A SURVEY OF RESULTS ON LINEAR CONVERGENCE FOR ITERATIVE PROXIMAL ALGORITHMS IN NONCONVEX SETTINGS

Russell Luke
(Heinz Bauschke, Charitha, Robert Hesse, Adrian Lewis, Jerôme Malick, Patrick Neumann, Hung Phan, and Shawn Wang)

Universität Göttingen

Limoges, May 19, 2015

Outline

Happy 80th year Terry!

Goals

Solve

$$0 \in F(x)$$

for $F : \mathbb{E} \rightrightarrows \mathbb{E}$ with $\mathbb{E}$ a Euclidean space.

- #1. Convergence (with rates and radii) of Picard iterations:

$$x^{k+1} \in Tx^k$$

- #2. Algorithms:
  - Feasibility: alternating projections
  - Optimization: Douglas-Rachford
- #2. Applications:
  - Ptychography/Phase retrieval
  - Rank constrained affine feasibility
  - Optimization with high-dimensional structured constraints

Building blocks

- Resolvent: $(\mathrm{Id} + \lambda F)^{-1}$
- Prox operator: for a function $f : X \to \overline{\mathbb{R}}$, define

$$\mathrm{prox}_{\eta f}(x) := \mathrm{argmin}_{\,y}\, f(y) + \frac{1}{2\eta}\|y - x\|^2$$

- Proximal reflector: $R_{\eta f} := 2\,\mathrm{prox}_{\eta f} - \mathrm{Id}$
- Projector: if $f = \iota_\Omega$ for $\Omega \subset X$ closed and nonempty, then $\mathrm{prox}_{\eta f}(\overline{x}) = P_\Omega \overline{x}$ where

$$\begin{aligned} P_\Omega x &:= \{\overline{x} \in \Omega \mid \|x - \overline{x}\| = \mathrm{dist}\,(x, \Omega)\} \\ \mathrm{dist}\,(x, \Omega) &:= \inf_{y \in \Omega} \|x - y\|. \end{aligned}$$

- Reflector: if $f = \iota_\Omega$ for some closed, nonempty set $\Omega \subset X$, then $R_\Omega := 2P_\Omega - \mathrm{Id}$

Algorithms

- Proximal point: $T_{PP} := (\mathrm{Id} + \lambda F)^{-1}$ (fixed step length)

and for $F = \partial f_1 + \partial f_2$, splitting algorithms:

- Forward-backward/Projected (sub)gradients:
  $T_{FB} := \mathrm{prox}_{f_1}(\mathrm{Id} - \lambda \partial f_2)$
- Backward-backward/Alternating projections: $T_{AP} := \mathrm{prox}_{f_1}\mathrm{prox}_{f_2}$
- Douglas-Rachford (ADMM):
  $T_{DR} := \mathrm{prox}_{f_1}\left(2\,\mathrm{prox}_{f_2} - \mathrm{Id}\right) - \mathrm{prox}_{f_2} + \mathrm{Id}$

Convergence analysis

Monotone/Convex

- Monotone proximal point: Martinet, Rokafellar
- Convex forward-backward: Eckstein, Combettes-Wajs, Combettes-Pesquet, Combettes-Dung-Vu, Nesterov, Beck-Teboulle
- Convex alternating projections: von Neumann, Aronszajn, Cheney-Goldstein, Bregman, Gubin-Polyak-Raik
- Convex Douglas-Rachford/ADMM: Lions-Mercier, Gabay

Nonmonotone/Nonconvex

- Nonmonotone proximal point: Spingarn, Pennanen, Combettes-Pennanen, Iusem-Pennanen-Svaiter, Aragón-Geoffroy, Aragón-Dontchev-Geoffroy, Attouch-Bolte
- Nonconvex alternating projections: Combettes-Trussell, Lewis-Malick, Lewis-L.-Malick, Bauschke-L.-Phan-Wang, **Hesse-L.**, Hesse-L.-Neumann, **Noll-Rondepierre**
- Nonconvex forward-backward: Attouch-Bolte-Redont-Soubeyran
- Nonconvex Douglas-Rachford: Borwein-Sims, **Hesse-L.**, Hesse-L.Neumann, Phan, Bauschke-Noll

Outline

Ambient regularity

(S, ϵ)-nonexpansive mappings

Let D and S be nonempty subsets of $\mathbb{E}$ and let T be a (multi-valued) mapping from D to $\mathbb{E}$.

- T is called (S, ε)-nonexpansive on D if, at each $\overline{x} \in S$, for all $x \in D$

$$\|x_+ - \overline{x}_+\| \leq \sqrt{1 + \varepsilon}\, \|x - \overline{x}\|$$
$$\forall\, \overline{x}_+ \in T\overline{x} \text{ and } \forall\, x_+ \in Tx.$$

- If this holds with $\epsilon = 0$ then T is called S-nonexpansive on D.
- If this holds with $\epsilon = 0$ and $S = D$ then T is called nonexpansive on D.

(S, ϵ)-firmly nonexpansive mappings

Let D and S be nonempty subsets of $\mathbb{E}$ and let T be a (multi-valued) mapping from D to $\mathbb{E}$.

- T is called (S, ε)-firmly nonexpansive on D if, at each $\overline{x} \in S$, for all $x \in D$

$$\|x_+ - \overline{x}_+\|^2 + \|(x - x_+) - (\overline{x} - \overline{x}_+)\|^2 \leq (1 + \varepsilon) \|x - \overline{x}\|^2$$
$$\forall\, \overline{x}_+ \in T\overline{x} \text{ and } \forall x_+ \in Tx.$$

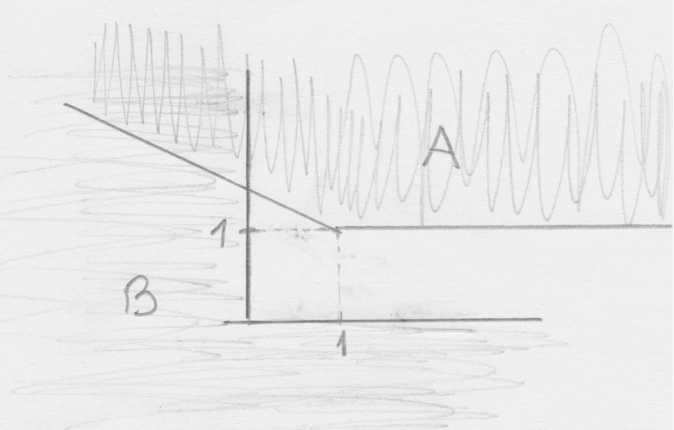
- If this holds with $\epsilon = 0$ then T is called S-firmly nonexpansive on D.
- If this holds with $\epsilon = 0$ and $S = D$ then T is called firmly nonexpansive on D.

Fixed points of set-valued mappings

The set of fixed points of a set-valued mapping $T : X \rightrightarrows X$ is defined by

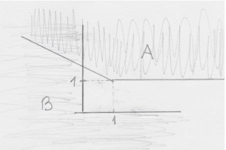
$$\{x \in X \mid x \in T(x)\}.$$

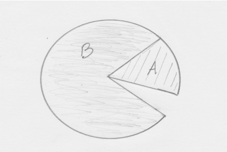
When $T := P_A P_B$ this includes the point $\{(1,1)\}$ in



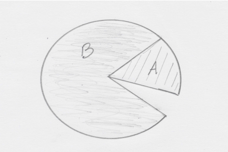
Alternating Projections: $T = P_A P_B$

(i) IS NOT (Fix T, ϵ)-firmly nonexpansive on A for any ϵ

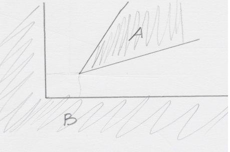




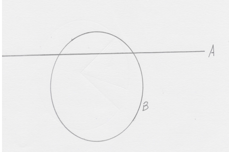
(ii) IS (S, ϵ)-firmly nonexpansive on A for



$(S, \epsilon) = (\{(0,0)\}, 0)$



$(S, \epsilon) = (\{(1,1)\}, 0)$



$(S, \epsilon) = (\{a^*\}, 1)$

$\square$

Regularity at the fixed point set

Basic Lemma

Let $T : D \rightrightarrows \mathbb{E}$ for $D \subset \mathbb{E}$ and let $S \subset \operatorname{Fix} T \subset \operatorname{ri} D$. Define $S_\delta := \delta\mathbb{B} + S \subset D$ for $\delta \in [0, \overline{\delta})$ fixed. Suppose

(a) T is (S, ϵ)-firmly nonexpansive on $S_{\overline{\delta}}$, and

(b) there exists $\lambda > 0$ such that T satisfies

$$\|x - x_+\| \geq \lambda \operatorname{dist}(x, S) \quad \forall\, x_+ \in Tx,\ \forall\, x \in S_{\overline{\delta}} \setminus \left(S_{\overline{\delta}} \cap (\operatorname{Fix} T + \delta\mathbb{B})\right). \tag{1}$$

Then, for all $x \in S_{\overline{\delta}} \setminus \left(S_{\overline{\delta}} \cap (\operatorname{Fix} T + \delta\mathbb{B})\right)$

$$\operatorname{dist}(x_+, \operatorname{Fix} T) \leq \sqrt{1 + \epsilon - \lambda^2}\, \operatorname{dist}(x, S) \quad \forall\, x_+ \in Tx. \tag{2}$$

(sub) Linear Convergence

Let $T : D \rightrightarrows D$ for $D \subset \mathbb{E}$ and let $S \subset \mathrm{Fix}\ T \subset \mathrm{ri}\, D$ be closed and nonempty. Define $S_\delta := (\delta\mathbb{B} + S) \cap D$. Suppose that for $\gamma \in [0, 1)$ fixed and for all $\overline{\delta} > 0$ small enough, there is a triplet $(\epsilon, \delta, \lambda) \in \mathbb{R}_+ \times [0, \gamma\overline{\delta}] \times (\sqrt{\epsilon}, \sqrt{\epsilon + 1}]$ such that

(a) T is (S, ϵ)-firmly nonexpansive on $S_{\overline{\delta}}$ and

(b) $\|x - x_+\| \geq \lambda\, \mathrm{dist}(x, S) \quad \forall x_+ \in Tx,\ \forall x \in S_{\overline{\delta}} \setminus \left(S_{\overline{\delta}} \cap (\mathrm{Fix}\ T + \delta\mathbb{B})\right)$.

Then for any x^0 close enough to S the iterates $x^{k+1} \in T(x^k)$ satisfy $\mathrm{dist}(x^k, \mathrm{Fix}\ T) \to 0$ as $k \to \infty$.

Condition (a) is relatively easy to verify. What about (b)?

The key insight into condition (b) is the connection to metric SUBregularity of set-valued mappings. This approach to the study of algorithms has been advanced by several authors [Pennanen, Klatte&Kummer, Aragón&Geoffroy, Dontchev&Rockafellar]

(strong) Metric (sub)-Regularity

(i) The mapping $\Phi : \mathbb{E} \rightrightarrows \mathbb{Y}$ is called metrically regular at of order w $\overline{x}$ for $\overline{y}$ if there is a constant $\kappa > 0$ together with neighborhood U of $\overline{x}$ and V of $\overline{y}$ such that

$$\operatorname{dist}\left(x, \Phi^{-1}(y)\right) \leq \kappa \operatorname{dist}^{w}\left(y, \Phi(x)\right) \quad \forall (x, y) \in U \times V. \tag{3}$$

The regularity modulus is the infimum of those constants $\kappa > 0$ such that (3) holds.

(ii) The mapping $\Phi : \mathbb{E} \rightrightarrows \mathbb{Y}$ is called metrically subregular of order w at $\overline{x}$ for $\overline{y}$ if $(\overline{x}, \overline{y}) \in \operatorname{gph} \Phi$ and there is a constant $\kappa > 0$ and neighborhoods U of $\overline{x}$ and V of $\overline{y}$ such that

$$\operatorname{dist}\left(x, \Phi^{-1}(\overline{y})\right) \leq \kappa \operatorname{dist}^{w}\left(\overline{y}, \Phi(x) \cap V\right) \quad \forall\, x \in U. \tag{4}$$

The subregularity modulus is the infimum of those constants $\kappa > 0$ such that (4) holds.

(sub) Linear Convergence with Metric Subregularity

Let $T : D \rightrightarrows D$ for $D \subset \mathbb{E}$, $\Phi := T - \mathrm{Id}$ and let $S \subset \mathrm{Fix}\ T \subset \mathrm{ri}\, D$ be closed and nonempty. Suppose that, for any $\overline{\delta} > 0$ small enough, there are $\gamma \in (0,1)$, $w \in (0,1]$, a nonnegative sequence of scalars $(\epsilon_i)_{i\in\mathbb{N}}$ and a constant $\kappa > 0$, such that, for all $i \in \mathbb{N}$, $S_{\overline{\delta}} := \left(\overline{\delta}\mathbb{B} + S\right) \cap D$,

$$\frac{1}{\sqrt{\epsilon_i + 1}} \leq \kappa \inf_{x \in S_{\gamma^i \overline{\delta}} \setminus \left(\left(\mathrm{Fix}\ T + \gamma^{i+1}\overline{\delta}\mathbb{B}\right) \cap S_{\gamma^i \overline{\delta}}\right)} \left\{ \inf_{x_+ \in Tx} \|x_+ - x\|^{w-1} \right\} < \frac{1}{\sqrt{\epsilon_i}}, \tag{5}$$

and

(a) T is (S, ϵ_i)-firmly nonexpansive on $S_{\gamma^i \overline{\delta}}$ and

(b) $\mathrm{dist}\left(x, \Phi^{-1}(0)\right) \leq \kappa\, \mathrm{dist}^w\left(0, \Phi(x)\right) \quad \forall x \in S_{\gamma^i \overline{\delta}} \setminus \left(\left(\mathrm{Fix}\ T + \gamma^{i+1}\overline{\delta}\mathbb{B}\right) \cap S_{\gamma^i \overline{\delta}}\right)$.

(sub) Linear Convergence with Metric Subregularity

Then, for any $x^0 \in S_{\overline{\delta}}$, the iterates $x^{j+1} \in T(x^j)$ converge to $\mathrm{Fix}\ T$ with

$$\mathrm{dist}\left(x^{j+1}, \mathrm{Fix}\ T\right) \leq \sqrt{1 + \epsilon_i - \left(\frac{1}{\kappa_i}\right)^2}\,\mathrm{dist}\left(x^j, S\right) \tag{6}$$

for all $x^j \in S_{\gamma^i \overline{\delta}} \setminus \left(\left(\mathrm{Fix}\ T + \gamma^{i+1}\overline{\delta}\mathbb{B}\right) \cap S_{\gamma^i \overline{\delta}}\right)$ where

$$\kappa_i := \kappa \inf_{x \in S_{\gamma^i \overline{\delta}} \setminus \left(\left(\mathrm{Fix}\ T + \gamma^{i+1}\overline{\delta}\mathbb{B}\right) \cap S_{\gamma^i \overline{\delta}}\right)} \left\{ \inf_{x_+ \in Tx} \|x_+ - x\|^{w-1} \right\}. \tag{7}$$

In particular, if $w = 1$, then convergence is at least linear with rate bounded above by $\sqrt{1 + \overline{\epsilon} - \left(\frac{1}{\kappa}\right)^2}$ where $\overline{\epsilon} := \sup_k \{\epsilon_k\}$.

Outline

Projection algorithms for feasibility

Regularity of sets

(i) A nonempty set $\Omega \subset \mathbb{E}$ is *(ε, δ)-subregular* at $\widehat{x}$ with respect to $S \subset \mathbb{E}$ if

$$\langle v, \overline{x} - y \rangle \leq \varepsilon \, \|v\| \, \|\overline{x} - y\| \tag{8}$$

holds for all $y \in \mathbb{B}_\delta(\widehat{x}) \cap \Omega$, $\overline{x} \in S \cap \mathbb{B}_\delta(\widehat{x})$, $v \in N^{\mathrm{prox}}_\Omega(y)$. The set Ω is said to be *(ε, δ)-subregular at $\widehat{x}$* if $S = \{\widehat{x}\}$.

(ii) If $S = \Omega$ in (i) then the set Ω is said to be *(ε, δ)-regular* at $\widehat{x}$.

(iii) A nonempty (locally) closed set $\Omega \subset \mathbb{E}$ is *Clarke regular* at a point $\overline{x} \in \Omega$ if, for all $\epsilon > 0$, $\exists\ \delta > 0$ such that $z \in \mathbb{B}_\delta(\widehat{x})$ and $x \in \Omega \cap \mathbb{B}_\delta(\widehat{x})$, and any $y \in P_\Omega(z)$,

$$\left\langle x - \widehat{x}, z - y \right\rangle \leq \epsilon \|x - \widehat{x}\| \|z - y\|.$$

(iv) If, for all $\epsilon > 0$, there exists a $\delta > 0$ such that (8) holds for all $y, \overline{x} \in \mathbb{B}_\delta(\widehat{x}) \cap \Omega$ and $v \in N^{\mathrm{prox}}_\Omega(y)$, then Ω is said to be *super-regular*.

(v) A nonempty (locally) closed set $\Omega \subset \mathbb{E}$ is *prox-regular* at a point $\overline{x} \in \Omega$ if the projector P_Ω is single-valued in a vicinity of $\overline{x}$.

Set regularity relations

(i) $\{(\epsilon, \delta) - \text{subregular sets}\} \supset \{(\epsilon, \delta) - \text{regular sets}\}$

(ii) $\{(\epsilon, \delta) - \text{regular sets}\} \supset \{\text{Clarke regular sets}\}$

(iii) $\{\text{Clarke regular sets}\} \supset \{\text{super-regular sets}\}$

(iv) $\{\text{super-regular sets}\} \supset \{\text{prox-regular sets}\}$

(v) $\{\text{prox-regular sets}\} \supset \{\text{closed convex sets}\}$

A more recent definition of 0-Hölder regular sets (relative to some other set) developed by Noll&Rondepierre (2015) satisfies

$$\{\text{0-Hölder regular sets}\} \supset \{(\epsilon, \delta) - \text{regular sets}\}$$

but haven't worked out the relation to $\{(\epsilon, \delta)\}$-subregular sets yet.

Projectors and reflectors onto (ε, δ)-subregular sets

Let $\Omega \subset \mathbb{E}$ be nonempty closed and (ε, δ)-subregular at $\widehat{x}$ with respect to $S \subseteq \Omega \cap \mathbb{B}_\delta(\widehat{x})$ and define

$$U := \left\{ x \in \mathbb{E} \;\middle|\; P_\Omega x \subset \mathbb{B}_\delta(\widehat{x}) \right\}. \tag{9}$$

- **(i)** The projector P_Ω is $(S, \widetilde{\varepsilon}_1)$-nonexpansive on U where $\widetilde{\varepsilon}_1 := 2\varepsilon + \varepsilon^2$.
- **(ii)** The projector P_Ω is $(S, \widetilde{\varepsilon}_2)$-firmly nonexpansive on $\mathbb{B}_\delta(\widehat{x})$ where $\widetilde{\varepsilon}_2 := 2\varepsilon + 2\varepsilon^2$.
- **(iii)** The reflector R_Ω is $(S, \widetilde{\varepsilon}_3)$-nonexpansive on $\mathbb{B}_\delta(\widehat{x})$ where $\widetilde{\varepsilon}_3 := 4\varepsilon + 4\varepsilon^2$.

Regularity of collections of sets

Linearly focusing collections of sets

The collection $\{\Omega_1, \Omega_2\}$ of closed subsets of $\mathbb{E}$ is said to be *Ω_1-weakly linearly focusing* at $\overline{x}$ when $P_{\Omega_1} P_{\Omega_2} \overline{x} \ni \overline{x}$ and there is a neighborhood U of $\overline{x}$ and a constant κ such that, for all $x \in U$,

$$\operatorname{dist}(x, E_1) \leq \kappa \operatorname{dist}(x, P_{\Omega_1} P_{\Omega_2} x) \tag{10}$$

where $E_1 := \{e \in \Omega_1 \mid P_{\Omega_1} P_{\Omega_2} e \ni e\}$. The collection is said to be *weakly linearly focusing* (no mention of Ω_1 or Ω_2) at $(\overline{x}_1, \overline{x}_2)$ when $(P_{\Omega_1} P_{\Omega_2} \overline{x}_1, P_{\Omega_2} P_{\Omega_1} \overline{x}_2) \ni (\overline{x}_1, \overline{x}_2)$ and there is a neighborhood U of $(\overline{x}_1, \overline{x}_2)$ and a constant κ such that, for all $(x_1, x_2) \in U$,

$$\operatorname{dist}((x_1, x_2), (E_1, E_2)) \leq \kappa \left\| \operatorname{dist}(x_1, P_{\Omega_1} P_{\Omega_2} x), \operatorname{dist}(x_2, P_{\Omega_2} P_{\Omega_1} x) \right\|. \tag{11}$$

where $E_2 := \{f \in \Omega_2 \mid P_{\Omega_2} P_{\Omega_1} e \ni e\}$.
The collection is said to be *strongly linearly focusing* at $\overline{x}$ when $\overline{x} \in \Omega_1 \cap \Omega_2$ and there is a neighborhood U of $(\overline{x}, \overline{x})$ and a constant κ such that, for all $x \in U$ (11) holds. The infimum of all κ such that (11) holds is called the *(weak/strong) focusing modulus*.

$\{\Omega_1, \Omega_2\}$ intersect transversally

$$\Longrightarrow$$

$\{\Omega_1, \Omega_2\}$ is linearly regular at $\bar{x} \in \Omega_1 \cap \Omega_2$

$$\Longrightarrow$$

$\{\Omega_1, \Omega_2\}$ is locally linearly regular at $\bar{x} \in \Omega_1 \cap \Omega_2$

$$\Longleftrightarrow$$

$\{\Omega_1, \Omega_2\}$ is *strongly linearly focusing* at $\bar{x}$

$$\Longleftrightarrow$$

$\psi := \mathbf{T} - \mathrm{Id}$ is metrically subregular at $(\bar{x}, \bar{x}, 0, 0) \in \operatorname{gph} \psi\,(\bar{x}, \bar{x})$

$$\Longrightarrow$$

$\Phi := T - \mathrm{Id}$ is metrically subregular at $(\bar{x}, 0) \in \operatorname{gph} \Phi\,(\bar{x}, \bar{x})$

$$\Longleftrightarrow$$

$\{\Omega_1, \Omega_2\}$ is Ω_1*-weakly linearly focusing* at $\bar{x}$

where $\mathbf{T} := (P_{\Omega_1} P_{\Omega_2}, P_{\Omega_2} P_{\Omega_1})$ and $T := P_{\Omega_1} P_{\Omega_2}$.

(sub) Linear Convergence of Alternating Projections

Let $T_{AP} : A \rightrightarrows A := P_A P_B$ for A, B closed, $\Phi_{AP} := T_{AP} - \mathrm{Id}$ and let $\overline{x} \in E := \{e \in A \mid P_A P_B e \ni e\}$. Suppose that, for any $\overline{\delta} > 0$ small enough, there are $\gamma \in (0, 1)$ and a pair $(\epsilon, \kappa) > 0$ such that

$$\frac{1}{\sqrt{\epsilon + 1}} \leq \kappa < \frac{1}{\sqrt{\epsilon}}, \tag{12}$$

and

(a) T is $(\{\overline{x}\}, \epsilon)$-firmly nonexpansive on $\overline{\delta}\mathbb{B}(\overline{x}) \cap A$ and

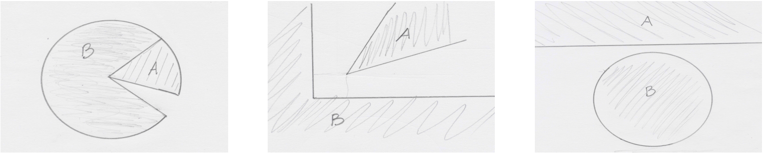
(b) (A, B) is A-weakly linearly focusing at $\overline{x}$.

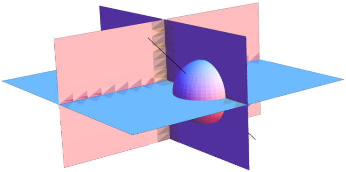
Then, for any $x^0 \in \overline{\delta}\mathbb{B}(\overline{x}) \cap A$, the iterates $x^{j+1} \in Tx^j$ converge linearly to E with rate $\sqrt{1 + \epsilon - \left(\frac{1}{\kappa}\right)^2}$.

Alternating Projections: $T_{AP} := P_A P_B$

Alternating projections converges locally linearly to Fix T_{AP}, with quantifiable rates and regions, in the following instances:

(i) any two lines in $\mathbb{R}^n$ whether they intersect or not.

(ii) 



$\square$

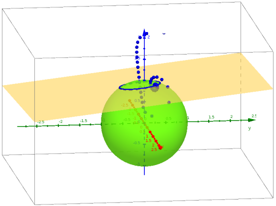
Optimization

Douglas-Rachford: $T_{DR} := \mathrm{prox}_{f_1}\left(2\,\mathrm{prox}_{f_2} - \mathrm{Id}\right) - \mathrm{prox}_{f_2} + \mathrm{Id}$

Let $W := \mathrm{aff}\{x^k\}_{k\in\mathbb{N}}$. The Douglas-Rachford algorithm converges locally linearly to $\mathrm{Fix}\ T_{DR} \cap W$ with quantifiable rates and regions, in the following instances:

(i) f_1 and f_2 are the indicator functions of sets in $\mathbb{R}^n$ with super-regular boundaries and the sets intersect essentially transversally: $T_1(\overline{x}) + T_2(\overline{x}) = W$. (For example, any two intersecting lines in $\mathbb{R}^n$.)

(ii) f_1 and f_2 are convex, piecewise linear-quadratic functions and that $\mathrm{Fix}\ T_{DR} \cap W \cap \mathcal{O} = \{\overline{x}\}$, i.e. is an isolated point. The rate of linear convergence is bounded above by $\sqrt{1-\kappa}$, where $\kappa = c^{-2} > 0$, for c a constant of metric subregularity of $T_{DR} - \mathrm{Id}$ at $\overline{x}$ for the neighborhood $\mathcal{O}$.

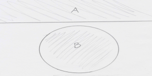


Proof logic.

convex, plc $\Longrightarrow$ single-valued polyhedral prox-operators $\Longrightarrow$ T_{DR} is polyhedral.
T_{DR} polyhedral + Fix $T_{DR} \cap W \cap \mathcal{O}$ an isolated point $\Longrightarrow$ metric subregularity [Dontchev& Rockafellar, Propositions 3I.1 and 3I.2]. $\square$

Outline

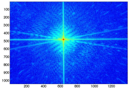
Application: phase retrieval



Phase retrieval

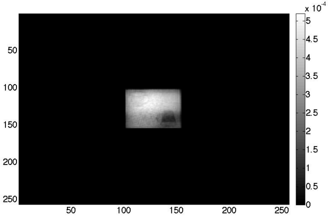
Given

- $|(Ax)_j|^2 = b_j$ for $b_j \in \mathbb{R}_+$ ($j = 1, 2, \ldots, m$) given by

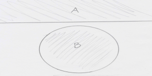


- some qualitative constraint ($x \in {\mathbb{R}^n}_+$ or supp $x \subset D$)

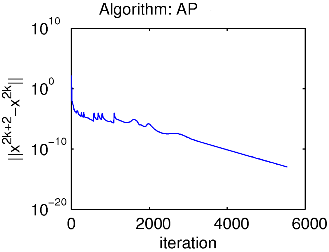
Find

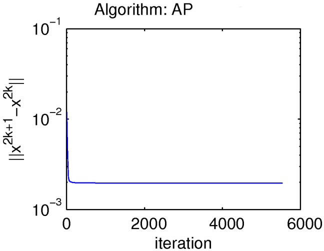


Application: phase retrieval

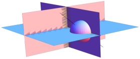


Phase retrieval





Application: sparse affine feasibility [Hesse-L.-Neumann, 2014]



$$\underset{x\in A_s}{\text{minimize}}\ \frac{1}{2}\operatorname{dist}(x,B)^2$$

where $A_s := \left\{x \in \mathbb{R}^n \middle|\ \|x\|_0 \le s\right\}$ and $B := \left\{x \in \mathbb{R}^n \middle|\ Mx = p\right\}$.

- Local convergence of alternating projections and Douglas-Rachford
- Global convergence of alternating projections under the assumption:

$$\begin{gathered} M \text{ is full rank and} \\ (1-\delta_{2s})\,\|x\|_2^2 \le \left\|M^\dagger Mx\right\|_2^2 \quad \forall\, x \in A_{2s} \end{gathered} \tag{13}$$

Denote $A_J := \operatorname{span}\{e_i |\ i \in J\}$ for $J \in \mathcal{J}_{2s} := \left\{J \in 2^{\{1,2,\ldots,n\}} \middle|\ J \text{ has } 2s \text{ elements}\right\}$. Then M satisfies (13) with $\delta_{2s} \in [0, \frac{\alpha-1}{\alpha})$ for some fixed $s > 0$ and $\alpha > 1 \implies$

$$(\forall J \in \mathcal{J}_{2s}) \qquad A_J \cap \ker(M) = \{0\}.$$

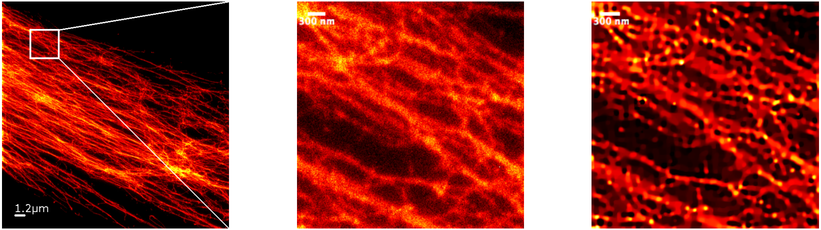
Image denoising/deconvolution

$$\underset{u\in\mathbb{R}^n}{\text{minimize}}\ J(u) + \rho \max\{F_\epsilon(Au)\}.$$

where J is convex, piecewise linear-quadratic, $A:\ \mathbb{R}^n \to \mathbb{R}^n$, and

$$F_\epsilon\ :\ \mathbb{R}^n \to 2^{\mathbb{R}^n}\ := v \mapsto (f_1(v) - \epsilon_1, f_2(v) - \epsilon_2, \ldots, f_M(v) - \epsilon_M)^T \qquad (14)$$

is convex quadratic.



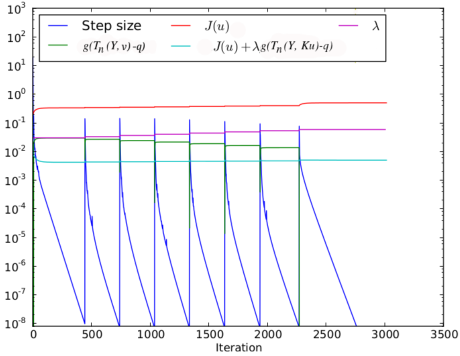
Solve with ADMM = Douglas-Rachford on the dual.

$$\underset{u\in\mathbb{R}^n}{\text{minimize}}\ J(u) + \rho \max\{F_\epsilon(Au)\}.$$

where J is convex, piecewise linear-quadratic, $A : \mathbb{R}^n \to \mathbb{R}^n$, and

$$F_\epsilon : \mathbb{R}^n \to 2^{\mathbb{R}^n} := v \mapsto (f_1(v) - \epsilon_1, f_2(v) - \epsilon_2, \ldots, f_M(v) - \epsilon_M)^T \quad (15)$$

is convex quadratic.



Outline

References I

T. Aspelmeier, Charitha, J. Fadili, D. R. Luke,
On Linear Convergence of the ADMM/Douglas Rachford Algorithms and Application to Statistical Imaging
Preprint (2015)

H.H. Bauschke, D.R. Luke, H.M. Phan, and X. Wang,
Restricted normal cones and the method of alternating projections: theory,
Set-Valued and Variational Analysis (2013)

H.H. Bauschke, D.R. Luke, H.M. Phan, and X. Wang,
Restricted normal cones and the method of alternating projections: applications,
Set-Valued and Variational Analysis (2013)

H.H. Bauschke, D.R. Luke, H.M. Phan, and X. Wang,
Restricted normal cones and sparsity optimization with affine constraints,
Foundations of Computational Mathematics (2013)

References II

J. M. Borwein and B. Sims.
The Douglas-Rachford algorithm in the absence of convexity.
Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Springer Optimization and its Applications, 49:93–109, 2011.

R. Hesse and D. R. Luke,
Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems,
SIAM J. Optimization, 23(4):2397–2419 (2013).

R. Hesse, D. R. Luke and P. Neumann
Alternating Projections and Douglas-Rachford for Sparse Affine Feasibility,
IEEE Transactions on Signal Processing, 62(18):4868–4881,2014.

References III

A.S. Lewis, D.R. Luke, and J. Malick,
Local linear convergence for alternating and averaged nonconvex projections,
Foundations of Computational Mathematics 9 (2009), 485–513.

A.S. Lewis and J. Malick,
Alternating projection on manifolds,
Mathematics of Operations Research 33 (2008), 216–234.

D. R. Luke,
Local Linear Convergence of Approximate Projections onto Regularized Sets,
Nonlinear Analysis, 75(2012):1531–1546