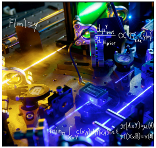
STRUCTURED (IN)FEASIBILITY: NONMONOTONE OPERATOR SPLITTING IN NONLINEAR SPACES

Russell Luke

Universität Göttingen

Optimization: Theory, Algorithms, Applications Lecture Series
Fields Institute and University of Waterloo



*SFB*1456

Deutsche Forschungsgemeinschaft

Outline I

Model Problems

Systems of Linear Inequalities

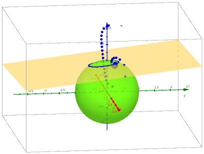
Given $A : \mathbb{R}^n \to \mathbb{R}^m$ and $b \in \mathbb{R}^m$ find $x \in \mathbb{R}^n$ so that

$$Ax \leq b$$

The Cone and Sphere Problem:

Find a point nearest (in some sense) to a cone and to spheres in the image of affine transformations.

Model Problems



Find a point $\overline{x}$ closest, in some sense, to sets C_j ($j = 1, 2, \ldots, m$).

Systems of Linear Inequalities

Given the vectors $a_j \in \mathbb{R}^n$ ($j = 1, 2, \ldots, m$) and the scalars $b_j \in \mathbb{R}$ define $C_j := \{x \mid a_j x \leq b_j\}$

Cone and Sphere

Given affine mappings $\mathcal{F}_j \; : \; \left(\mathbb{R}^d\right)^n \to \left(\mathbb{R}^d\right)^n$ ($j = 1, 2, \ldots, m$) with $\mathcal{F}_j(z) = \widehat{z} = \left(\widehat{z}_1, \widehat{z}_2, \ldots, \widehat{z}_n\right)$ for $\widehat{z}_i \in \mathbb{R}^d$. Seek vectors satisfying measurement/data given by

$$C_j := \left\{ z \in \left(\mathbb{R}^d\right)^n \mid \left\|\left(\mathcal{F}_j(z)\right)_i\right\| = b_{ij}, \quad \forall\, i = 1, 2, \ldots, n \right\}.$$

Qualitative constraints: support, nonnegativity, sparsity, symmetry modeled with the cone C_0

Model Categories

I Feasibility:

$$\text{Find } x^* \in \cap_{j=0}^{m} C_j$$

II Product Space Formulations

$$\text{Find } X^* \in D \cap (C_0 \times C_1 \times \cdots \times C_m)$$

where $D = \{X = (x_1, x_2, \ldots, x_m) \mid x_i = x_j \in \mathbb{X}\}$.

III Smooth Optimization

$$\underset{x \in \mathbb{X}}{\text{minimize}} \sum_{j=0}^{m} \text{dist}^2(x, C_j)$$

IV Constrained Optimization

$$\begin{array}{ll} \underset{x \in \mathbb{X}}{\text{minimize}} & f(x) \\ \text{subject to} & x \in \Omega \end{array}$$

Algorithms

Let

$$\iota_\Omega(x) := \begin{cases} 0 & \text{when } x \in \Omega \subset \mathbb{X} \\ +\infty & \text{else,} \end{cases}$$

$$\begin{array}{ll} \underset{x\in\mathbb{X}}{\text{minimize}} & f(x) \\ \text{subject to} & x \in \Omega \end{array} \iff \underset{x\in\mathbb{X}}{\text{minimize}}\ f(x) + \iota_\Omega(x)$$

$$\downarrow$$

$$\text{solve } 0 \in \partial\left(f(x) + \iota_\Omega(x)\right)$$

$$\downarrow$$

$$\text{Find } x \in T(x)$$

$$\downarrow$$

$$x^{k+1} \in T(x^k)$$

Main question: starting with the fixed point iteration, when can one follow the implications upward?

Answer #1: when f and Ω are convex and $\mathbb{X}$ is a linear space.

More interesting cases: when f and/or Ω are non-convex and/or $\mathbb{X}$ is a non-Euclidean space.

Splitting Algorithms

Let $(x^k)_{k\in\mathbb{N}}$ be generated by the fixed point iteration

$$x^{k+1} \in T(x^k)$$

or

$$x^{k+1} \in T_\alpha(x^k) := ((1-\alpha)\,\mathrm{Id} \oplus \alpha T)\,(x^k)$$

where $T : C \rightrightarrows C \subset \mathbb{X}$ is a composition/average of resolvents* (implicit steps) and descent mappings* (explicit steps) with Fix $T \neq \emptyset$: for example

$$T := (\alpha_1 r \oplus (1-\alpha_1)\,\mathrm{Id}) \circ (\alpha_2 g \oplus (1-\alpha_2)\,\mathrm{Id}).$$

* for instance, $g = -\nabla f$ (descent) and $r = (\mathrm{Id} + \nabla f)^{-1}$ (resolvent)

Outline I

Problem #1: Inconsistent convex feasibility

Linear Systems of Equations

Given $A \in \mathbb{R}^{m \times n}$ and $b_j \in \mathbb{R}^m$ solve

$$Ax = b$$

Problem #1: Inconsistent convex feasibility

Linear Systems of Equations

Given $A \in \mathbb{R}^{m\times n}$ and $b_j \in \mathbb{R}^m$ solve

$$Ax = b \quad \Longleftrightarrow \quad x \in \bigcap_{j=1,2,\ldots,m} L_j$$

$$\text{where} \quad L_j = \left\{ y \;\middle|\; \left\langle a_j,\ y \right\rangle = b_j \right\}$$

Problem #1: Inconsistent convex feasibility

Linear Systems of Equations

Given $A \in \mathbb{R}^{m \times n}$ and $b_j \in \mathbb{R}^m$ solve

$$Ax = b \quad \Longleftrightarrow \quad x \in \bigcap_{j=1,2,\ldots,m} L_j$$

$$\text{where} \quad L_j = \left\{ y \,\middle|\, \langle a_j,\ y \rangle = b_j \right\}$$

Cyclic Projections

$$x^{k+1} = P_{L_1} P_{L_2} \cdots P_{L_m} x^k$$

Problem #1

Linear Systems of Equations

Given $A \in \mathbb{R}^{m \times n}$ and $b_j \in \mathbb{R}^m$ find

$$x \in L := \bigcap_{j=1,2,\ldots,m} L_j \neq \emptyset \quad \Longrightarrow \quad \text{Cyclic Projections:}$$

$$\text{where } L_j = \{y \mid \langle a_j,\ y\rangle = b_j\} \qquad x^{k+1} = P_{L_1} P_{L_2} \cdots P_{L_m} x^k$$

Problem #1

Linear Systems of Equations

Given $A \in \mathbb{R}^{m \times n}$ and $b_j \in \mathbb{R}^m$ find

$$x \in L := \bigcap_{j=1,2,\ldots,m} L_j \neq \emptyset \quad \Longrightarrow \quad \text{Cyclic Projections:}$$

$$\text{where } L_j = \{y \mid \langle a_j,\ y\rangle = b_j\} \qquad x^{k+1} = P_{L_1} P_{L_2} \cdots P_{L_m} x^k$$

Convergence Theory

If A is full rank then cyclic projections converges either finitely or linearly to some $\overline{x} \in L$ from any starting point.

Problem #1

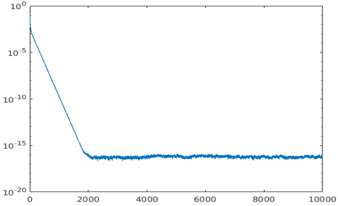
Linear Systems of Equations

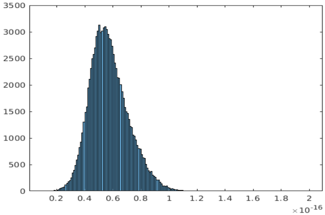
Given $A \in \mathbb{R}^{m\times n}$ and $b_j \in \mathbb{R}^m$ find

$$x \in L := \bigcap_{j=1,2,\dots,m} L_j \quad \Longrightarrow \quad \text{Cyclic Projections:}$$

$$\text{where } L_j = \{y \mid \langle a_j,\ y\rangle = b_j\} \qquad x^{k+1} = P_{L_1}P_{L_2}\cdots P_{L_m}x^k$$

$m = 50, n = 60$





Problem #1

Linear Systems of Equations

Given $A \in \mathbb{R}^{m \times n}$ and $b_j \in \mathbb{R}^m$ find

$$x \in L := \bigcap_{j=1,2,\ldots,m} L_j \quad \Longrightarrow \quad \text{Cyclic Projections with errors:}$$

where $L_j = \{ y \mid \langle a_j,\ y \rangle = b_j \}$ $\qquad x^{k+1} = P_{L_1} P_{L_2} \cdots P_{L_m} x^k + \epsilon_k$

Problem #1

Linear Systems of Equations

Given $A \in \mathbb{R}^{m\times n}$ and $b_j \in \mathbb{R}^m$ find

$$x \in L := \bigcap_{j=1,2,\dots,m} L_j \quad \Longrightarrow \quad \text{Cyclic Projections with errors:}$$

where $L_j = \{y \mid \langle a_j,\ y\rangle = b_j\}$ $\qquad x^{k+1} = P_{L_1}P_{L_2}\cdots P_{L_m}x^k + \epsilon_k$

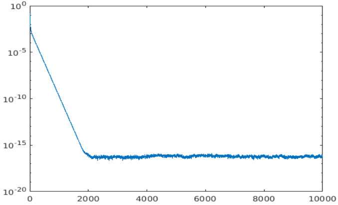
Convergence Theory

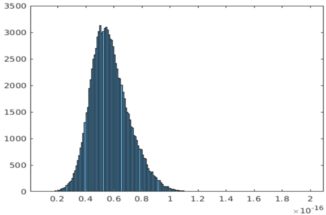
If A is full rank then cyclic projections with vanishing errors converges.

Problem #1

Does not explain:

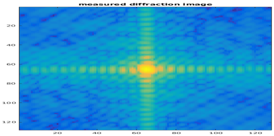
$m = 50, n = 60$

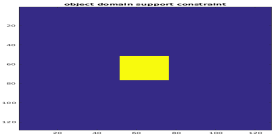




Problem #2: inconsistent, nonconvex feasibility

Coherent Diffraction Imaging





$$\mathcal{F}_1 \;:\; \left(\mathbb{R}^2\right)^{128} \to \left(\mathbb{R}^2\right)^{128} \qquad \text{(Fresnel transform)}$$

$$\mathcal{F}_1(z) = \widehat{z} \text{ for } \widehat{z} \in \mathbb{R}^{2\times 128}$$

define

$$C_1 \;:=\; \left\{ z \in \left(\mathbb{R}^2\right)^{128} \;\middle|\; \left\|(\mathcal{F}_1(z))_i\right\| = b_i, \quad \forall\; i = 1, 2, \dots, 128 \right\}$$

$$C_0 \;=\; (\mathbb{R}_+ \times \{0\})^{128}$$

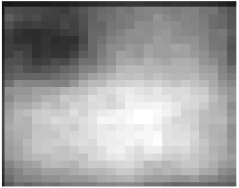
Most efficient numerical methods: $x^{k+1} \in T_* x^k$ where

$$T_{CP} \;:=\; P_{C_0} P_{C_1}$$

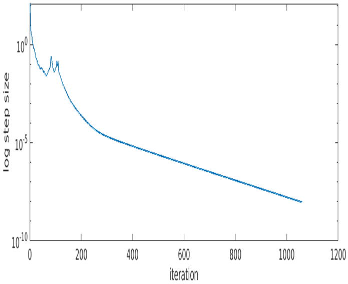
$$T_{DR} \;:=\; \tfrac{1}{2}\left((2P_{C_0} + \mathrm{Id}) \circ (2P_{C_1} + \mathrm{Id}) + \mathrm{Id}\right)$$

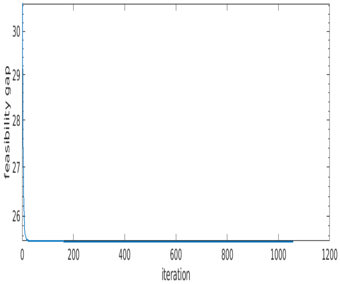
$$T_{DR\lambda} \;:=\; \tfrac{\lambda}{2}\left((2P_{C_0} + \mathrm{Id}) \circ (2P_{C_1} + \mathrm{Id}) + \mathrm{Id}\right) + (1-\lambda)P_{C_1}$$

Problem #2

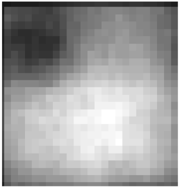


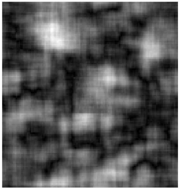
CP [L.Tam& Thao, 2018]



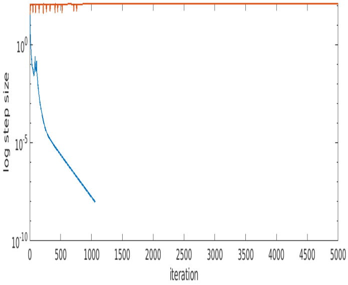


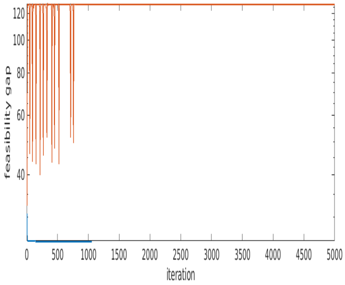
Problem #2





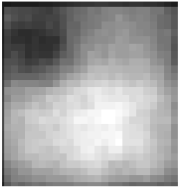
CP [L.Tam& Thao, 2018] DR



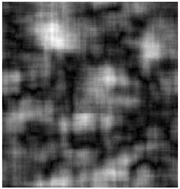


Problem #2

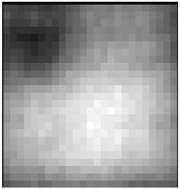
Relaxed Douglas-Rachford



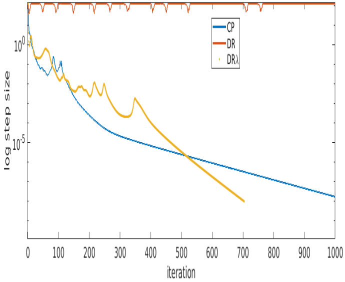
CP [L.Tam& Thao, 2018]

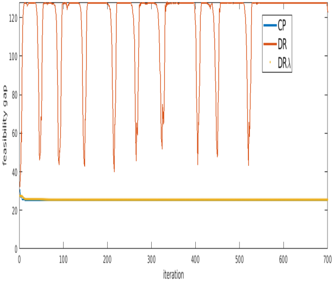


DR



DRλ[L.Martins, 2020]





Goals

- When does $\{y \mid 0 \in \partial(f(y) + \iota_\Omega(y))\} = \operatorname{argmin}_{x\in\Omega} f$?
  - How do you calculate $\partial(f(y) + \iota_\Omega(y))$?
  - How do you calculate $[\partial(f(\cdot) + \iota_\Omega(\cdot))]^{-1}(0)$?
- How do you construct T to find $\overline{x} \in \{y \mid 0 \in \partial(f(y) + \iota_\Omega(y))\}$?
- When does $\operatorname{Fix} T \sim \{\overline{x} \mid 0 \in \partial(f(\overline{x}) + \iota_\Omega(\overline{x}))\}$?
- When does the sequence defined by $x^{k+1} \in Tx^k$ converge to a point $\overline{x} \in \operatorname{Fix} T$?
- How fast does $x^k \to \overline{x} \in \operatorname{Fix} T$?
- At iteration K, what is $d(x^K, \operatorname{Fix} T)$?

Outline I

p-uniformly convex spaces

p-uniformly convex metric spaces

Let $p \in (1, \infty)$. A metric space (G, d) is p-uniformly convex with constant $c > 0$ whenever it is uniquely geodesic, and at any $x, y, z \in G$

$$(\forall t \in [0,1]) \quad d(z,(1-t)x \oplus ty)^p \leq (1-t)d(z,x)^p + td(z,y)^p - \frac{c}{2}t(1-t)d(x,y)^p.$$

Distance and Projectors

Distance and projector

Let Ω be a nonempty subset of $\mathbb{X}$. Then

$$\mathrm{dist}_\Omega\colon \mathbb{X} \to \mathbb{R}\colon x \mapsto \inf_{y\in\mathbb{X}} \mathrm{dist}(x,y) + \iota_\Omega(y)$$

is the distance function of the set Ω and

$$P_\Omega : \mathbb{X} \rightrightarrows \mathbb{X} : x \mapsto \{y \in \Omega \mid \mathrm{dist}_\Omega(x) = \mathrm{dist}(x,y)\}$$

is the corresponding (metric) projector. An element $y \in P_\Omega(x)$ is called a (metric) projection.
The inverse projector P_Ω^{-1} is defined by

$$P_\Omega^{-1}(y) := \{x \in \mathbb{X} \mid P_\Omega(x) \ni y\}\,.$$

Projectors

Nice properties

Let C be a nonempty closed subset of (G, d), a p-uniformly convex space. If C is convex, then

(i) P_C is pointwise α-firmly nonexpansive (defined later) at points in C and

(ii) $P_C(x)$ is a singleton, for every $x \in G$.

Projectors

Projector onto a sphere

Let $z \in \mathbb{E}$ and $\rho \in \mathbb{R}_{++}$. Set $\mathbb{S}_\rho(z) := \{x \mid \|x - z\| = \rho\}$. Then

$$(\forall x \in \mathbb{E}) \quad P_{\mathbb{S}_\rho(z)}(x) = \begin{cases} \rho \frac{x-z}{\|x-z\|} + z, & \text{if } x \neq z; \\ \mathbb{S}_\rho(z), & \text{otherwise.} \end{cases}$$

while

$$P^{-1}_{\mathbb{S}_\rho(z)}(y) = \{x \mid x = \rho\,(y - z) + z, \text{ for } \rho \geq 0\,\}$$

(i) $P_{\mathbb{S}_\rho(z)}$ is neither pointwise α-firmly nonexpansive

(ii) nor everywhere a singleton

Proximal Mappings

p-prox mappings

In a complete p-uniformly convex space the p-proximal mapping of a proper and lower semicontinuous function f is defined by

$$\operatorname{prox}^p_{f,\lambda}(x) := \operatorname{argmin}_{y \in G} f(y) + \frac{1}{p\lambda^{p-1}} d(y,x)^p \quad (\lambda > 0).$$

Example: projectors

For

$$\iota_\Omega(x) := \begin{cases} 0 & \text{when } x \in \Omega \subset G \\ +\infty & \text{else,} \end{cases}$$

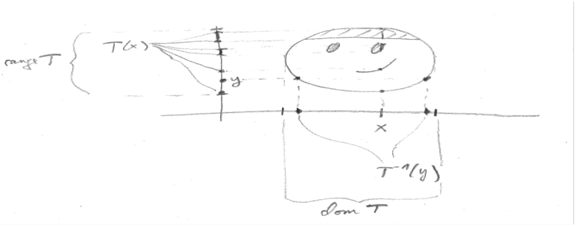
then

$$\operatorname{prox}^p_{\iota_\Omega} = P_\Omega(x).$$

Multi-mappings

Set-valued mappings

A set- or multi-valued mapping T from X to the set of subsets of Y, 2^Y, is denoted $T : X \rightrightarrows Y$. The inverse of a multi-valued mapping is defined by $T^{-1}(y) = \{x \mid T(x) \ni y\}$. The domain is the set $\mathrm{dom}\ T := \{x \mid T(x) \neq \emptyset\}$ and the range is $\mathrm{range}\, T := \{y \mid \exists\, x \in X \text{ with } y \in T(x)\}$. T is said to be closed-valued at x when $T(x)$ is closed.



Fixed points of set-valued mappings

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

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

Elements of fixed point theory

Pointwise almost nonexpansive mappings

Let (G, d) be a metric space and $D \subset G$. The mapping $T : D \rightrightarrows G$ is said to be pointwise almost nonexpansive at $x_0 \in D$ on D if there exists a constant $\epsilon \in [0, 1)$ such that

$$d(x^+, x_0^+) \leq \sqrt[p]{1+\epsilon}\, d(x, x_0)$$
$$\forall\, x^+ \in Tx \text{ and } \forall x_0^+ \in Tx_0 \text{ whenever } x \in D.$$

When the above inequality holds for all $x_0 \in D$ then F is said to be almost nonexpansive on D. If $\epsilon = 0$ in (1) the qualifier "almost" is dropped.

Single-valuedness of pointwise almost nonexpansive mappings

Any mapping $T : D \rightrightarrows G$ $(D \subset G)$ that is pointwise almost nonexpansive at $y \in \operatorname{Fix} T$ is single-valued there.

Elements of fixed point theory

Pointwise almost α-firmly nonexpansive mappings

Let (G, d) be a p-uniformly convex metric space with constant c ($p \in (1, \infty), c > 0$) and $D \subset G$. The mapping T is said to be pointwise almost α-firmly nonexpansive at $x_0 \in D$ on D whenever there exist constants $\alpha \in (0, 1)$ and $\epsilon \in [0, 1)$ such that

$$\begin{aligned} d^p(x^+, x_0^+) \quad &\leq \quad (1+\epsilon) d^p(x, x_0) - \frac{1-\alpha}{\alpha} \psi_T^{(p,c)}(x, x_0) \\ \text{where} \qquad & \psi_T^{(p,c)}(x, x_0) := \tfrac{c}{4} \left(d^p(x^+, x) + d^p(x_0^+, x_0) - d^p(x^+, x_0) \right. \\ & \qquad \left. + d^p(x, x_0) + d^p(x^+, x_0^+) - d^p(x, x_0^+) \right) \\ & \forall\, x^+ \in Tx \text{ and } \forall x_0^+ \in Tx_0 \quad \forall x \in D. \end{aligned} \tag{1}$$

When the above inequality holds for all $x_0 \in D$ then T is said to be almost α-firmly nonexpansive on D. If $\epsilon = 0$ in (1), the qualifier "almost" is dropped.

We call $\psi_T^{(p,c)}$ the transport discrepancy.

Elements of fixed point theory

On Euclidean spaces

- prox_f for f convex is α-firmly nonexpansive everywhere with $\alpha = 1/2$ - no violation, not pointwise
- Projectors onto convex sets: α-firmly nonexpansive everywhere with $\alpha = 1/2$ - no violation, not pointwise
- Projectors onto smooth manifolds: almost α-firmly nonexpansive on neighborhoods, with violation arbitrarily small for small enough neighborhoods [L., Tam, Thao, 2018]
- $T := \text{Id} - \lambda \nabla f$: steepest gradient descent with step length λ is almost α-firmly nonexpansive with violation arbitrarily small for small enough steps [L., Tam, Thao, 2018]

Elements of fixed point theory

On p-uniformly convex spaces

- prox_f for f convex is pointwise almost α-firmly nonexpansive at fixed points with constant and violation dependent on the curvature constant c of the space [Berdellima, Lauster,L., 2021]
- Projectors onto convex sets: pointwise α-firmly nonexpansive at fixed points with $\alpha = 1/2$ [Berdellima, Lauster,L., 2021]
- $T := \beta \,\text{Id} \oplus (1-\beta)\,\text{prox}^p_{f,\lambda}$: nonlinear gradient descent with step length λ is pointwise almost α-firmly nonexpansive with constant and violation dependent on β and the curvature constant c [Lauster, L. 2021]

Elements of fixed point theory

Calculus

When fixed points coincide, convex combinations and compositions of almost α-firmly nonexpansive mappings are almost α-firmly nonexpansive. In CAT(0) spaces the operators need not have common fixed points.

- This means that we can apply these ideas to all the leading algorithms for the problems of interest.
- In convex settings, pointwise α-firm nonexpansiveness at fixed points is all that is needed to prove convergence of fixed point iterations.

Inverse Regularity

Metric regularity on a set

Let (G_1, d_1) and (G_2, d_2) be metric spaces and let $\mathcal{T} : G_1 \rightrightarrows G_2$, $U_1 \subset G_1$, $U_2 \subset G_2$. The mapping $\mathcal{T}$ is called metrically regular on $U_1 \times U_2$ with gauge ρ whenever

$$d_1\left(x, \mathcal{T}^{-1}(y)\right) \leq \rho(d_2(y, \mathcal{T}(x)))$$

holds for all $x \in U_1$ and $y \in U_2$ with $0 < \rho(d_2(y, \mathcal{T}(x)))$ where $\mathcal{T}^{-1}(y) := \{z \mid \mathcal{T}(z) = y\}$. When the set U_2 consists of a single point, $U_2 = \{\bar{y}\}$, then $\mathcal{T}$ is said to be metrically subregular for $\bar{y}$ on U_1 with gauge ρ.

[Aze (06), Klatte&Kummer (09), Ioffe (11, 13), Ngai& Théra (04, 08)], Dontchev&Rockafellar (14), Kruger, L. &Thao (16)]

Generic Quantitative Convergence [L., Tam, & Thao 2018; Berdellima, Lauster& L. 2020/21]

From the transport discrepancy $\psi_T^{(p,c)}$ and a subset $S \subset G$ we construct the following surrogate mapping $\mathcal{T}_S : G \to \mathbb{R}_+ \cup \{+\infty\}$ by

$$\mathcal{T}_S(x) := \left(\inf_{y \in S} \psi_T^{(p,c)}(x, y) \right)^{1/p}.$$

Basic Result

Let (G, d) be a p-uniformly convex space, let $T : D \rightrightarrows D$ for $D \subseteq G$ and $T(D)$ boundedly compact, and let Fix T be nonempty. Assume

(i) T is pointwise almost α-firmly nonexpansive at all points $y \in S =$ Fix T with the same constant α on D and arbitrarily small violation on small enough neighborhoods of Fix T;

(ii) $\mathcal{T}_S$ is metrically subregular for 0 on D with gauge ρ, that is,

$$d(x, \mathcal{T}_S^{-1}(0)) \leqslant \rho(|\mathcal{T}_S(x)|), \quad \forall x \in D. \tag{2}$$

Then for any $x_0 \in D$ close enough to Fix T, the sequence $(x_k)_{k \in \mathbb{N}}$ defined by $x_{k+1} := Tx_k$ converges in the metric d to some $x^* \in$ Fix T with rate characterized by ρ and the violation ϵ.

Error bound

In particular, if (*ii*) holds with the constant κ satisfying $\sqrt{\frac{1-\alpha}{(1+\epsilon)\alpha}} \leq \kappa < \sqrt{\frac{1-\alpha}{\epsilon\alpha}}$, then for all $x^0 \in (\mathcal{O} \cap D) \setminus (\mathrm{Fix}\ T)$ the iteration $x^{k+1} \in T(x^k) \to \overline{x} \in \mathrm{Fix}\ T$ R-linearly with rate

$$c := \left(1 + \epsilon - \frac{1-\alpha}{\alpha\kappa^2}\right)^{1/2} < 1.$$

In other words, when we stop our algorithm at iteration k we can provide an upper bound on the distance of x^k to the model solution $x^* \in S$.

Proof sketch

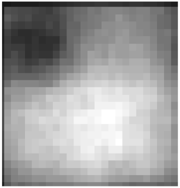
Start with the inequality defining metric subregularity, (2), and insert (1), the defining inequality from assumption (i), to get convergence to zero of the distance of the iterates to $\mathrm{Fix}\ T$. To get convergence of the iterates to a point $x^* \in \mathrm{Fix}\ T$, show that it is a Cauchy sequence. □

Verifying condition (i) is a masters thesis.

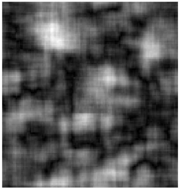
(ii) is a PhD dissertation, but necessary for convergence with rate characterized by ρ [L, Teboulle, Thao, 2019; Berdellima, Lauster, L, 2021]

Back to Example #2: inconsistent nonconvex feasibility

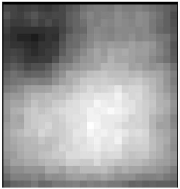
Explains inconsistent (realistic) phase retrieval



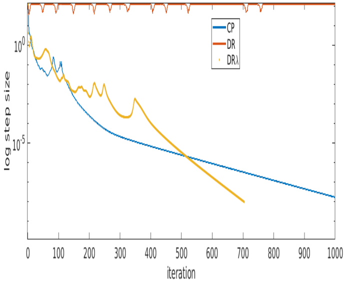
CP [L.Tam& Thao, 2018]

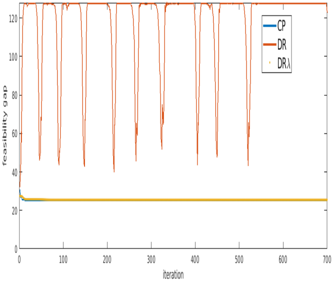


DR



DRλ[L.Martins, 2020]





Stochastic Setting:

The above results can be lifted to a stochastic setting:

- (G, d) is a Hilbert space
- $D \subset G$ is compact
- I is an index set (possibly uncountable)
- $T_i : D \to D$ is continuous for all $i \in I$
- $\mathcal{P}$ is the Markov operator with transition kernel p where

$$(x \in D)(A \in \mathcal{B}(D)) \qquad p(x, A) := \mathbb{P}(T_\xi x \in A)$$

- Define $\mathscr{T} : \mathscr{P}(D) \to \mathbb{R}_+ \cup \{+\infty\}$ by

$$\mathscr{T}(\mu) := \inf_{\pi \in \operatorname{inv} \mathcal{P}} \inf_{\gamma \in C_*(\mu, \pi)} \left(\int_{D \times D} \mathbb{E}_\xi \left[\psi_{T_\xi}(x, y) \right] \gamma(dx, dy) \right)^{1/2}.$$

Random Function Iterations:
Given $X^0 \sim \mu_0$

$$X^{k+1} = T_{\xi_k} X^k$$

Markov Chain:
Given μ_0

$$\mu_{k+1} = \mu_0 \mathcal{P}^k \overset{?}{\to} \operatorname{inv} \mathcal{P}$$

Stochastic Setting [Hermer, Sturm, L. arXiv 2020]

Linear/geometric convergence of Markov Chains

Suppose

(a) T_ξ is almost α-firmly nonexpansive in expectation with constant $\alpha \in (0,1)$ on D:

$$\mathbb{E}_\xi \left[d^2(T_\xi(x), T_\xi(y)) \right] \le (1+\epsilon) d^2(x,y) - \frac{1-\alpha}{\alpha} \mathbb{E}_\xi \left[\psi_{T_\xi}(x,y) \right], \quad \forall x, y \in D;$$

(b) $\mathscr{T}$ is metrically subregular with respect to W_2 for 0 on $\mathscr{P}_2(D)$ with constant κ:

$$d_{W_2}(\mu, \operatorname{inv} \mathcal{P}) \le \kappa |\mathscr{T}(\mu)| \quad \forall \mu \in \mathscr{P}_2(D).$$

where $\kappa \ge \sqrt{(1-\alpha)/(\alpha(1+\epsilon))}$

Then for any $\mu_0 \in \mathscr{P}_2(D)$ the sequence $(\mu_k)_{k \in \mathbb{N}}$ converges R-linearly to some $\pi^{\mu_0} \in \operatorname{inv} \mathcal{P}$:

$$d_{W_2}(\mu_{k+1}, \operatorname{inv} \mathcal{P}) \le c\, d_{W_2}(\mu_k, \operatorname{inv} \mathcal{P})$$

where $c := \sqrt{1 + \epsilon - \left(\frac{1-\alpha}{\kappa^2 \alpha}\right)} < 1$. If $\operatorname{inv} \mathcal{P}$ consists of a single point then convergence is Q-linear (geometric).

Back to Example #1

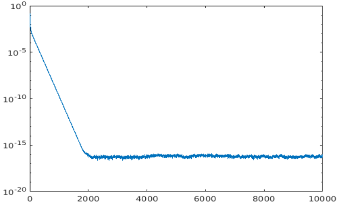
Explains iterative methods for linear systems of equations

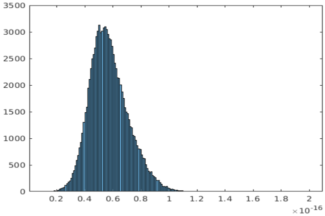
Given $A \in \mathbb{R}^{m\times n}$ and $b_j \in \mathbb{R}^m$ find

$$x \in L := \bigcap_{j=1,2,\ldots,m} L_j \quad \Longrightarrow \quad \text{Cyclic Projections:}$$

$$\text{where } L_j = \{ y \mid \langle a_j,\ y \rangle = b_j \} \qquad x^{k+1} = P_{L_1} P_{L_2} \cdots P_{L_m} x^k$$

$m = 50, n = 60$





TO DO

- In p-uniformly convex spaces: the prox mappings are assumed to have common fixed points for the calculus of α-firmly nonexpansive mappings to hold. Must determine the calculus for the inconsistent case
- In p-uniformly convex spaces: when does metric subregularity come "for free"?
- In p-uniformly convex spaces: algorithms on manifolds/nonlinear spaces (e.g. angle-resolved photon emission spectroscopy, femtosecond X-ray tomography...)
- In p-uniformly convex spaces and Wasserstein spaces: $T : D \rightrightarrows D$? (can only handle truly multi-valued mappings in Euclidean spaces)
- In p-uniformly convex spaces and Wasserstein spaces: when does Fix T correspond to critical points?
- In p-uniformly convex spaces and Wasserstein spaces: how do you characterize critical points?
- In Wasserstein spaces: new algorithms and methods for MCMC
- In Wasserstein spaces: randomized algorithms for large-scale, distributed optimization (e.g. femtosecond X-ray tomography, machine learning)

Outline I

References I

- A. Berdellima, F. Lauster and D. R. Luke,
  α-firmly nonexpansive operators on metric spaces
  J. Fixed Point Theory (in revision)
  https://arxiv.org/abs/2104.11302

- N. Hermer, D. R. Luke, and A. K. Sturm,
  Random Function Iterations for Stochastic Fixed Point Problems
  https://arxiv.org/abs/2007.06479
  (submitted July 2020)

- N. Hermer, D. R. Luke, and A. K. Sturm,
  Random Function Iterations for Consistent Stochastic Feasibility
  Numerical Functional Analysis and Optimization 40(4):386–420 (2019)
  DOI: 10.1080/01630563.2018.1535507.

References II

F. Lauster and D. R. Luke,
Convergence of proximal splitting algorithms in CAT(Îº) spaces and beyond
Fixed Point Theory Algorithms Sci Eng (2021) 2021:13
DOI 10.1186/s13663-021-00698-0

D. R. Luke.,
ProxToolbox,
https://num.math.uni-goettingen.de/proxtoolbox/

D. R. Luke, A.-L. Martins,
Convergence Analysis of the Relaxed Douglas Rachford Algorithm,
SIAM J. on Optimization, 30(1):542–584(2020).

D. R. Luke, Nguyen H. T., M. K. Tam
Quantitative convergence analysis of iterated expansive, set-valued mappings
Mathematics of Operations Research 43(4):1143–1176 (2018).
DOI https://doi.org/10.1287/moor.2017.0898.

References III

D. R. Luke, S. Sabach and M. Teboulle,
Optimization on Spheres: Models and Proximal Algorithms with Computational Performance Comparisons,
SIAM J. Mathematics of Data Science, 1(3), 408-445 (2019).

T. Salditt, A. Egner, and D.R. Luke (eds).
Nanoscale Photonic Imaging,
Springer Cham. (2020).

Outline I

ProxToolbox

https://num.math.uni-goettingen.de/proxtoolbox/

https://gitlab.gwdg.de/nam/ProxPython

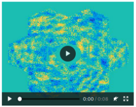
ProxToolbox

- ProxToolbox
  - Python version
    - Older versions
  - Matlab version
  - Additional binary data
  - Acknowledgements
- Getting started
  - Structure
    - I. proxtoolbox folder
    - II. InputData
    - III. Demos
      - Demos Accompanying Publications:

The ProxToolbox is a collection of modules for solving mathematical problems using fixed point iterations with proximal operators. It was used to generate many if not all of the numerical experiments conducted in the papers in the ProxMatlab Literature folder.

For a complete listing of papers with links, go to the [groups publications] (https://vaopt.math.uni-goettingen.de/en/publications.php).

This site is maintained by the Working Group in Variational Analysis of the Institute for Numerical and Applied Mathematics at the University of Göttingen.



Video: Representative iterate of a noisy JWST test wavefront recovered with the Douglas-Rachford algorithm

Python version

1.0.3:

- Sources
- Documentation

The documentation includes a tutorial.

Older versions

- **0.2.2**: Source, Documentation
- **0.1**: Source, Documentation

Matlab version

3.0

For help with the Matlab version, see the :code: `README` file in the ProxMatlab source.

Model Category I: feasibility

$$\text{Find } x^* \in \cap_{j=0}^m C_j$$

- cyclic projectors: $T := P_{C_0} P_{C_1} \cdots P_{C_m}$
- Douglas-Rachford (Lions 1979):
  $T_{DR} := \frac{1}{2}\left(R_{C_0} R_{C_1} + \mathrm{Id}\right)$ where $R_C := 2P_C - \mathrm{Id}$
- Relaxed Dougals-Rachford, DRλ (L. 2005, L. Martins 2020):
  $T_{DR\lambda} := \frac{\lambda}{2}\left(R_{C_0} R_{C_1} + \mathrm{Id}\right) + (1-\lambda)P_B$
- Douglas-Rachford-Alternating-Projectors (DRAP, Thao Nguyen 2019)

$$T_{DRAP} := P_{C_0}\left(\left(1+\lambda\right) P_{C_1} - \lambda\, \mathrm{Id}\right) - \lambda\left(P_{C_1}\, \mathrm{Id} - \mathrm{Id}\right)$$

* Algorithms in red only work for two sets.
* The sets C_j are nonconvex so the projectors and fixed point mappings T_* are set-valued.

Model Category I: feasibility

Multi-set extensions of Douglas-Rachford:

- Cyclic Douglas-Rachford (Tam& Borwein 2014):

$$T_{CDR} = T_{DR}^{01} \circ T_{DR}^{12} \circ \cdots \circ T_{DR}^{m0}$$

- Anchored Cyclic Douglas-Rachford (Bauschke,Noll& Phan 2015):

$$T_{CDR} = T_{DR}^{01} \circ T_{DR}^{02} \circ \cdots \circ T_{DR}^{0m}$$

- Cyclic DRλ (L., Martins, Tam 2018, L. Sabach, Teboulle 2019):

$$T = T_{DR\lambda}^{01} \cdot T_{DR\lambda}^{12} \cdot \cdots \cdot T_{DR\lambda}^{m0}.$$

Another motivation for T_{CDR} was to deal with inconsistency:

$$\cap_{j=0}^{m} C_j = \emptyset$$

Model Category I: feasibility

Before moving to the next model category, brief mention of alternating directions method of multipliers (ADMM_1, Glowinski& Marroco 1975) via nonsmooth constrained optimization and the the nonsmooth augmented Lagrangian:

$$\min_{x, z_j \in (\mathbb{R}^d)^n} \left\{ \iota_{C_0}(x) + \sum_{j=1}^{m} \iota_{C_j}(z_j) \mid z_j = x, \; j = 1, 2, \ldots, m \right\},$$

$$\widetilde{L}_\eta(x, z_j, v_j) := \iota_{C_0}(x) + \sum_{j=1}^{m} \left(\iota_{C_j}(z_j) + \langle v_j, \; x - z_j \rangle + \frac{\eta}{2} \|x - z_j\|^2 \right),$$

where

$$\iota_C(x) := \begin{cases} 0 & \text{wherever } x \in C \\ +\infty & \text{else} \end{cases}.$$

Model Category II: product space formulations

$$\text{Find } Z^* \in C \cap D,$$

where $Z^* = (z_0^*, z_1^*, \ldots, z_m^*)$, $C := C_0 \times C_1 \times \cdots \times C_m$ and D is the diagonal set of $\left(\mathbb{R}^d\right)^{n(m+1)}$ which is defined by $\left\{Z = (z, z, \ldots, z) : \ z \in \left(\mathbb{R}^d\right)^n\right\}$.

Two important features:

(i) the projector onto the set C is easily computed (in parallel):

$$P_C(Z) = \left(P_{C_0}(z_0), P_{C_1}(z_1), \ldots, P_{C_m}(z_m)\right),$$

(ii) D is a subspace which also has simple projection given by $P_D(Z) = \bar{Z}$ where

$$\bar{Z}_j = \frac{1}{m+1} \sum_{j=0}^{m} z_j.$$

Model Category II: product space

$$\text{Find } x^* \in \cap_{j=0}^m C_j$$

- alternating/averaged projectors: $T_{AvP} := P_D P_C = \frac{1}{m+1}\sum_{j=0}^m P_{C_j}$
- projected gradients: $T_{PG} := P_C\left(\mathrm{Id} - \lambda \nabla\, \mathrm{dist}_D^2\right) = P_C P_D \; (\lambda = 1/2)$
- Relaxed Dougals-Rachford, DRλ:
  $T_{DR\lambda} := \frac{\lambda}{2}\left(R_D R_C + \mathrm{Id}\right) + (1-\lambda)P_C$
- Douglas-Rachford-Alternating-Projectors (DRAP)

$$T_{DRAP} := P_D\left(\left(1+\lambda\right)P_C - \lambda\,\mathrm{Id}\right) - \lambda\left(P_C\,\mathrm{Id} - \mathrm{Id}\right)$$

We see from projected gradients that this strategy is a hidden way of smoothing the feasibility problem...

Model Category III: smooth optimization

$$\underset{z\in(\mathbb{R}^d)^n}{\text{minimize}}\ f(z) := \tfrac{1}{2(m+1)} \textstyle\sum_{j=0}^m \operatorname{dist}^2(z, C_j)$$

$$\Longleftrightarrow$$

$$\underset{z\in(\mathbb{R}^d)^n}{\text{minimize}}\ \tfrac{1}{2(m+1)} \textstyle\sum_{j=0}^m \left\|\mathcal{F}\mathcal{P}_j z - b_j\right\|^2$$

Note that

$$\nabla f(z) := \frac{1}{m+1} \sum_{j=0}^{m} \left(\operatorname{Id} - P_{C_j}\right)(z).$$

So that

$$T_{AvP} := \frac{1}{m+1} \sum_{j=0}^{m} P_{C_j} = (\operatorname{Id} - \nabla f),$$

which is just the method of steepest descent without stepsize optimization.

Model Category III: smooth optimization

Replacing the indicator functions in the nonsmooth ADMM_1 model with squared distance functions yields (L.Sabach, Teboulle, Zatlawey, 2017)

$$z^{k+1} = \frac{1}{m}\sum_{j=1}^{m}\left(u_j^k + \frac{1}{\rho_j}\left(z^k - z^{k-1}\right)\right).$$

$$u_j^{k+1} = P_{C_j}\left(u_j^k + \frac{1}{\rho_j}\left(2z^k - z^{k-1}\right)\right) \quad (j = 1, 2, \ldots, m)$$

which is recognizable as a two-step averaged projections recursion AvP^2.

Summary and theoretical status

		Global		Local	
	Literature	crit./fixed pts	rates	crit/fixed pts	rates
Model Category I					
CP	L. Tam, Thao 2018	+	0	+	+
CDR	Borwein& Tam, 2014, 2015	0	-	0	-
CDRλ	L., Martins& Tam 2018	0	-	0	-
$ADMM_1$	Bolte, Sabach, Teboulle 2018	+	0	+	0
Model Category II					
DR	Hesse&L. 2013; L., Tam& Thao 2018	0	-	+	+
DRλ	L.& Martins 2020,	0	-	+	+
DRAP	Thao 2018	0	-	+	+
Model Category III					
Wirtinger	Candes et al 2014	-	-	+	-
AP/AvP/PG	many	+	+	+	+
AvP^2	L., Sabach, Teboulle, Zatlawey 2017	+	-	+	-
DyRePr	Beck, Teboulle, Chikishev 2008, L. Burke& Lyons 2002	-	-	0	-
QNAvP	L. Burke& Lyons 2002	-	-	+	0

Table: State of the theory for the algorithms in this study. A $+$ indicates that the theory is well developed for settings that cover the cone and sphere problem; a 0 indicates that the theory is developed for certain settings - convex, for instance - but that the setting of the cone and sphere problem remains open. A $-$ indicates that the theory remains unexplored.

On Death, Taxes... and Metric Subregularity

A brief detour:

Linear monotonicity

$(x_k)_{k\in\mathbb{N}}$ is said to be linearly monotone with respect to Ω if there is a $c \in [0,1]$ such that

$$(\forall k \in \mathbb{N}) \qquad \operatorname{dist}(x_{k+1}, \Omega) \leq c \operatorname{dist}(x_k, \Omega).$$

Compare to Féjer monotonicity:

- $(x_k)_{k\in\mathbb{N}}$ is said to be Féjer monotone with respect to Ω if there is a $c \in [0,1]$ such that

$$(\forall k \in \mathbb{N})(\forall y \in \Omega) \qquad \operatorname{dist}(x_{k+1}, y) \leq c \operatorname{dist}(x_k, y).$$

On Death, Taxes... and Metric Subregularity

Linearly monotone fixed point sequences $\implies$ metric subregularity [L.-Teboulle-Thao, 2018]

Let $T : \Lambda \rightrightarrows \Lambda \subset \mathbb{E}$ with $\mathrm{Fix}\ T \neq \emptyset$. Fix $\delta \in (0, \infty]$. Generate the sequence $(x_k)_{k\in\mathbb{N}}$ by $x_{k+1} \in T(x_k)$.

- For all $x_0 \in ((\mathrm{Fix}\ T + \delta\mathbb{B}) \cap \Lambda) \setminus \mathrm{Fix}\ T$ the sequence $(x_k)_{k\in\mathbb{N}}$ is linearly monotone with respect to $\mathrm{Fix}\ T \cap \Lambda$ with constant $c \in (0, 1)$

  $\implies$

  $\Phi := T - \mathrm{Id}$ is metrically subregular on $(\mathrm{Fix}\ T + \delta\mathbb{B}) \setminus \mathrm{Fix}\ T$ for 0 relative to Λ with constant $\kappa \leq \frac{1}{1-c}$.
- If, in addition, T is pointwise almost averaged on $(\mathrm{Fix}\ T + \delta\mathbb{B}) \cap \Lambda$

  $\implies$

  $x_k \to \overline{x} \in \mathrm{Fix}\ T \cap \Lambda$ R-linearly with rate c.

History of $DR\lambda$

- [L. 2005] DRλ introduced (called RAAR) and global convergence for convex (inconsistent) feasibilty.
- [L. 2008] DRλ shown to be equivalent to Douglas-Rachford applied to the problem

$$\underset{x \in C_1}{\text{minimize}} \ \frac{\lambda}{2(1-\lambda)} \operatorname{dist}^2(x, C_0)$$

  Fixed points characterized[1]; local nonexpansivity and C_0 convex $\implies$ local convergence to Fix $T_{DR\lambda}$
- [Hesse-L. (2013), Phan (2016)] Showed local linear convergence of Douglas-Rachford for super-regular sets with transversal intersection
- [Li-Pong 2016] rediscovered DRλ and studied convergence when
  - C_0 is convex, and at least one of the sets is compact
  - C_0 convex, one of the sets is compact, and both semi-algebraic
  - C_0 is convex, and at least one of the sets is compact, and both semi-algebraic and $C_0 \cap C_1 \neq \emptyset$
- [Dao and Phan 2018] showed local R-linear convergence for superregular sets with linearly regular intersection.

[1] modulo a slight error

A Simple Counterexample

$$A := \left\{ x = (x_1, x_2) \in \mathbb{R}^2 \;\middle|\; x_1^2 + x_2^2 = 1 \right\}, \text{ and } B := \{(0,0)\}.$$

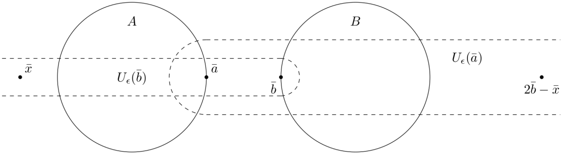
$$\text{Fix } T_{DR\lambda} = \emptyset \text{ for all } \lambda \in (0,1).$$

We assume throughout that Fix $T_{DR\lambda} \neq \emptyset$

$DR\lambda$ in [L-Martins. 2020]

- Sufficient regularity conditions (super-regularity at a distance) to guarantee that $T_{DR\lambda}$ is single-valued at $\overline{x} \in \text{Fix } T_{DR\lambda}$ and almost averaged for any given ϵ on a small enough neighborhood of $\overline{x}$
- Characterization of Fix $T_{DR\lambda}$ when C_0 and C_1 are closed and $T_{DR\lambda}$ is single-valued on Fix $T_{DR\lambda}$
- Sufficient regularity conditions (subtransversality of $\{C_0, C_1, C_0', C_1'\}$ and a technical condition) to guarantee that $T_{DR\lambda}$ is metrically subregular at $\overline{x} \in \text{Fix } T_{DR\lambda}$
- Sufficient conditions for local linear convergence of $T_{DR\lambda}$ for C_0 and C_1 nonconvex and nonintersecting.

$$\text{Fix } T_{DR\lambda} = \{\overline{x}\} = \left\{(2,0) - \frac{\lambda}{1-\lambda}(1,0)\right\}$$



Super-regularity at a distance

super-regularity [Lewis-L.-Malick,2009]

Let $\Omega \subseteq \mathbb{R}^n$ and $\overline{x} \in \Omega$. The set Ω is said to be super-regular at $\overline{x}$ if it is locally closed at $\overline{x}$ and, for every $\epsilon > 0$, there is a $\delta > 0$ such that for all $(x, 0) \in \operatorname{gph} N_\Omega \cap \{(\mathbb{B}_\delta(\overline{x}), 0)\}$

$$\langle y' - y, x - y \rangle \leq \varepsilon \, ||y' - y|| \|x - y\|, \quad (\forall y' \in \mathbb{B}_\delta(\overline{x}))\,(\forall y \in P_\Omega(y'))\,.$$

ϵ-subregularity [Daniilidis-L.Tam, 2018]

A set Ω is ϵ-subregular relative to Λ at $\overline{x}$ for $(x, v) \in \operatorname{gph} N_\Omega$ if it is locally closed at $\overline{x}$ and, there exists an $\epsilon > 0$ together with a neighborhood U_ϵ of $\overline{x}$, such that

$$\langle v - (y' - y),\ y - x \rangle \leq \epsilon \, \|v - (y' - y)\| \, \|y - x\| \quad (\forall y' \in \Lambda \cap U_\epsilon)(\forall y \in P_\Omega(y')$$

Ω is subregular relative to Λ at $\overline{x}$ for $(x, v) \in \operatorname{gph} N_\Omega$ if it is locally closed and for all $\epsilon > 0$ there exists U_ϵ such that the above holds.

Super-regularity at a distance

ϵ-subregularity [Daniilidis-L.Tam, 2018]

A set Ω is ϵ-subregular relative to Λ at $\overline{x}$ for $(x, v) \in \operatorname{gph} N_\Omega$ if it is locally closed at $\overline{x}$ and, there exists an $\epsilon > 0$ together with a neighborhood U_ϵ of $\overline{x}$, such that

$$\langle v - (y' - y),\ y - x\rangle \leq \epsilon \|v - (y' - y)\| \|y - x\| \quad (\forall y' \in \Lambda \cap U_\epsilon)(\forall y \in P_\Omega(y')$$

Ω is subregular relative to Λ at $\overline{x}$ for $(x, v) \in \operatorname{gph} N_\Omega$ if it is locally closed and for all $\epsilon > 0$ there exists U_ϵ such that teh above holds.

Super-regularity at a distance [L.-Martins, 2020]

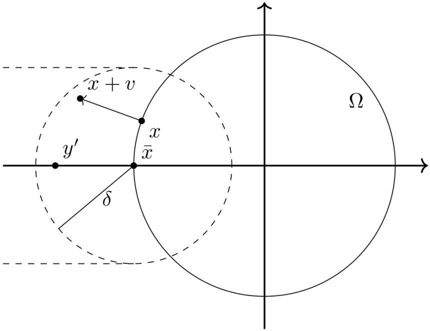
A set Ω is called ϵ-super-regular at a distance relative to Λ at $\overline{x}$ if it is ϵ-subregular relative to Λ at $\overline{x}$ for all $(x, v) \in V_\epsilon$ where

$$V_\epsilon := \left\{(x, v) \in \operatorname{gph} N_\Omega^{\text{prox}} \mid x + v \in U_\epsilon,\ x \in P_\Omega(x + v)\right\}. \qquad (3)$$

The set Ω is called super-regular at a distance relative to Λ at $\overline{x}$ if it is ϵ-super-regular relative to Λ at $\overline{x}$ for all $\epsilon > 0$.

Super-regularity at a distance: remarks

- $\bar{x}$ is not necessarily in Ω
- Super-regularity first instroduced in [Lewis-L.Malick, 2009] only refers to points in Ω
- Super-regularity at a distance $\implies$ super-regularity



Super-regularity at a distance $\implies$ almost averagedness of projectors

[L.-Martins 2020]
Let U be a neighborhood of $\overline{x} \in \Omega$. Let $\Lambda := P_\Omega^{-1}(\overline{x}) \cap U$. If Ω is ϵ-super-regular at a distance at $\overline{x}$ relative to Λ with constant ϵ on the neighborhood U, then

- If $\epsilon \in [0, 1)$, then P_Ω is pointwise almost nonexpansive at each $y' \in \Lambda$ with violation $\widetilde{\epsilon}$ on U for $\widetilde{\epsilon} := 4\epsilon/\left(1-\epsilon\right)^2$
- If $\epsilon \in [0, 1)$, then P_Ω is pointwise almost averaged with $\alpha = 1/2$ at each $y' \in \Lambda$ with violation $\widetilde{\epsilon}_2$ on U for $\widetilde{\epsilon}_2 := 4\epsilon(1+\epsilon)/\left(1-\epsilon\right)^2$
- R_Ω is pointwise almost nonexpansive at each $y' \in \Lambda$ with violation $\widetilde{\epsilon}_3 := 8\epsilon(1+\epsilon)/(1-\epsilon)^2$ on U

From these facts we can conclude almost averagedness of $T_{DR\lambda}$ with arbitrarily small violation ϵ.

Metric subregularity of $T_{DR\lambda}$

To show metric subregularity of

$$T_{DR\lambda} = \left(\frac{\lambda}{2} \left(R_{C_0} R_{C_1} + \mathrm{Id} \right) + (1 - \lambda) P_{C_1} \right)$$

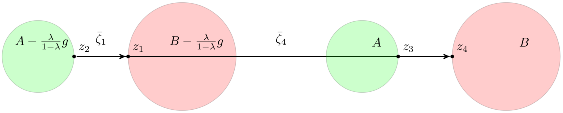
there are two routes:

(a) Compute the coderivative of $T_{DR\lambda}$ and verify that this is injective at $\bar{z} \in \operatorname{Fix} \, T_{DR\lambda}$, or

(b) Conclude metric subregularity of

$$T_{\bar{\zeta}} : \mathbb{E}^4 \rightrightarrows \mathbb{E}^4 \, : u \mapsto \left\{ \left(u_1^+, u_1^+ - \bar{\zeta}_1, u_1^+ - \bar{\zeta}_1 - \bar{\zeta}_2, u_1^+ + \bar{\zeta}_4 \right) \; \middle| \; u_1^+ \in T_{DR\lambda} u$$

from subtransversality of $\{B - \frac{\lambda}{1-\lambda} g, A - \frac{\lambda}{1-\lambda} g, A, B\}$ and another technical condition.



We chose path (b).

Geometry of Set Feasibility

Subtransversal collection of sets [Kruger-L.Thao, 18]

Let $\{\Omega_1, \Omega_2, \ldots, \Omega_m\}$ be a collection of nonempty closed subsets of $\mathbb{E}$ and define $\Psi : \mathbb{E}^m \rightrightarrows \mathbb{E}^m$ by $\Psi(x) := P_\Omega(\Pi x) - \Pi x$ where $\Omega := \Omega_1 \times \Omega_2 \times \cdots \times \Omega_m$, the projection P_Ω is with respect to the Euclidean norm on $\mathbb{E}^m$ and $\Pi : x = (x_1, x_2, \ldots, x_m) \mapsto (x_2, x_3, \ldots, x_m, x_1)$ is the permutation mapping on the product space $\mathbb{E}^m$ for $x_j \in \mathbb{E}$ $(j = 1, 2, \ldots, m)$. Let $\overline{x} = (\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_m) \in \mathbb{E}^m$ and $\overline{y} \in \Psi(\overline{x})$. The collection of sets is said to be subtransversal with constant κ relative to $\Lambda \subset \mathbb{E}^m$ at $\overline{x}$ for $\overline{y}$ if Ψ is metrically subregular at $\overline{x}$ for $\overline{y}$ on some neighborhood U of $\overline{x}$ with constant κ relative to Λ.

Metric subregularity from subtransversality

[L.Martins, 2020]

Let $\lambda \in (0,1)$, $\overline{x} \in \operatorname{Fix}\ T_{DR\lambda}$ with $T_{DR\lambda}$ being single-valued at $\overline{x}$ and set $g := P_B(\overline{x}) - P_A(P_B(\overline{x}))$. Furthermore, let $\overline{\zeta} \in \mathcal{Z}(\overline{x}, g)$ and $\overline{u} = (\overline{u}_1, \overline{u}_2, \overline{u}_3, \overline{u}_4) \in W_0(g)$ satisfy $\overline{\zeta} = \overline{u} - \Pi\overline{u}$ with $\overline{u}_1 = \overline{x}$. Let $T_{\overline{\zeta}}$ be defined as above and define $\Phi_{\overline{\zeta}} := T_{\overline{\zeta}} - \operatorname{Id}$. Suppose the following hold:

(i) the collection of sets $\left\{B - \frac{\lambda}{1-\lambda}g, A - \frac{\lambda}{1-\lambda}g, A, B\right\}$ is subtransversal at $\overline{u}$ for $\overline{\zeta}$ relative to $\Lambda \subseteq W\left(\overline{\zeta}\right)$ with constant κ and neighborhood U of $\overline{u}$;

(ii) there exists a positive constant σ such that

$$\operatorname{dist}\left(\overline{\zeta}, \Psi_g(u)\right) \leq \sigma \operatorname{dist}\left(0, \Phi_{\overline{\zeta}}(u)\right), \quad \forall u \in \Lambda \cap U \text{ with } u_1 \in B - \tfrac{\lambda}{1-\lambda}g.$$

Then the mapping $\Phi_{\overline{\zeta}} := T_{\overline{\zeta}} - \operatorname{Id}$ is metrically subregular for 0 on U relative to $\Lambda \cap \mathcal{N}$ with constant $\bar{\kappa} = \kappa\sigma$, where $\mathcal{N} := \left\{z \in \mathbb{E}^4 \;\middle|\; P_A(2z_4 + \frac{\lambda}{1-\lambda})g = z_3\right\}$.

Metric subregularity from subtransversality

Remark

The technical condition (ii) above only appears for inconsistent feasibility.

Harvest time

Local linear convergence of $T_{DR\lambda}$

Under the assumptions discussed above, there exists an $\epsilon' \leq \epsilon$ and a neighborhood $U' \subset U$ ($U' = U_1' \times U_2' \times U_3' \times U_4' \subset \mathbb{E}^4$) on which the sequence $\left(u^k\right)_{k\in\mathbb{N}}$ generated by $u^{k+1} \in T_{\overline{\zeta}} u^k$ seeded by a point $u^0 \in W\left(\overline{\zeta}\right) \cap U'$ with $u_1^0 \in U_1' \cap \left(B - \frac{\lambda}{1-\lambda} g\right)$ satisfies

$$\operatorname{dist}\left(u^{k+1}, \operatorname{Fix}\ T_{\overline{\zeta}} \cap S\right) \leq c \operatorname{dist}\left(u^k, S\right) \quad (\forall k \in \mathbb{N})$$

for

$$c := \sqrt{1 + \epsilon' - \frac{1}{2\overline{\kappa}^2}} < 1$$

where $\overline{\kappa} = \kappa\sigma$ with κ and σ. Consequently, $\operatorname{dist}\left(u^k, \tilde{u}\right) \to 0$ for some $\tilde{u} \in \operatorname{Fix}\ T_{\overline{\zeta}} \cap S$, and hence

$$\operatorname{dist}\left(u_1^k, \tilde{u}_1\right) \to 0$$

at least R-linearly with rate $c < 1$. If $\operatorname{Fix}\ T_{DR\lambda} \cap S_1$ is a singleton, then convergence is Q-linear.