On the behavior of the Douglas–Rachford algorithm in possibly nonconvex settings

Minh N. Dao

Priority Research Centre for Computer-Assisted Research Mathematics and its Applications (CARMA)

THE UNIVERSITY OF NEWCASTLE AUSTRALIA

Splitting Algorithms, Modern Operator Theory, and Applications

Oaxaca, September 17–22, 2017

Based on joint works with

Heinz H. Bauschke	University of British Columbia Okanagan
Scott B. Lindstrom	University of Newcastle
Walaa M. Moursi	University of British Columbia Okanagan
Dominikus Noll	University of Toulouse
Hung M. Phan	University of Massachusetts Lowell
Matthew K. Tam	University of Göttingen

Outline

Outline

1. Introduction
2. Behavior of DR algorithm in possibly inconsistent case
3. Finite convergence
4. A Lyapunov-type approach to convergence theory
5. Local linear convergence

Fundamental feasibility problem

Unless stated otherwise,

$$X : \text{a real Hilbert space},$$
$$A, B : \text{closed (possibly nonconvex) subsets of } X.$$

The fundamental feasibility problem asks to

$$\text{find } x \in A \cap B.$$

In the inconsistent case, i.e., $A \cap B = \varnothing$, it can be naturally formulated as finding a best approximation pair relative to A and B:

$$\text{find } (a, b) \in A \times B \text{ such that } \|a - b\| = \inf \|A - B\|.$$

- arises in a wide range of applications including image recovery and encoding algorithms.
- is typically approached by projection methods which combine projectors and their variants in a suitable way to generate a sequence converging to a solution of the problem.

Fundamental feasibility problem

Unless stated otherwise,

$$X : \text{a real Hilbert space},$$
$$A, B : \text{closed (possibly nonconvex) subsets of } X.$$

The fundamental feasibility problem asks to

$$\text{find } x \in A \cap B.$$

In the inconsistent case, i.e., $A \cap B = \varnothing$, it can be naturally formulated as finding a best approximation pair relative to A and B:

$$\text{find} \quad (a, b) \in A \times B \quad \text{such that} \quad \|a - b\| = \inf \|A - B\|.$$

- arises in a wide range of applications including image recovery and encoding algorithms.
- is typically approached by projection methods which combine projectors and their variants in a suitable way to generate a sequence converging to a solution of the problem.

Projectors and relaxed projectors

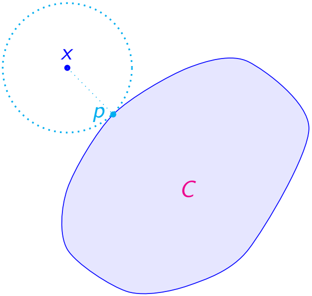
Let C be a nonempty closed set in X and $\lambda \in \mathbb{R}_+$. The projector onto C is

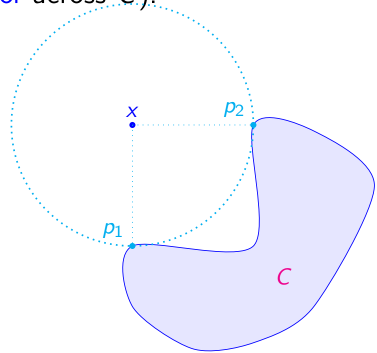
$$P_C \colon X \rightrightarrows C \colon x \mapsto P_C x := \operatorname{argmin}_{c \in C} \|x - c\|.$$

The relaxed projector for C with parameter λ is defined by

$$P_C^\lambda := (1-\lambda)\operatorname{Id} + \lambda P_C.$$

- $P_C^0 = \operatorname{Id}$, $P_C^1 = P_C$,
- $P_C^2 = R_C := 2P_C - \operatorname{Id}$ (the reflector across C).





Projectors and relaxed projectors

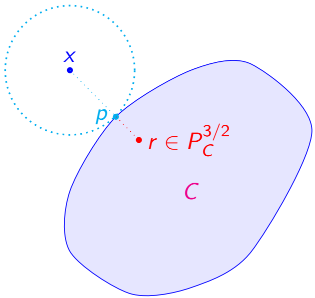
Let C be a nonempty closed set in X and $\lambda \in \mathbb{R}_+$. The projector onto C is

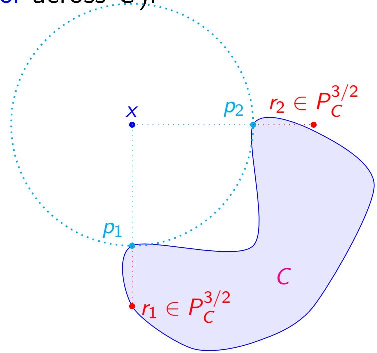
$$P_C \colon X \rightrightarrows C \colon x \mapsto P_C x := \operatorname{argmin}_{c \in C} \|x - c\|.$$

The relaxed projector for C with parameter λ is defined by

$$P_C^\lambda := (1-\lambda)\operatorname{Id} + \lambda P_C.$$

- $P_C^0 = \operatorname{Id}$, $P_C^1 = P_C$,
- $P_C^2 = R_C := 2P_C - \operatorname{Id}$ (the reflector across C).





Douglas–Rachford (DR) algorithm

Use the DR operator

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

to generate a DR sequence $(z_n)_{n\in\mathbb{N}}$ by

$$(\forall n \in \mathbb{N}) \quad z_{n+1} \in T_{A,B} z_n, \quad \text{where } z_0 \in X.$$

- $T_{A,B}$ is single-valued when A and B are convex.
- $z \in \operatorname{Fix} T_{A,B} := \{z \in X \mid z \in T_{A,B} z\} \quad \Rightarrow \quad (\exists a \in P_A z)\ a \in A \cap B.$

Douglas–Rachford (DR) algorithm

Use the DR operator

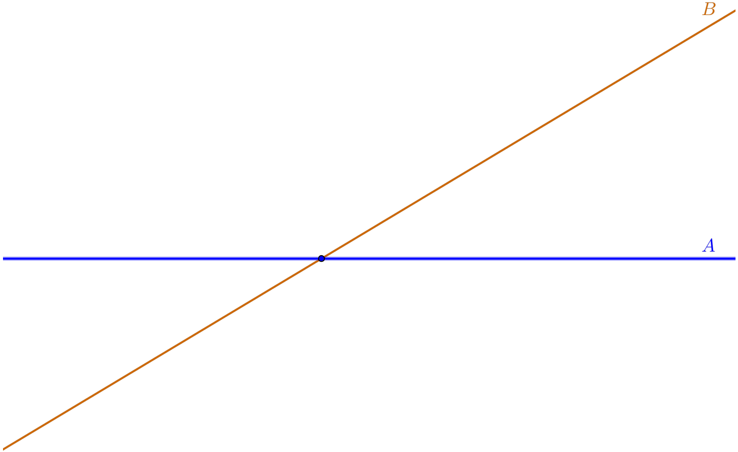
$$T_{A,B} := \frac{1}{2}(\mathrm{Id} + R_B R_A)$$

to generate a DR sequence $(z_n)_{n\in\mathbb{N}}$ by

$$(\forall n \in \mathbb{N}) \quad z_{n+1} \in T_{A,B} z_n, \quad \text{where } z_0 \in X.$$

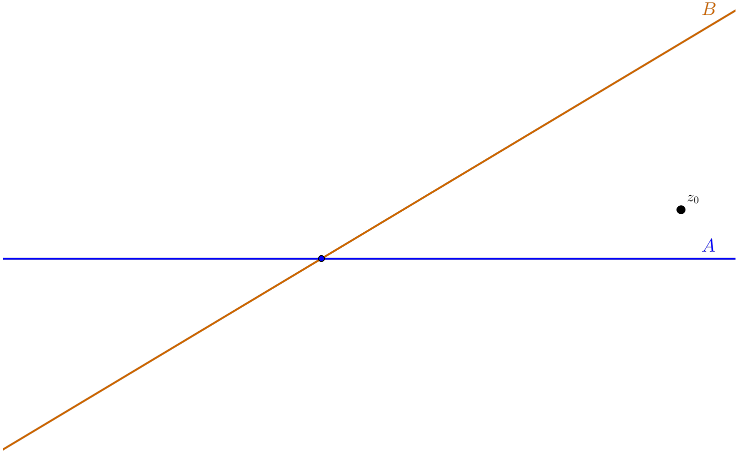
- $T_{A,B}$ is single-valued when A and B are convex.
- $z \in \operatorname{Fix} T_{A,B} := \{z \in X \mid z \in T_{A,B} z\} \quad \Rightarrow \quad (\exists a \in P_A z)\ a \in A \cap B$.

Reflect, reflect, average. . .



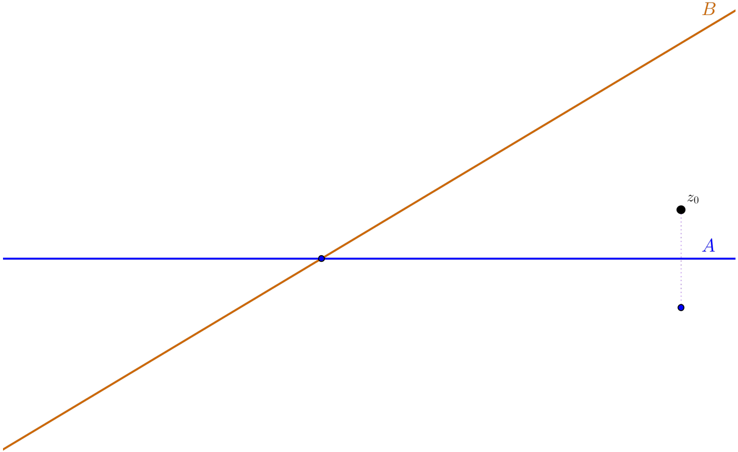
While $(z_n)_{n\in\mathbb{N}}$ "spirals" towards the origin, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ occasionally gets very close to the origin!

Reflect, reflect, average. . .



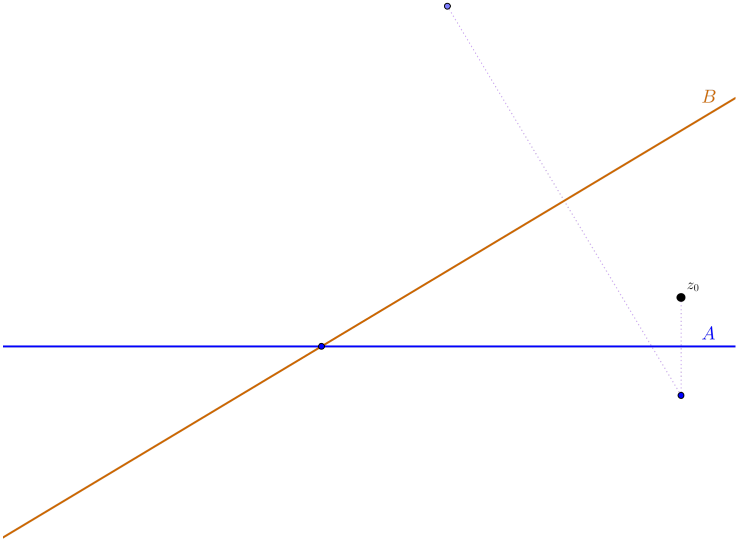
While $(z_n)_{n\in\mathbb{N}}$ "spirals" towards the origin, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ occasionally gets very close to the origin!

Reflect, reflect, average. . .



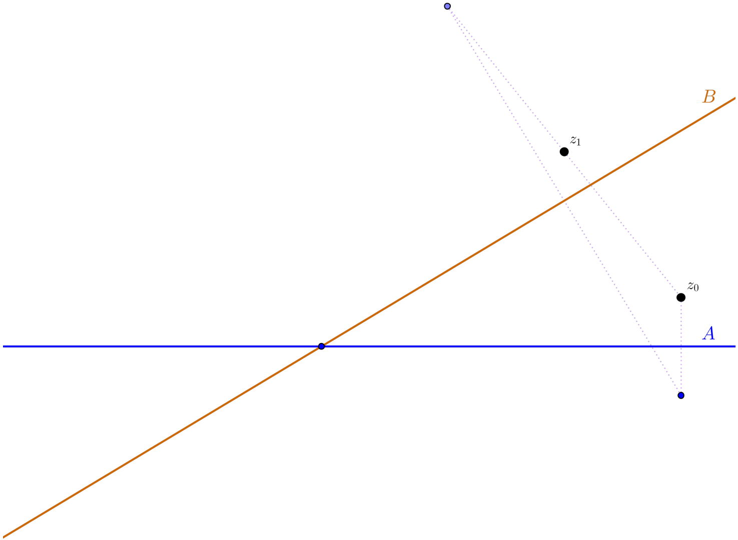
While $(z_n)_{n\in\mathbb{N}}$ "spirals" towards the origin, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ occasionally gets very close to the origin!

Reflect, reflect, average. . .



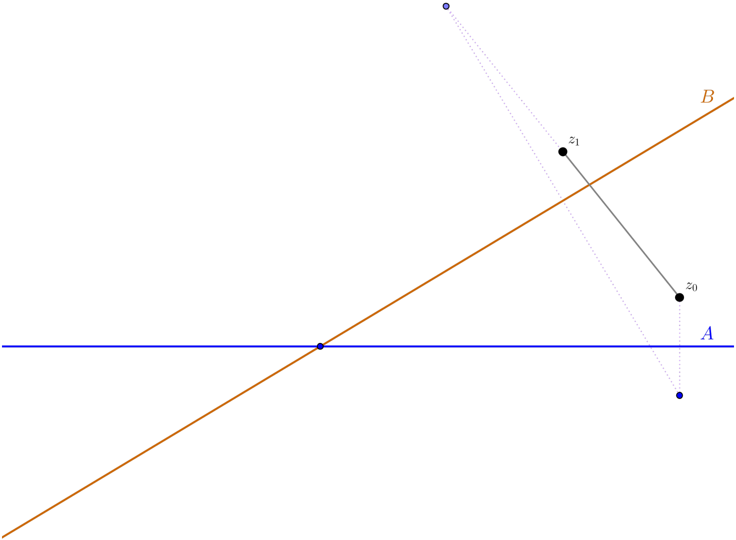
While $(z_n)_{n\in\mathbb{N}}$ "spirals" towards the origin, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ occasionally gets very close to the origin!

Reflect, reflect, average. . .



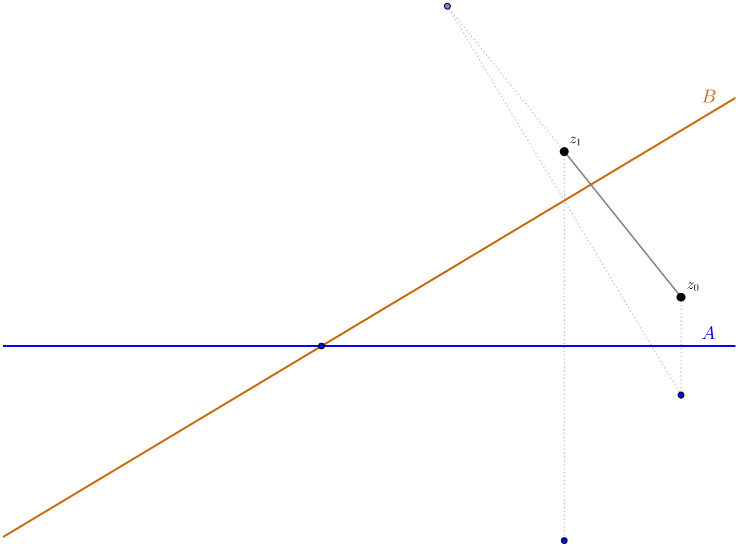
While $(z_n)_{n\in\mathbb{N}}$ "spirals" towards the origin, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ occasionally gets very close to the origin!

Reflect, reflect, average. . .



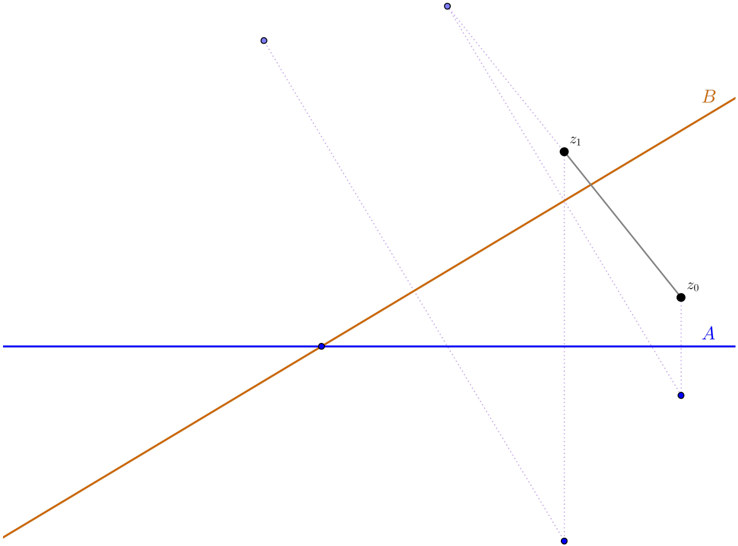
While $(z_n)_{n\in\mathbb{N}}$ "spirals" towards the origin, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ occasionally gets very close to the origin!

Reflect, reflect, average. . .



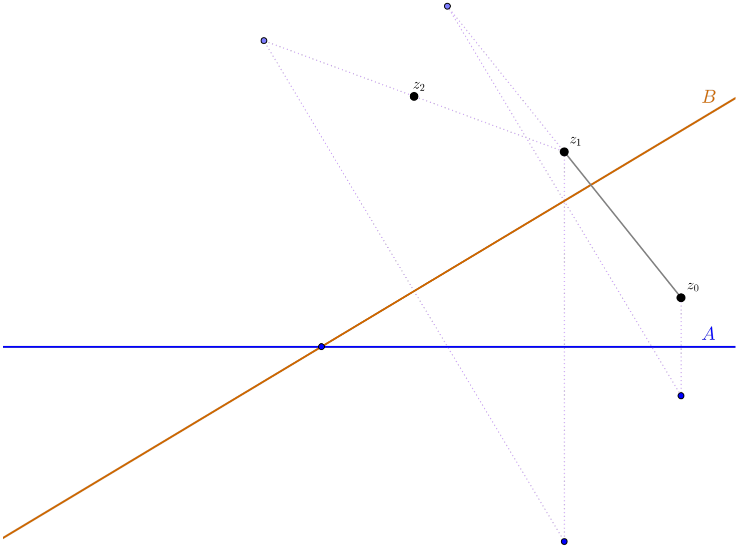
While $(z_n)_{n\in\mathbb{N}}$ "spirals" towards the origin, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ occasionally gets very close to the origin!

Reflect, reflect, average. . .



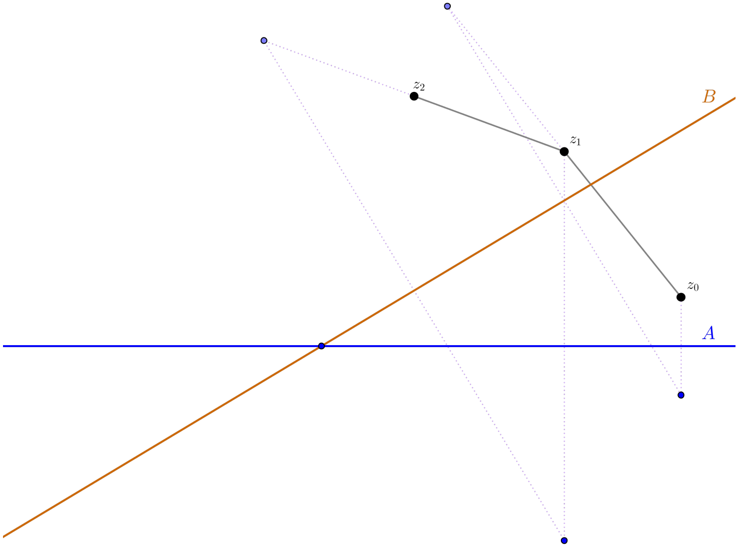
While $(z_n)_{n\in\mathbb{N}}$ "spirals" towards the origin, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ occasionally gets very close to the origin!

Reflect, reflect, average. . .



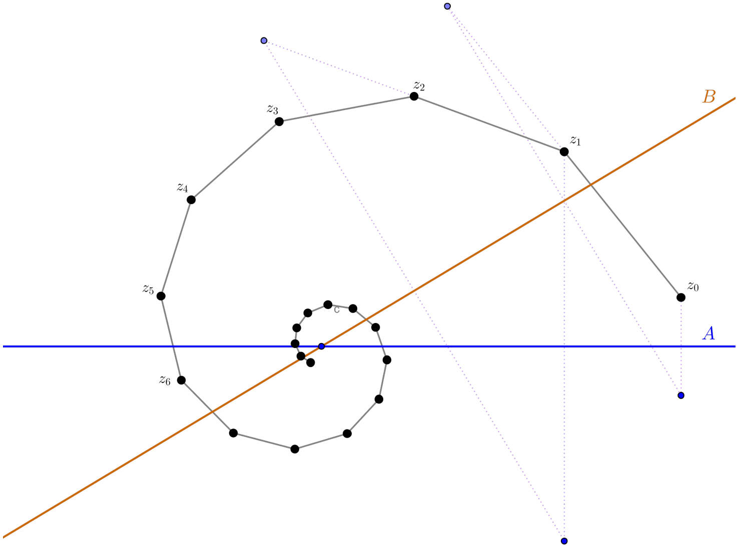
While $(z_n)_{n\in\mathbb{N}}$ "spirals" towards the origin, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ occasionally gets very close to the origin!

Reflect, reflect, average. . .



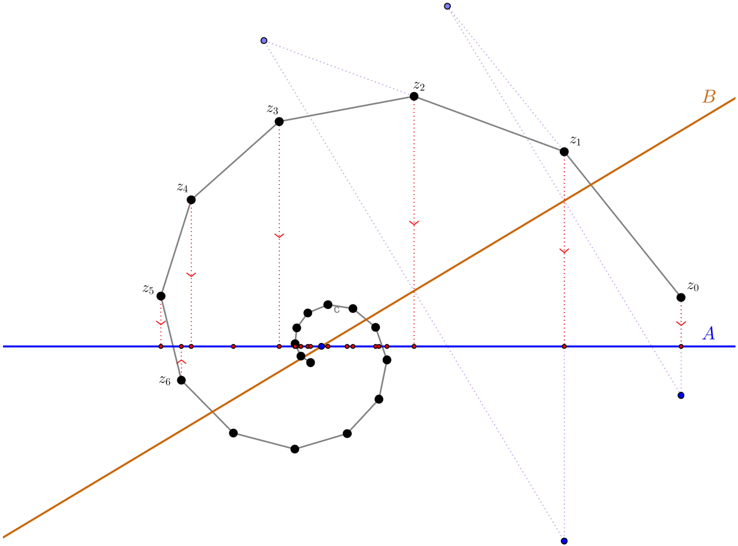
While $(z_n)_{n\in\mathbb{N}}$ "spirals" towards the origin, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ occasionally gets very close to the origin!

Reflect, reflect, average...



While $(z_n)_{n\in\mathbb{N}}$ "spirals" towards the origin, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ occasionally gets very close to the origin!

Reflect, reflect, average. . .



While $(z_n)_{n\in\mathbb{N}}$ "spirals" towards the origin, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ occasionally gets very close to the origin!

Outline

Convex-convex case

Suppose that A and B are convex.

Fact (Lions–Mercier 1979, Svaiter 2011, Bauschke–Combettes–Luke 2004)

If $A \cap B \neq \varnothing$, then $z_n \rightharpoonup z \in \operatorname{Fix} T = (A \cap B) + N_{A-B}(0)$ and $P_A z_n \rightharpoonup P_A z \in A \cap B$; otherwise, $\|z_n\| \to +\infty$.

Now assume that $g := P_{\overline{B-A}} 0 \in B - A$, or equivalently,

$$E := A \cap (B - g) \neq \varnothing \quad \text{and} \quad F := (A + g) \cap B \neq \varnothing.$$

Fact (Bauschke–Combettes–Luke 2004)

The sequence $(P_A z_n)_{n \in \mathbb{N}}$ is bounded and its weak cluster points lie in E.

Here $N_{A-B}(0)$ is the normal cone of the set $A - B = \{a - b \mid a \in A, b \in B\}$.

Affine-convex case

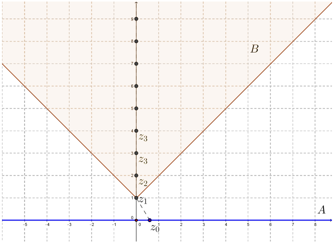
Theorem (Bauschke–D–Moursi 2016, When A is a closed affine subspace)

1. $P_A z_n \rightharpoonup a \in E = A \cap (B - g)$.
2. *No general conclusion can be drawn about the sequence* $(P_B z_n)_{n\in\mathbb{N}}$.

This is strengthened by Bauschke–Moursi 2017 for A convex.

Example ($X = \mathbb{R}^2$, $A = \mathbb{R} \times \{0\}$, $B = \text{epi}(|\cdot| + 1)$)

For $z_0 \in [-1, 1] \times \{0\}$, $z_n = (0, n) \in B$, and $\|P_B z_n\| = \|z_n\| = n \to \infty$.



Theorem (Bauschke–D–Moursi 2016, When B is a closed affine subspace)

1. $P_A z_n \rightharpoonup a \in E = A \cap (B - g)$.
2. $P_B z_n \rightharpoonup b \in F = (A + g) \cap B$.

Affine-convex case

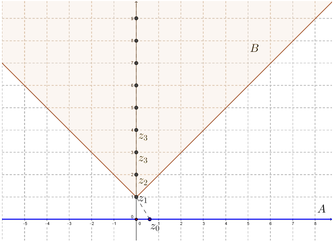
Theorem (Bauschke–D–Moursi 2016, When A is a closed affine subspace)

1. $P_A z_n \rightharpoonup a \in E = A \cap (B - g)$.
2. *No general conclusion can be drawn about the sequence* $(P_B z_n)_{n\in\mathbb{N}}$.

This is strengthened by Bauschke–Moursi 2017 for A convex.

Example ($X = \mathbb{R}^2$, $A = \mathbb{R} \times \{0\}$, $B = \text{epi}(|\cdot| + 1)$)

For $z_0 \in [-1, 1] \times \{0\}$, $z_n = (0, n) \in B$, and $\|P_B z_n\| = \|z_n\| = n \to \infty$.



Theorem (Bauschke–D–Moursi 2016, When B is a closed affine subspace)

1. $P_A z_n \rightharpoonup a \in E = A \cap (B - g)$.
2. $P_B z_n \rightharpoonup b \in F = (A + g) \cap B$.

Spingarn's method

Consider the problem to find a *least-squares solution* of $\bigcap_{j=1}^{M} C_j$, i.e., to

$$\text{find minimizers of } \sum_{j=1}^{M} d_{C_j}^2, \tag{1}$$

where $C_1, \ldots, C_M$ are nonempty closed convex (possibly nonintersecting) subsets of X with corresponding distance functions $d_{C_1}, \ldots, d_{C_M}$.

Set $\mathbf{X} := X^M$, $\mathbf{A} := \{(x, \ldots, x) \in \mathbf{X} \mid x \in X\}$, and $\mathbf{B} := C_1 \times \cdots \times C_M$. Assume that $\mathbf{g} = (g_1, \ldots, g_M) := P_{\overline{\mathbf{B}-\mathbf{A}}} 0 \in \mathbf{B} - \mathbf{A}$. Then

$$\mathbf{E} := \mathbf{A} \cap (\mathbf{B} - \mathbf{g}) \neq \varnothing, \quad \text{and } (x, \ldots, x) \in \mathbf{A} \cap (\mathbf{B} - \mathbf{g}) \Leftrightarrow x \in \bigcap_{j=1}^{M} (C_j - g_j).$$

Corollary

Let $(\mathbf{z}_n)_{n\in\mathbb{N}}$ *be a DR sequence for* $(\mathbf{A}, \mathbf{B})$. *Then*

$$P_{\mathbf{A}} \mathbf{z}_n \rightharpoonup \mathbf{z} = (z, \ldots, z) \in \mathbf{A} \cap (\mathbf{B} - \mathbf{g}),$$

where $z \in \bigcap_{j=1}^{M} (C_j - g_j)$ *and z is a least-squares solution of* (1).

This generalizes work by Spingarn (1987) who considered only halfspaces and whose proof was much more complicated.

Spingarn's method

Consider the problem to find a *least-squares solution* of $\bigcap_{j=1}^{M} C_j$, i.e., to

$$\text{find minimizers of } \sum_{j=1}^{M} d_{C_j}^2, \tag{1}$$

where $C_1, \ldots, C_M$ are nonempty closed convex (possibly nonintersecting) subsets of X with corresponding distance functions $d_{C_1}, \ldots, d_{C_M}$.

Set $\mathbf{X} := X^M$, $\mathbf{A} := \{(x, \ldots, x) \in \mathbf{X} \mid x \in X\}$, and $\mathbf{B} := C_1 \times \cdots \times C_M$.
Assume that $\mathbf{g} = (g_1, \ldots, g_M) := P_{\overline{\mathbf{B}-\mathbf{A}}} 0 \in \mathbf{B} - \mathbf{A}$. Then

$$\mathbf{E} := \mathbf{A} \cap (\mathbf{B} - \mathbf{g}) \neq \varnothing, \quad \text{and } (x, \ldots, x) \in \mathbf{A} \cap (\mathbf{B} - \mathbf{g}) \Leftrightarrow x \in \bigcap_{j=1}^{M} (C_j - g_j).$$

Corollary

Let $(\mathbf{z}_n)_{n\in\mathbb{N}}$ be a DR sequence for $(\mathbf{A}, \mathbf{B})$. Then

$$P_{\mathbf{A}} \mathbf{z}_n \rightharpoonup \mathbf{z} = (z, \ldots, z) \in \mathbf{A} \cap (\mathbf{B} - \mathbf{g}),$$

where $z \in \bigcap_{j=1}^{M} (C_j - g_j)$ and z is a least-squares solution of (1).

This generalizes work by Spingarn (1987) who considered only halfspaces and whose proof was much more complicated.

Spingarn's method

Consider the problem to find a *least-squares solution* of $\bigcap_{j=1}^{M} C_j$, i.e., to

$$\text{find minimizers of } \sum_{j=1}^{M} d_{C_j}^2, \tag{1}$$

where $C_1, \dots, C_M$ are nonempty closed convex (possibly nonintersecting) subsets of X with corresponding distance functions $d_{C_1}, \dots, d_{C_M}$.

Set $\mathbf{X} := X^M$, $\mathbf{A} := \{(x, \dots, x) \in \mathbf{X} \mid x \in X\}$, and $\mathbf{B} := C_1 \times \cdots \times C_M$. Assume that $\mathbf{g} = (g_1, \dots, g_M) := P_{\overline{\mathbf{B}-\mathbf{A}}} 0 \in \mathbf{B} - \mathbf{A}$. Then

$$\mathbf{E} := \mathbf{A} \cap (\mathbf{B} - \mathbf{g}) \neq \varnothing, \quad \text{and } (x, \dots, x) \in \mathbf{A} \cap (\mathbf{B} - \mathbf{g}) \Leftrightarrow x \in \bigcap_{j=1}^{M} (C_j - g_j).$$

Corollary

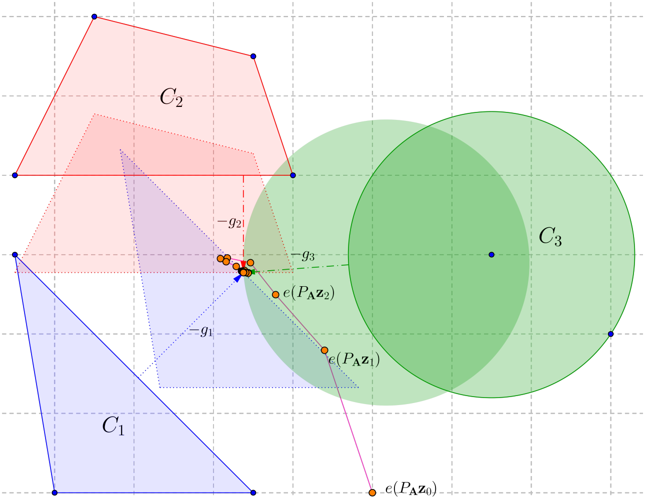
Let $(\mathbf{z}_n)_{n\in\mathbb{N}}$ *be a DR sequence for* $(\mathbf{A}, \mathbf{B})$*. Then*

$$P_{\mathbf{A}} \mathbf{z}_n \rightharpoonup \mathbf{z} = (z, \dots, z) \in \mathbf{A} \cap (\mathbf{B} - \mathbf{g}),$$

where $z \in \bigcap_{j=1}^{M} (C_j - g_j)$ *and* z *is a least-squares solution of* (1).

This generalizes work by Spingarn (1987) who considered only halfspaces and whose proof was much more complicated.

Find a point in the generalized intersection



Three closed convex sets are shown along with their translations forming the generalized intersection. The first few terms of $(e(P_{\mathbf{A}}\mathbf{z}_n))_{n\in\mathbb{N}}$ (yellow points) are also depicted. Here $e : \mathbf{A} \to \mathbb{R}^2 : (x, x, x) \mapsto x$.

Outline

Presence of Slater's condition ($A \cap \operatorname{int} B \neq \varnothing$)

From now on, X is finite-dimensional.

Lemma

If A and B are convex and $0 \in \operatorname{int}(A - B)$, then $z_n \to z \in A \cap B$; the convergence is finite provided that $z \in A \cap \operatorname{int} B$.

Theorem (Bauschke–D–Noll–Phan 2016, Bauschke–D 2017)

Suppose that $A \cap \operatorname{int} B \neq \varnothing$. Then the DR algorithm converges finitely to a point in $A \cap B$ in each of the following cases:

1. *A is an affine subspace and B is a polyhedron.*
2. *$A \in \{X \times \{0\}, X \times \mathbb{R}_+, X \times \mathbb{R}_-\}$ and $B = \operatorname{epi} f$, where $f\colon X \to \left]-\infty, +\infty\right]$ is convex, l.s.c., and proper.*
3. *A is a hyperplane/halfspace and B is a finite intersection of closed balls B_j such that $(\forall x \in A \cap \operatorname{bdry} B)(\exists! B_j)\ x \in \operatorname{bdry} B_j$.*

int C: the interior of C.

Absence of Slater's condition

- In the case of an affine subspace and a polyhedron, if the Slater's condition is replaced by "$A \cap \operatorname{ri} B \neq \varnothing$", then finite convergence fails in general, e.g., the case of two lines in $\mathbb{R}^2$.
- If $A \in \{X \times \{0\}, X \times \mathbb{R}_-\}$ and $B = \operatorname{epi} f$, where $\inf_X f \geq 0$ and f is differentiable at its minimizers, then $(P_A z_n)_{n\in\mathbb{N}}$ and hence $(z_n)_{n\in\mathbb{N}}$ do not converge finitely whenever $z_0 = (x_0, \rho_0) \in B$ with $x_0 \notin \operatorname{argmin} f$.

Theorem (Bauschke–D 2017)

Suppose that A is a hyperplane/halfspace and that $A \cap B \neq \varnothing$. Then the DR sequence converges finitely to a point in $z \in \operatorname{Fix} T_{A,B}$ with $P_A z \in A \cap B$ in each of the following cases:

1. *B is a halfspace of X.*
2. *$X = \mathbb{R}^2$, and B is a polyhedral set.*

---

ri C: the interior of C relative to the affine hull of C.

Absence of Slater's condition

- In the case of an affine subspace and a polyhedron, if the Slater's condition is replaced by "$A \cap \operatorname{ri} B \neq \varnothing$", then finite convergence fails in general, e.g., the case of two lines in $\mathbb{R}^2$.
- If $A \in \{X \times \{0\}, X \times \mathbb{R}_-\}$ and $B = \operatorname{epi} f$, where $\inf_X f \geq 0$ and f is differentiable at its minimizers, then $(P_A z_n)_{n\in\mathbb{N}}$ and hence $(z_n)_{n\in\mathbb{N}}$ do not converge finitely whenever $z_0 = (x_0, \rho_0) \in B$ with $x_0 \notin \operatorname{argmin} f$.

Theorem (Bauschke–D 2017)

Suppose that A is a hyperplane/halfspace and that $A \cap B \neq \varnothing$. Then the DR sequence converges finitely to a point in $z \in \operatorname{Fix} T_{A,B}$ with $P_A z \in A \cap B$ in each of the following cases:

1. *B is a halfspace of X.*
2. *$X = \mathbb{R}^2$, and B is a polyhedral set.*

---

ri C: the interior of C relative to the affine hull of C.

When one set is finite

Suppose that B is a finite subset of X and let $(z_n)_{n\in\mathbb{N}}$ be a DR sequence for (A, B).

Theorem (Bauschke–D 2017)

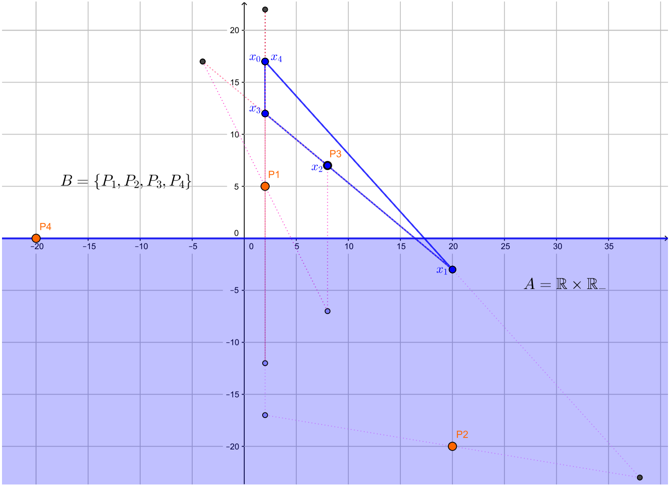
If A is an affine subspace/a halfspace, $A \cap B \neq \varnothing$, and the sequence $(z_n)_{n\in\mathbb{N}}$ is asymptotically regular, i.e., $z_n - z_{n+1} \to 0$, then $(z_n)_{n\in\mathbb{N}}$ converges in finitely many steps to a point $z \in \operatorname{Fix} T_{A,B}$ with $P_A z \in A \cap B$.

Theorem (Bauschke–D 2017)

If A is a hyperplane/halfspace and B is contained in one of two halfspaces generated by A, then either

1. *$(z_n)_{n\in\mathbb{N}}$ converges finitely to a point $z \in \operatorname{Fix} T_{A,B}$ with $P_A z \in A \cap B$, or*
2. *$A \cap B = \varnothing$ and $\|z_n\| \to +\infty$ in which case $(P_A z_n)_{n\in\mathbb{N}}$ converges finitely to a best approximation solution $a \in A$ relative to A and B.*

Without asymptotic regularity or “one-side” property



A 4-cycle of the DR algorithm for a halfspace and a finite set. Interchanging the roles of two sets gives finite convergence, as shown by Aragón Artacho–Borwein–Tam 2016.

Periodic behavior

Theorem (Bauschke–D–Lindstrom 2017)

Suppose that A is a hyperplane and that $B = \{b_1, b_2\}$, where b_1 and b_2 do not belong to the same halfspace generated by A. Let $(z_n)_{n\in\mathbb{N}}$ be a DR sequence for (A, B). Then

1. *$(z_n)_{n\in\mathbb{N}}$ does not converge.*
2. *$(z_n)_{n\in\mathbb{N}}$ cycles after certain steps regardless the starting point if and only if there exist $k_1, k_2 \in \mathbb{N} \smallsetminus \{0\}$ such that $k_1 d_A(b_1) = k_2 d_A(b_2)$.*

Outline

The case of a line and a circle

Fact (Borwein–Sims 2011)

Let $\alpha \in [0,1[$. Then the DR algorithm for

$$A = \mathbb{R} \times \{\alpha\} \quad \textit{and} \quad B = \left\{(x,\rho) \in \mathbb{R}^2 \;\middle|\; x^2 + \rho^2 = 1\right\}$$

is locally convergent around $(\pm\sqrt{1-\alpha^2}, \alpha)$.

Conjecture (BS11): The DR algorithm is actually globally convergent. This has since been resolved in the affirmative by Benoist (2015).

Idea: Consider $V\colon \mathbb{R}^2 \to]-\infty, +\infty]$ given by

$$V(x,\rho) := \tfrac{1}{2}x^2 - (1-\alpha)\ln|x| + \alpha\sqrt{1-x^2} - \alpha\ln(1+\sqrt{1-x^2}) + \tfrac{1}{2}(\rho-\alpha)^2.$$

Then V decreases along DR sequences: $V(T_{A,B}z) \le V(z)$ with equality if and only if $z \in \operatorname{Fix} T_{A,B}$.

The case of a line and a circle

Fact (Borwein–Sims 2011)

Let $\alpha \in [0,1[$. Then the DR algorithm for

$$A = \mathbb{R} \times \{\alpha\} \quad \text{and} \quad B = \left\{(x,\rho) \in \mathbb{R}^2 \;\middle|\; x^2 + \rho^2 = 1\right\}$$

is locally convergent around $(\pm\sqrt{1-\alpha^2}, \alpha)$.

Conjecture (BS11): The DR algorithm is actually globally convergent. This has since been resolved in the affirmative by Benoist (2015).

Idea: Consider $V\colon \mathbb{R}^2 \to\,]-\infty,+\infty]$ given by

$$V(x,\rho) := \tfrac{1}{2}x^2 - (1-\alpha)\ln|x| + \alpha\sqrt{1-x^2} - \alpha\ln(1+\sqrt{1-x^2}) + \tfrac{1}{2}(\rho-\alpha)^2.$$

Then V decreases along DR sequences: $V(T_{A,B}z) \le V(z)$ with equality if and only if $z \in \operatorname{Fix} T_{A,B}$.

Finding a zero of a function

In this section,

$$f\colon X \to [-\infty, +\infty] \text{ is proper with closed graph.}$$

Consider the feasibility problem in $X \times \mathbb{R}$ with constraints

$$A = X \times \{0\} \quad \text{and} \quad B = \operatorname{gra} f := \big\{(x,\rho) \in X \times \mathbb{R} \;\big|\; f(x) = \rho\big\},$$

which can be cast as

$$\text{find } x \in X \text{ such that } f(x) = 0.$$

- B is generally not convex unless f is affine.
- **For a line and a circle:** Up to symmetry, take $f(x) = -\sqrt{1-x^2} + \alpha$.

Finding a zero of a function

In this section,

$$f\colon X \to [-\infty, +\infty] \text{ is proper with closed graph.}$$

Consider the feasibility problem in $X \times \mathbb{R}$ with constraints

$$A = X \times \{0\} \quad \text{and} \quad B = \operatorname{gra} f := \big\{(x, \rho) \in X \times \mathbb{R} \;\big|\; f(x) = \rho\big\},$$

which can be cast as

$$\text{find } x \in X \text{ such that } f(x) = 0.$$

- B is generally not convex unless f is affine.
- **For a line and a circle:** Up to symmetry, take $f(x) = -\sqrt{1 - x^2} + \alpha$.

A Lyapunov-type approach

Definition (Lyapunov-type function)

A function $V\colon X\times\mathbb{R}\to\,]-\infty,+\infty]$ is a Lyapunov-type function for f on a nonempty convex subset D of X if it can be expressed in the form

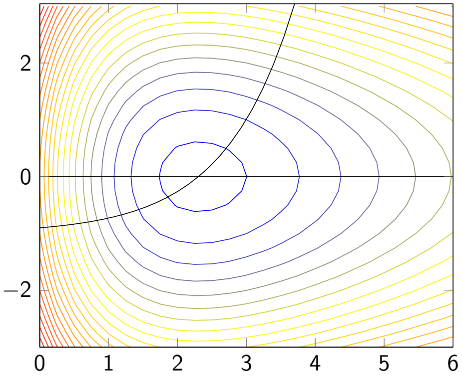
$$V(x,\rho)=F(x)+\frac{1}{2}\rho^2$$

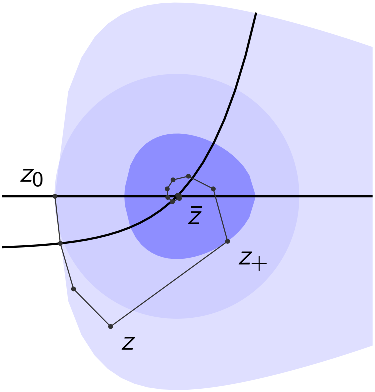
for some proper coercive convex function $F\colon D\to\,]-\infty,+\infty]$ whose subdifferential satisfies

$$(\forall x\in D)\quad \partial F(x)\supseteq\begin{cases}\left\{\frac{f(x)}{\|x^*\|^2}x^*\ \middle|\ x^*\in\partial^0 f(x)\right\} & \text{if } 0\notin\partial^0 f(x),\\ \{0\} & \text{if } f(x)=0.\end{cases}$$

$\partial^0 f:=\partial f\cup-\partial(-f)$: the symmetric (limiting) subdifferential of f.

A Lyapunov-type approach: Some intuition





A Lyapunov-type function for $f(x) = \frac{1}{10}\exp(x) - 1$, which guarantees global convergence of the DR algorithm to $\bar{z} := (\ln(10), 0)$.

Convergence theorem

Write $z_n = (x_n, \rho_n) \in X \times \mathbb{R}$. Suppose that there exists a Lyapunov-type function for f on D, that f is locally Lipschitz continuous on $D \smallsetminus f^{-1}(0)$, and that

$$(\exists n_0 \in \mathbb{N})(\forall n \geq n_0) \quad x_n \in D \quad \text{and} \quad x_{n+1} \notin (\partial^0 f)^{-1}(0) \smallsetminus f^{-1}(0).$$

Theorem (D–Tam 2017)

The DR sequence $(z_n)_{n\in\mathbb{N}}$ is bounded and asymptotically regular, and each of its cluster points $\bar{z}$ satisfy $P_A\bar{z} \in A \cap B$. Suppose, in addition, that $\overline{D} \cap f^{-1}(0) = \{\bar{x}\}$ is contained in D. Then

1. $z_n \to \bar{z}$ *with* $P_A\bar{z} \in A \cap B$.
2. $z_n \to \bar{z} = (\bar{x}, 0) \in A \cap B$ *provided that* $0 \notin \partial^0 f(\bar{x})$ *and* $f|_D$ *is continuous at* $\bar{x}$.

Linear convergence

Corollary

Suppose that $\overline{D} \cap f^{-1}(0) = \{\bar{x}\} \subseteq D$. *Then*

1. *If f is continuously differentiable around* $\bar{x}$ *with* $\nabla f(\bar{x}) \neq 0$, *then* $z_n \to \bar{z} = (\bar{x}, 0) \in A \cap B$ *with R-linear rate.*
2. *If* $X = \mathbb{R}$ *and f is twice strictly differentiable at* $\bar{x}$ *with* $f'(\bar{x}) \neq 0$, *then* $z_n \to \bar{z} = (\bar{x}, 0) \in A \cap B$ *with Q-linear rate*

$$\kappa := \frac{1}{\sqrt{1 + |f'(\bar{x})|^2}}.$$

A sequence $(z_n)_{n\in\mathbb{N}}$ is said to converge to a point $\bar{z}$

- with *R-linear rate* $\kappa \in [0,1[$ if $(\exists \eta \in \mathbb{R}_+)(\forall n \in \mathbb{N})\ \|z_n - \bar{z}\| \leq \eta\kappa^n$;
- with *Q-linear rate* $\kappa \in [0,1[$ if $\limsup_{n\to\infty} \frac{\|z_{n+1}-\bar{z}\|}{\|z_n-\bar{z}\|} \leq \kappa$.

Some examples

- $X = \mathbb{R}$, $f(x) = \alpha \exp(x) - \beta$ with $(\alpha, \beta) \in \mathbb{R}^2_{++}$. One possible F is

$$F(x) := \int \frac{f(x)}{f'(x)} dx = \int \left(1 - \frac{\beta}{\alpha} \exp(-x)\right) dx = x + \frac{\beta}{\alpha} \exp(-x)$$

$\longrightarrow$ Global Q-linear convergence with rate $\kappa = 1/\sqrt{1+\beta^2}$.

- $X = \mathbb{R}$, $p \in]1, +\infty[$,

$$f(x) := \begin{cases} x^p & \text{if } x \geq 0, \\ x & \text{if } x < 0, \end{cases} \qquad \partial^0 f(x) = \begin{cases} px^{p-1} & \text{if } x \geq 0, \\ [0,1] & \text{if } x = 0, \\ 1 & \text{if } x < 0. \end{cases}$$

Note that f is nonconvex and nonsmooth at $x = 0$. Define F by

$$F(x) := \begin{cases} \frac{1}{2p} x^2 & \text{if } x \geq 0, \\ \frac{1}{2} x^2 & \text{if } x < 0, \end{cases}$$

then V is a Lyapunov-type function for f on $\mathbb{R}$.

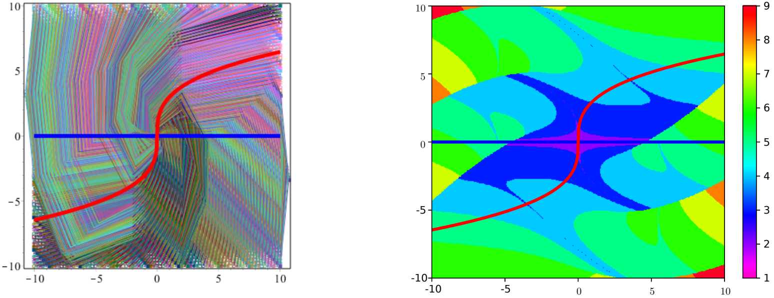
Some examples

Suppose $f = \alpha\|\cdot\|^p$ for $\alpha \in \mathbb{R} \smallsetminus \{0\}$ and $p \in]0, +\infty[$. Then whenever $x \neq 0$, we have $\partial^0 f(x) = \{\alpha p\|x\|^{p-2}x\}$ and

$$\frac{f(x)}{\|\nabla f(x)\|}\nabla f(x) = \frac{\alpha\|x\|^p}{\alpha^2 p^2\|x\|^{2p-2}}\alpha p\|x\|^{p-2}x = \frac{1}{p}x,$$

which leads to $F(x) = \frac{1}{2p}\|x\|^2$. The global convergence follows.

The same function F works for $f = \alpha|\cdot|^p\,\mathrm{sgn}(\cdot)$ on $\mathbb{R}$.



Illustrations of the DR algorithm for $f(x) = 3\sqrt[3]{x}$ on $[-10, 10] \times [-10, 10]$.

Outline

1 Introduction

2 Behavior of DR algorithm in possibly inconsistent case

3 Finite convergence

4 A Lyapunov-type approach to convergence theory

5 Local linear convergence

Generalized DR operator

Let $\lambda, \mu \in]0, 2]$ and $\alpha \in]0, 1]$. The generalized DR operator for (A, B) with parameters (λ, μ, α) is defined by

$$T_{\lambda,\mu}^{\alpha} := (1-\alpha)\,\mathrm{Id} + \alpha P_B^{\mu} P_A^{\lambda}.$$

- $T_{1,1}^{1} = P_B P_A$ is the classical alternating projection (AP) operator.
- $T_{2,2}^{1/2} = \frac{1}{2}(\mathrm{Id} + R_B R_A)$ is the classical DR operator.
- $T_{2,2\alpha}^{1/2} = (1-\alpha)P_A + \frac{\alpha}{2}(\mathrm{Id} + R_B R_A)$ is the relaxed avaraged alternating reflection (RAAR) operator.
- If B is an affine subspace of X, then

$$T_{1+\alpha,1+\alpha}^{1/(1+\alpha)} = (1-\alpha)P_B P_A + \frac{\alpha}{2}(\mathrm{Id} + R_B R_A)$$

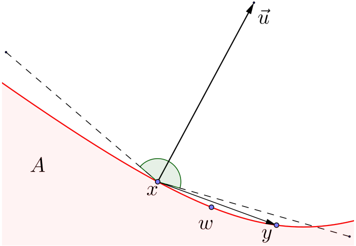
(a convex combination of the classical AP and DR operators).

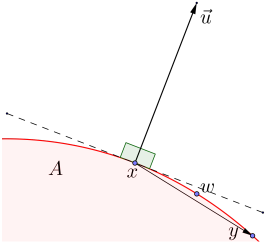
Regularity of sets

Let $\varepsilon \in \mathbb{R}_+$ and $\delta \in \mathbb{R}_{++}$. A set C is said to be (ε, δ)-regular at $w \in X$ if

$$\forall x, y \in C \cap \mathbb{B}_\delta(w), \forall u \in N_C^{\text{prox}}(x) : \quad \langle u, y - x \rangle \le \varepsilon \|u\| \cdot \|y - x\|$$

and superregular at w if $\forall \varepsilon \in \mathbb{R}_{++}, \exists \delta \in \mathbb{R}_{++}$: C is (ε, δ)-regular at w.





- Convex sets and sets with "smooth" boundary are superregular.

$N_C^{\text{prox}}(x) := \text{cone}(P_C^{-1}(x) - x) = \left\{ \lambda(z - x) \;\middle|\; z \in P_C^{-1}(x),\ \lambda \in \mathbb{R}_+ \right\}.$

Key properties

Lemma

Let $\varepsilon_1 \in [0, 1/3]$, $\varepsilon_2 \in [0, 1[$ *and set* $\gamma := 1 - \alpha + \alpha\big(1 + \frac{\lambda\varepsilon_1}{1-\varepsilon_1}\big)\big(1 + \frac{\mu\varepsilon_2}{1-\varepsilon_2}\big)$, $\beta := \frac{1-\alpha}{\alpha}$. *If* A *and* B *are* (ε_1, δ)*- and* $(\varepsilon_2, \sqrt{2}\delta)$*-regular at* $w \in A \cap B$, *then* $T^{\alpha}_{\lambda,\mu}$ *is* $(A \cap B \cap \mathbb{B}_\delta(w), \gamma, \beta)$*-quasi firmly Fejér monotone on* $\mathbb{B}_{\delta/2}(w)$ *in the sense that*

$$\forall x \in \mathbb{B}_{\delta/2}(w),\ \forall x_+ \in T^{\alpha}_{\lambda,\mu}x,\ \forall \overline{x} \in A \cap B \cap \mathbb{B}_\delta(w):$$
$$\|x_+ - \overline{x}\|^2 + \beta\|x - x_+\|^2 \le \gamma\|x - \overline{x}\|^2.$$

Lemma

Let $\varepsilon \in [0, 1/3]$. *If* A *is superregular at* w *and* $\{A, B\}$ *is strongly regular at* $w \in A \cap B$, *then there exist* $\delta \in \mathbb{R}_{++}$, $\nu \in \mathbb{R}_{++}$ *such that* $T^{\alpha}_{\lambda,\mu}$ *is* $(A \cap B, \nu)$*-quasi coercive on* $\mathbb{B}_{\delta/2}(w)$ *in the sense that*

$$\forall x \in \mathbb{B}_{\delta/2}(w),\ \forall x_+ \in T^{\alpha}_{\lambda,\mu}x: \quad \|x - x_+\| \ge \nu d_{A\cap B}(x).$$

$\{A, B\}$ is strongly regular at $w \in A \cap B$ if $N_A(w) \cap (-N_B(w)) = \{0\}$.

Local linear convergence

Let A and B be closed subsets of X with $A \cap B \neq \varnothing$. Suppose that $\{A, B\}$ is superregular and strongly regular at some point $w \in A \cap B$.

Fact (Phan 2016)

When started at a point sufficiently close to w, the DR sequence converges R-linearly to a point in $A \cap B$.

Theorem (D–Phan 2016)

Let $\lambda, \mu \in \,]0, 2]$ and $\alpha \in \,]0, 1[$. Then when started at a point sufficiently close to w, the generalized DR sequence generated by $T^{\alpha}_{\lambda,\mu}$ converges R-linearly to a point in $A \cap B$.

Some key references

- F.J. Aragón Artacho, J.M. Borwein, and M.K. Tam, Global behavior of the Douglas–Rachford method for a nonconvex feasibility problem, *Journal of Global Optimization* 65 (2016), 309–327.
- H.H. Bauschke, P.L. Combettes, and D.R. Luke, Finding best approximation pairs relative to two closed convex sets in Hilbert spaces, *Journal of Approximation Theory* 127 (2004), 178–192.
- J. Benoist, The Douglas–Rachford algorithm for the case of the sphere and the line, *Journal of Global Optimization* 63 (2015), 363–380.
- J.M. Borwein and B. Sims, The Douglas–Rachford algorithm in the absence of convexity, In: *Fixed-point Algorithms for Inverse Problems in Science and Engineering*, Springer, 2011, pp 93–109.
- R. Hesse and D.R. Luke, Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems, *SIAM Journal on Optimization* 23 (2013), 2397–2419.
- H.M. Phan, Linear convergence of the Douglas–Rachford method for two closed sets, *Optimization* **65** (2016), 369–385.

Contact

The manuscripts corresponding to this talk are

- H.H. Bauschke and MND, On the finite convergence of the Douglas–Rachford algorithm for solving (not necessarily convex) feasibility problems in Euclidean spaces, *SIAM Journal on Optimization* 27 (2017), 507–537.
- H.H. Bauschke, MND, and S.B.. Lindstrom, On the behavior of the Douglas–Rachford iteration, In preparation.
- H.H. Bauschke, MND, and W.M. Moursi, The Douglas–Rachford algorithm in the affine-convex case, *Operations Research Letters* 44 (2016), 379–382.
- H.H. Bauschke, MND, D. Noll, and H.M. Phan, On Slater's condition and finite convergence of the Douglas–Rachford algorithm for solving convex feasibility problems in Euclidean spaces, *Journal of Global Optimization* 65 (2016), 329–349.
- MND and H.M. Phan, Linear convergence of projection algorithms, (2016), arXiv:1609.00341.
- MND and M.K. Tam, A Lyapunov-type approach to convergence of the Douglas–Rachford algorithm, (2017), arXiv:1706.04846.

THANK YOU VERY MUCH!

daonminh@gmail.com / minh.dao@newcastle.edu.au