2 + \|J_{A^{-1}} x - J_{A^{-1}} y\|^2.$$

Shadows' convergence: Useful identities

We now have

$$\|J_A x - J_A y\|^2 + \|J_{A^{-1}} x - J_{A^{-1}} y\|^2 - \|J_A Tx - J_A Ty\|^2 - \|J_{A^{-1}} Tx - J_{A^{-1}} Ty\|^2$$
$$= \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 + \underbrace{2\langle J_A Tx - J_A Ty, J_{A^{-1}} Tx - J_{A^{-1}} Ty\rangle}_{\geq 0}$$
$$+ \underbrace{2\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle}_{\geq 0}.$$

Hence we conclude that

$$\|J_A Tx - J_A Ty\|^2 + \|J_{A^{-1}} Tx - J_{A^{-1}} Ty\|^2 \leq \|J_A x - J_A y\|^2 + \|J_{A^{-1}} x - J_{A^{-1}} y\|^2.$$

Working in $X \times X$, we can just write

$$\|(J_A Tx, J_{A^{-1}} Tx) - (J_A Ty, J_{A^{-1}} Ty)\|^2 \leq \|(J_A x, J_{A^{-1}} x) - (J_A y, J_{A^{-1}} y)\|^2.$$

Shadows convergence: A simplified proof

Recall that the so-called Kuhn–Tucker set is defined by

$$\mathcal{S} := \mathcal{S}_{(A,B)} := \{(z,k) \in X \times X \mid -k \in Bz, k \in Az\} \subseteq Z \times K.$$

Shadows convergence: A simplified proof

Recall that the so-called Kuhn–Tucker set is defined by

$$\mathcal{S} := \mathcal{S}_{(A,B)} := \{(z,k) \in X \times X \mid -k \in Bz, k \in Az\} \subseteq Z \times K.$$

Theorem

Suppose that $Z = \operatorname{zer}(A+B) \neq \varnothing$. Let $x \in X$ and let $(z,k) \in \mathcal{S}$. Then the following hold:

(i) *For every $n \in \mathbb{N}$, we have*
$\|(J_A T^{n+1}x, J_{A^{-1}} T^{n+1}x) - (z,k)\|^2 \leq \|(J_A T^n x, J_{A^{-1}} T^n x) - (z,k)\|^2$,
i..e., $(J_A T^n x, J_{A^{-1}} T^n x)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $\mathcal{S}$.

Shadows convergence: A simplified proof

Recall that the so-called Kuhn–Tucker set is defined by

$$\mathcal{S} := \mathcal{S}_{(A,B)} := \{(z,k) \in X \times X \mid -k \in Bz, k \in Az\} \subseteq Z \times K.$$

Theorem

Suppose that $Z = \operatorname{zer}(A+B) \neq \varnothing$. Let $x \in X$ and let $(z,k) \in \mathcal{S}$. Then the following hold:

(i) *For every $n \in \mathbb{N}$, we have*
$\|(J_A T^{n+1}x, J_{A^{-1}} T^{n+1}x) - (z,k)\|^2 \leq \|(J_A T^n x, J_{A^{-1}} T^n x) - (z,k)\|^2$,
i..e., $(J_A T^n x, J_{A^{-1}} T^n x)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $\mathcal{S}$.

(ii) $(J_A T^n x, J_{A^{-1}} T^n x)_{n \in \mathbb{N}}$ *converges weakly to some point in $\mathcal{S}$.*

Shadows convergence: A simplified proof

Recall that the so-called Kuhn–Tucker set is defined by

$$\mathcal{S} := \mathcal{S}_{(A,B)} := \{(z,k) \in X \times X \mid -k \in Bz, k \in Az\} \subseteq Z \times K.$$

Theorem

Suppose that $Z = \operatorname{zer}(A+B) \neq \varnothing$. Let $x \in X$ and let $(z,k) \in \mathcal{S}$. Then the following hold:

(i) *For every $n \in \mathbb{N}$, we have*
$\|(J_A T^{n+1}x, J_{A^{-1}} T^{n+1}x) - (z,k)\|^2 \leq \|(J_A T^n x, J_{A^{-1}} T^n x) - (z,k)\|^2$,
i..e., $(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\mathcal{S}$.

(ii) *$(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in $\mathcal{S}$.*

Proof.

(i): We have $z + k \in \operatorname{Fix} T$ (details omitted).

Shadows convergence: A simplified proof

Recall that the so-called Kuhn–Tucker set is defined by

$$\mathcal{S} := \mathcal{S}_{(A,B)} := \{(z,k) \in X \times X \mid -k \in Bz, k \in Az\} \subseteq Z \times K.$$

Theorem

Suppose that $Z = \operatorname{zer}(A+B) \neq \varnothing$. Let $x \in X$ and let $(z,k) \in \mathcal{S}$. Then the following hold:

(i) *For every $n \in \mathbb{N}$, we have* $\|(J_A T^{n+1}x, J_{A^{-1}} T^{n+1}x) - (z,k)\|^2 \leq \|(J_A T^n x, J_{A^{-1}} T^n x) - (z,k)\|^2$, *i..e., $(J_A T^n x, J_{A^{-1}} T^n x)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $\mathcal{S}$.*

(ii) *$(J_A T^n x, J_{A^{-1}} T^n x)_{n \in \mathbb{N}}$ converges weakly to some point in $\mathcal{S}$.*

Proof.

(i): We have $z + k \in \operatorname{Fix} T$ (details omitted). Therefore $(\forall n \in \mathbb{N})$
$(z,k) = (J_A(z+k), J_{A^{-1}}(z+k))$

Shadows convergence: A simplified proof

Recall that the so-called Kuhn–Tucker set is defined by

$$\mathcal{S} := \mathcal{S}_{(A,B)} := \{(z,k) \in X \times X \mid -k \in Bz, k \in Az\} \subseteq Z \times K.$$

Theorem

Suppose that $Z = \operatorname{zer}(A+B) \neq \varnothing$. Let $x \in X$ and let $(z,k) \in \mathcal{S}$. Then the following hold:

(i) *For every $n \in \mathbb{N}$, we have*
$\|(J_A T^{n+1}x, J_{A^{-1}} T^{n+1}x) - (z,k)\|^2 \leq \|(J_A T^n x, J_{A^{-1}} T^n x) - (z,k)\|^2$,
i..e., $(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\mathcal{S}$.

(ii) *$(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in $\mathcal{S}$.*

Proof.

(i): We have $z + k \in \operatorname{Fix} T$ (details omitted). Therefore $(\forall n \in \mathbb{N})$
$(z,k) = (J_A(z+k), J_{A^{-1}}(z+k)) = (J_A T^n(z+k), J_{A^{-1}} T^n(z+k))$.

Shadows convergence: A simplified proof

Recall that the so-called Kuhn–Tucker set is defined by

$$\mathcal{S} := \mathcal{S}_{(A,B)} := \{(z,k) \in X \times X \mid -k \in Bz, k \in Az\} \subseteq Z \times K.$$

Theorem

Suppose that $Z = \operatorname{zer}(A+B) \neq \varnothing$. Let $x \in X$ and let $(z,k) \in \mathcal{S}$. Then the following hold:

(i) *For every $n \in \mathbb{N}$, we have* $\|(J_A T^{n+1}x, J_{A^{-1}} T^{n+1}x) - (z,k)\|^2 \leq \|(J_A T^n x, J_{A^{-1}} T^n x) - (z,k)\|^2$, *i..e., $(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\mathcal{S}$.*

(ii) *$(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in $\mathcal{S}$.*

Proof.

(i): We have $z + k \in \operatorname{Fix} T$ (details omitted). Therefore $(\forall n \in \mathbb{N})$ $(z,k) = (J_A(z+k), J_{A^{-1}}(z+k)) = (J_A T^n(z+k), J_{A^{-1}} T^n(z+k))$. Apply $\|(J_A Tx, J_{A^{-1}} Tx) - (J_A Ty, J_{A^{-1}} Ty)\|^2 \leq \|(J_A x, J_{A^{-1}} x) - (J_A y, J_{A^{-1}} y)\|^2$ with (x, y) replaced by $(T^n x, z+k)$.

Shadows convergence: A simplified proof

Recall that the so-called Kuhn–Tucker set is defined by

$$\mathcal{S} := \mathcal{S}_{(A,B)} := \{(z,k) \in X \times X \mid -k \in Bz, k \in Az\} \subseteq Z \times K.$$

Theorem

Suppose that $Z = \operatorname{zer}(A+B) \neq \varnothing$. Let $x \in X$ and let $(z,k) \in \mathcal{S}$. Then the following hold:

(i) *For every $n \in \mathbb{N}$, we have* $\|(J_A T^{n+1}x, J_{A^{-1}} T^{n+1}x) - (z,k)\|^2 \leq \|(J_A T^n x, J_{A^{-1}} T^n x) - (z,k)\|^2$, *i..e., $(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\mathcal{S}$.*

(ii) *$(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in $\mathcal{S}$.*

Proof.

(i): We have $z+k \in \operatorname{Fix} T$ (details omitted). Therefore $(\forall n \in \mathbb{N})$ $(z,k) = (J_A(z+k), J_{A^{-1}}(z+k)) = (J_A T^n(z+k), J_{A^{-1}} T^n(z+k))$. Apply $\|(J_A Tx, J_{A^{-1}} Tx) - (J_A Ty, J_{A^{-1}} Ty)\|^2 \leq \|(J_A x, J_{A^{-1}} x) - (J_A y, J_{A^{-1}} y)\|^2$ with (x,y) replaced by $(T^n x, z+k)$. (ii): We prove the weak cluster points of the bounded sequence $(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ lie in $\mathcal{S}$ (details omitted). Now combine with (i) and use the classical Fejér monotonicity principle. □

References

Attouch, H., and Théra, M. (1996). A general duality principle for the sum of two operators, *Journal of Convex Analysis* 3:1–24.

Bauschke, H.H. and Combettes, P.L. (2017). *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer.

Attouch, H., and Théra, M. (1996). A general duality principle for the sum of two operators, *Journal of Convex Analysis* 3:1–24.

Bauschke, H.H. and Moursi, W.M. (2017). On the Douglas–Rachford algorithm. *Math. Program. (Ser. A)*. 164:263–284

Bauschke, H.H. and Moursi, W.M. (2016). The Douglas–Rachford algorithm for two (not necessarily intersecting) affine subspaces. *SIAM J. Optim.*, 26:968–985.

Bauschke, H.H., Hare, W.L., and Moursi, W.M. (2016). On the Range of the Douglas–Rachford operator. *Math. Oper. Res.*, 41:884–879.

Bauschke, H.H., Dao, M.N., and Moursi, W.M. (2016). The Douglas–Rachford algorithm in the affine-convex case. *Oper. Res. Lett.*, 44:379–382.

Bauschke, H.H., Hare, W.L., and Moursi, W.M. (2014). Generalized solutions for the sum of two maximally monotone operators. *SIAM J. Control Optim.*, 52:1034–1047.

Combettes, P.L. (2004). Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization* 53:475–504.

Eckstein, J. (1989). *Splitting Methods for Monotone Operators with Applications to Parallel Optimization*, Ph.D. thesis, MIT.

THANK YOU !!