Optimization and intrinsic geometry

Dmitriy Drusvyatskiy
School of ORIE, Cornell University

Joint work with
A. Daniilidis (Chile), A.D. Ioffe (Technion), A.S. Lewis (Cornell)

July 31, 2013

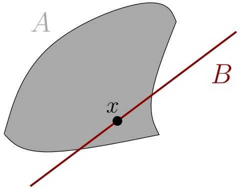
Outline

- **Method of alternating projections**

- **Active sets in optimization**

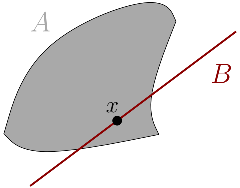
Method of alternating projections

Given sets $A, B \subset \mathbf{R}^n$, find some point $x \in A \cap B$.



Method of alternating projections

Given sets $A, B \subset \mathbf{R}^n$, find some point $x \in A \cap B$.

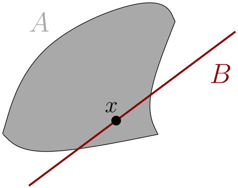


Distance and projection:

$$d_B(x) = \min_{y \in B} |x - y| \quad \text{and} \quad P_B(x) = \{\text{nearest points of } B \text{ to } x\}.$$

Method of alternating projections

Given sets $A, B \subset \mathbf{R}^n$, find some point $x \in A \cap B$.



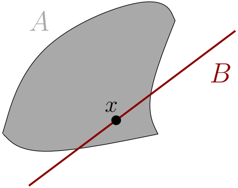
Distance and projection:

$$d_B(x) = \min_{y \in B} |x - y| \quad \text{and} \quad P_B(x) = \{\text{nearest points of } B \text{ to } x\}.$$

Finding points in P_A and P_B is often easy!

Method of alternating projections

Given sets $A, B \subset \mathbf{R}^n$, find some point $x \in A \cap B$.



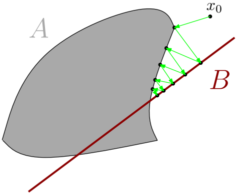
Distance and projection:

$$d_B(x) = \min_{y \in B} |x - y| \quad \text{and} \quad P_B(x) = \{\text{nearest points of } B \text{ to } x\}.$$

Finding points in P_A and P_B is often easy!

Method of alternating projections (von Neumann '33):

$$\begin{array}{lcl} x_{k+1} & \in & P_B(x_k) \\ x_{k+2} & \in & P_A(x_{k+1}) \end{array}$$



Method of alternating projections

Compressed sensing:

$$\{x : \operatorname{rank} x \leq r\} \bigcap \{x : Ax = b\}$$

Compressed sensing:

$$\{x : \operatorname{rank} x \le r\} \bigcap \{x : Ax = b\}$$

$P_{\{x:\,\operatorname{rank} x \le r\}}(y) \iff$ set $n - r$ smallest coordinates in $|\cdot|$ to zero.

Method of alternating projections

Compressed sensing:

$$\{x : \operatorname{rank} x \leq r\} \bigcap \{x : Ax = b\}$$

$P_{\{x:\, \operatorname{rank} x \leq r\}}(y) \iff$ set $n - r$ smallest coordinates in $|\cdot|$ to zero.

Low-order control:

$$\{X \succeq 0 : \operatorname{rank} X \leq r\} \bigcap \{X : \mathcal{A}(X) = b\}$$

Method of alternating projections

Compressed sensing:

$$\{x : \operatorname{rank} x \leq r\} \bigcap \{x : Ax = b\}$$

$P_{\{x:\, \operatorname{rank} x \leq r\}}(y) \Longleftrightarrow$ set $n - r$ smallest coordinates in $|\cdot|$ to zero.

Low-order control:

$$\{X \succeq 0 : \operatorname{rank} X \leq r\} \bigcap \{X : \mathcal{A}(X) = b\}$$

$P_{\{X \succeq 0:\, \operatorname{rank} X \leq r\}}(Y) \Longleftrightarrow$ diagonalize, set $n - r$ smallest eigenvalues to zero, set negative eigenvalues to zero.

Successes

Alternating projection heuristics remain popular!

Successes

Alternating projection heuristics remain popular!

Simple:

- require a few lines of code,
- no data structures, not memory or space intensive

Successes

Alternating projection heuristics remain popular!

Simple:

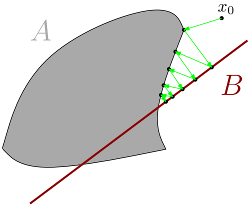
- require a few lines of code,
- no data structures, not memory or space intensive

Versatile:

- Inverse eigenvalue problems (Chen, Chu '96)
- Pole placement (Orsi, Yang '06)
- Information theory (Tropp, Dhillon, Heath, Strohmer '05)
- Low-order control design (Grigoriadis, Skelton '96)
- Image processing (Bauschke, Combettes, Luke '02)
- Hubble telescope (NASA '95)

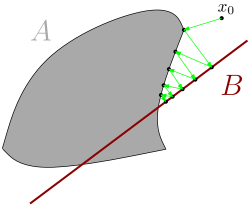
Convergence of alternating projections

$$\boxed{\begin{array}{lcl} x_{k+1} & \in & P_B(x_k) \\ x_{k+2} & \in & P_A(x_{k+1}) \end{array}}$$



Convergence of alternating projections

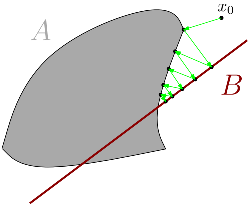
$$\begin{array}{lll} x_{k+1} & \in & P_B(x_k) \\ x_{k+2} & \in & P_A(x_{k+1}) \end{array}$$



The "angle" between A and B drives the convergence!

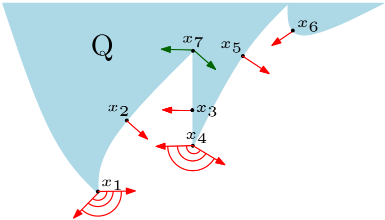
Convergence of alternating projections

$$\begin{array}{|lcl|}\hline x_{k+1} & \in & P_B(x_k) \\ x_{k+2} & \in & P_A(x_{k+1}) \\ \hline\end{array}$$



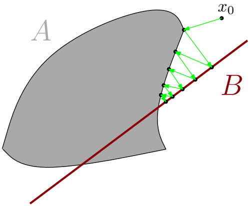
The "angle" between A and B drives the convergence!

Key tool: Normal cones $N_A(x)$ and $N_B(x)$



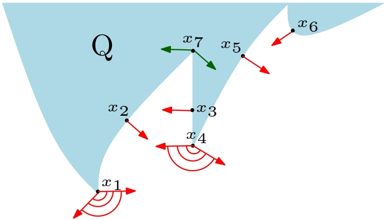
Convergence of alternating projections

$$\boxed{\begin{array}{lcl} x_{k+1} & \in & P_B(x_k) \\ x_{k+2} & \in & P_A(x_{k+1}) \end{array}}$$



The "angle" between A and B drives the convergence!

Key tool: Normal cones $N_A(x)$ and $N_B(x)$



Convergence theory: Bauschke, Borwein, Combettes, Deutsch, Lewis, Luke, Malick, Phan, Trussell, von Neumann, Wang etc. . .

Convergence of alternating projections

Transversality: $N_B(\bar{x}) \cap -N_A(\bar{x}) = \{0\}.$

Transversality: $N_B(\bar{x}) \cap -N_A(\bar{x}) = \{0\}.$

Eg:

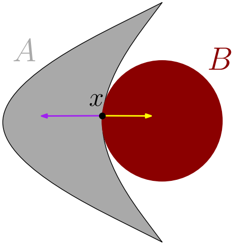


Figure: Not transverse

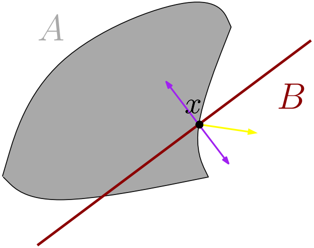


Figure: Transverse

Convergence of alternating projections

Transversality: $$N_B(\bar{x}) \cap -N_A(\bar{x}) = \{0\}.$$

Eg:

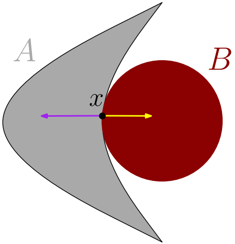


Figure: Not transverse

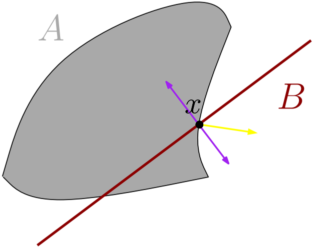


Figure: Transverse

Convergence of alternating projections (D-Ioffe-Lewis '13):

A and B transverse at $\bar{x}$ $\implies$ local **R**-linear convergence.

Slope and error bounds

Common problem: Estimate

$$\mathrm{dist}\,(x, [f \leq r]) \quad \text{(difficult)}$$

Slope and error bounds

Common problem: Estimate

$$\mathrm{dist}\,(x, [f \le r]) \quad \text{(difficult)}$$

"The residual":

$$f(x) - r \qquad \text{(easy)}$$

Slope and error bounds

Common problem: Estimate

$$\operatorname{dist}(x, [f \leq r]) \quad \text{(difficult)}$$

"The residual":

$$f(x) - r \qquad \text{(easy)}$$

Desirable quality: Exists κ with

$$\operatorname{dist}(x, [f \leq r]) \leq \kappa(f(x) - r)$$

Slope and error bounds

Common problem: Estimate

$$\operatorname{dist}(x, [f \le r]) \quad \text{(difficult)}$$

"The residual":

$$f(x) - r \qquad \text{(easy)}$$

Desirable quality: Exists κ with

$$\operatorname{dist}(x, [f \le r]) \le \kappa(f(x) - r)$$

Eg: Hoffman's error bound for linear programming.

Slope and error bounds

Common problem: Estimate

$$\operatorname{dist}(x, [f \le r]) \quad \text{(difficult)}$$

"The residual":

$$f(x) - r \qquad \text{(easy)}$$

Desirable quality: Exists κ with

$$\operatorname{dist}(x, [f \le r]) \le \kappa(f(x) - r)$$

Eg: Hoffman's error bound for linear programming.

Slope: "Fastest instantaneous rate of decrease"

$$\boxed{|\nabla f|(\bar{x}) := \limsup_{x \to \bar{x}} \frac{f(\bar{x}) - f(x)}{|\bar{x} - x|}}$$

Slope and error bounds

Common problem: Estimate

$$\mathrm{dist}\,(x, [f \le r]) \quad \text{(difficult)}$$

"The residual":

$$f(x) - r \qquad \text{(easy)}$$

Desirable quality: Exists κ with

$$\mathrm{dist}\,(x, [f \le r]) \le \kappa(f(x) - r)$$

Eg: Hoffman's error bound for linear programming.

Slope: "Fastest instantaneous rate of decrease"

$$\boxed{|\nabla f|(\bar{x}) := \limsup_{x \to \bar{x}} \frac{f(\bar{x}) - f(x)}{|\bar{x} - x|}}$$

Restrict $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ to a "slice" $f^{-1}(a, b)$.

Lemma (Error bound)

The following are equivalent.

Non-criticality:

$$|\nabla f| \ge \frac{1}{\kappa}$$

Error-bound:

$$\mathrm{dist}\,(x, [f \le r]) \le \kappa(f(x) - r), \quad \text{when } r \in (a, f(x))$$

Slope and error bounds

Common problem: Estimate

$$\operatorname{dist}(x, [f \le r]) \quad \text{(difficult)}$$

"The residual":

$$f(x) - r \qquad \text{(easy)}$$

Desirable quality: Exists κ with

$$\operatorname{dist}(x, [f \le r]) \le \kappa(f(x) - r)$$

Eg: Hoffman's error bound for linear programming.

Slope: "Fastest instantaneous rate of decrease"

$$\boxed{|\nabla f|(\bar{x}) := \limsup_{x \to \bar{x}} \frac{f(\bar{x}) - f(x)}{|\bar{x} - x|}}$$

Restrict $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$ to a "slice" $f^{-1}(a, b)$.

Lemma (Error bound)

The following are equivalent.

Non-criticality:

$$|\nabla f| \ge \frac{1}{\kappa}$$

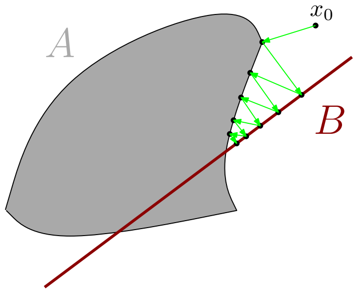
Error-bound:

$$\operatorname{dist}(x, [f \le r]) \le \kappa(f(x) - r), \quad \text{when } r \in (a, f(x))$$

- Observed by Azé-Corvellec '04, Ioffe '00.

Coupling function:

$$\psi(x, y) = \delta_A(x) + |x - y| + \delta_B(y).$$



Coupling function:

$$\boxed{\psi(x,y) = \delta_A(x) + |x-y| + \delta_B(y).}$$

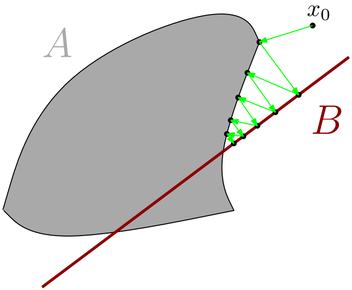
Error bound:

$$N_B(\bar{x}) \cap -N_A(\bar{x}) = \{0\}$$

$$\Downarrow$$

$$\boxed{\max\{|\nabla\psi_x|(y),\ |\nabla\psi_y|(x)\} \geq \kappa}$$

for $x \in A$ and $y \in B$, not in $A \cap B$.



Coupling function:

$$\psi(x, y) = \delta_A(x) + |x - y| + \delta_B(y).$$

Error bound:

$$N_B(\bar{x}) \cap -N_A(\bar{x}) = \{0\}$$

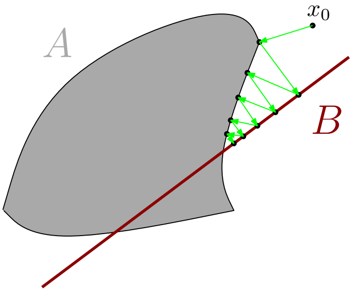
$$\Downarrow$$

$$\max\{|\nabla \psi_x|(y),\ |\nabla \psi_y|(x)\} \geq \kappa$$

for $x \in A$ and $y \in B$, not in $A \cap B$.

$$\Downarrow$$

Local linear convergence



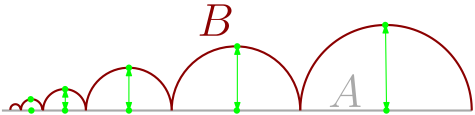
Convergence

- Transversality is necessary but not verifiable.

Convergence

- Transversality is necessary but not verifiable.

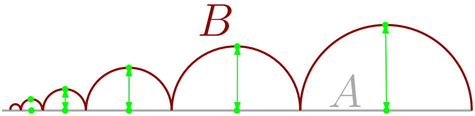
Eg:



Convergence

- Transversality is necessary but not verifiable.

Eg:

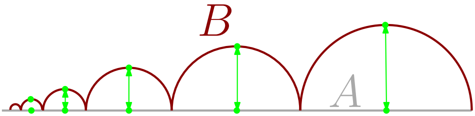


What about sublinear convergence?

Convergence

- Transversality is necessary but not verifiable.

Eg:



What about sublinear convergence?

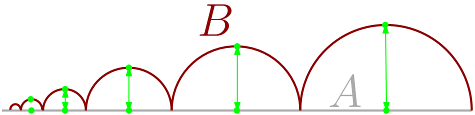
$Q \subset \mathbf{R}^n$ is semi-algebraic if is a finite union of

solution sets to finitely many polynomial inequalities.

Convergence

- Transversality is necessary but not verifiable.

Eg:



What about sublinear convergence?

$Q \subset \mathbf{R}^n$ is semi-algebraic if is a finite union of

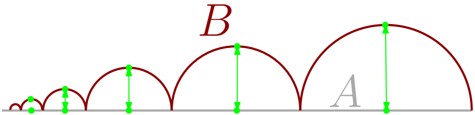
solution sets to finitely many polynomial inequalities.

Eg: semi-definite representable sets (Nesterov-Nemirovskii)

Convergence

- Transversality is necessary but not verifiable.

Eg:



What about sublinear convergence?

$Q \subset \mathbf{R}^n$ is semi-algebraic if is a finite union of

solution sets to finitely many polynomial inequalities.

Eg: semi-definite representable sets (Nesterov-Nemirovskii)

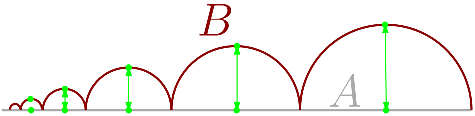
Theorem (D-Ioffe-Lewis)

A and B are semi-algebraic, $A \cap B$ is compact, x_0 near $A \cap B$ $\implies$ alternating projections converge.

Convergence

- Transversality is necessary but not verifiable.

Eg:



What about sublinear convergence?

$Q \subset \mathbf{R}^n$ is semi-algebraic if is a finite union of

solution sets to finitely many polynomial inequalities.

Eg: semi-definite representable sets (Nesterov-Nemirovskii)

Theorem (D-Ioffe-Lewis)

A and B are semi-algebraic, $A \cap B$ is compact, x_0 near $A \cap B$
$\Longrightarrow$ alternating projections converge.

Generic transversality (D-Ioffe-Lewis):

If A and B are semi-algebraic, then

$A + a$ and $B + b$ are transverse for a.e. (a, b)

Why semi-algebraicity?

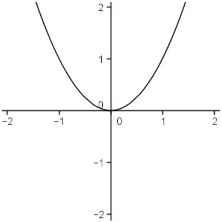


Figure: $f(x) = x^2$

Why semi-algebraicity?

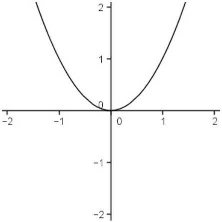


Figure: $f(x) = x^2$

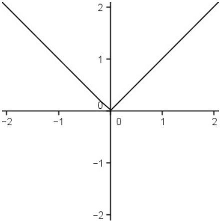


Figure: $\sqrt{f(x)} = |x|$

Why semi-algebraicity?

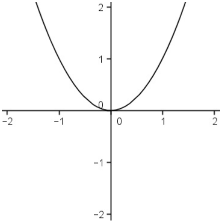


Figure: $f(x) = x^2$

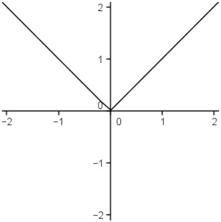


Figure: $\sqrt{f(x)} = |x|$

Desingularization: (Bolte-Daniilidis-Lewis '07)
For semi-algebraic f, there exists "nice" ϕ with

$$|\nabla(\phi \circ f)|(x) \geq 1 \quad \text{for } x \notin \operatorname{crit} f.$$

Why semi-algebraicity?

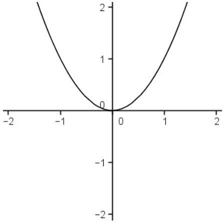


Figure: $f(x) = x^2$

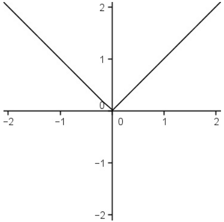


Figure: $\sqrt{f(x)} = |x|$

Desingularization: (Bolte-Daniilidis-Lewis '07)
For semi-algebraic f, there exists "nice" ϕ with

$$|\nabla(\phi \circ f)|(x) \geq 1 \quad \text{for } x \notin \operatorname{crit} f.$$

Error bounds always applicable for semi-algebraic functions!

Why semi-algebraicity?

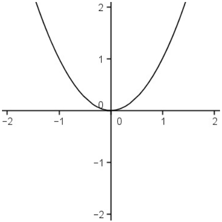


Figure: $f(x) = x^2$

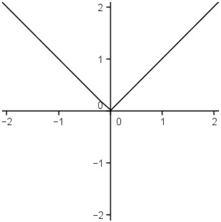


Figure: $\sqrt{f(x)} = |x|$

Desingularization: (Bolte-Daniilidis-Lewis '07)
For semi-algebraic f, there exists "nice" ϕ with

$$|\nabla(\phi \circ f)|(x) \geq 1 \quad \text{for } x \notin \operatorname{crit} f.$$

Error bounds always applicable for semi-algebraic functions!

- Apply to

$$\psi(x, y) = \delta_A(x) + |x - y| + \delta_B(y).$$

Open questions

General paradigm:

$$\text{no convexity} \quad \Longrightarrow \quad \text{no global convergence.}$$

Open questions

General paradigm:

$$\text{no convexity} \quad \Longrightarrow \quad \text{no global convergence.}$$

Variants of alternating projections work globally!

Open questions

General paradigm:

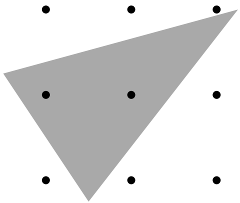
$$\text{no convexity} \quad \Longrightarrow \quad \text{no global convergence.}$$

Variants of alternating projections work globally!

- Integer programming:

$$\mathbb{Z}^2 \cap \{x : Ax \leq b\}$$

(eg: sudoku, 3-SAT, 4 queens problem, etc . . .)

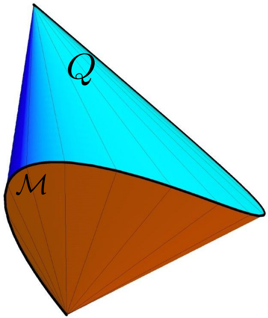


Ongoing work with Artacho, Borwein.

Active sets in optimization.

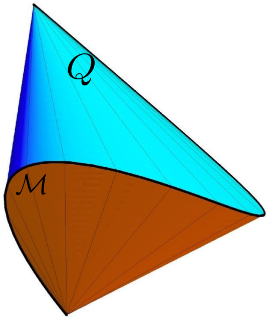
Active sets in optimization

Figure: Q is 4×4 Toeplitz spectrahedron



Active sets in optimization

Figure: Q is 4×4 Toeplitz spectrahedron



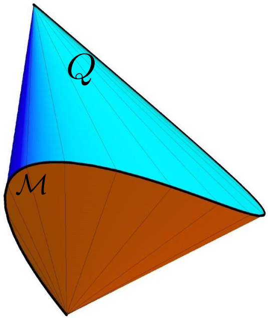
Definition (Partial Smoothness)

A set Q is partly smooth relative to $\mathcal{M} \subset Q$ if

1. **(Smoothness)** $\mathcal{M}$ is a smooth manifold,
2. **(Sharpness)** $N_{\mathcal{M}} = \operatorname{span} N_Q$ on $\mathcal{M}$,
3. **(Continuity)** N_Q varies continuously on $\mathcal{M}$.

Active sets in optimization

Figure: Q is 4×4 Toeplitz spectrahedron



Definition (Partial Smoothness)

A set Q is partly smooth relative to $\mathcal{M} \subset Q$ if

1. **(Smoothness)** $\mathcal{M}$ is a smooth manifold,
2. **(Sharpness)** $N_{\mathcal{M}} = \operatorname{span} N_Q$ on $\mathcal{M}$,
3. **(Continuity)** N_Q varies continuously on $\mathcal{M}$.

(Originates in Lewis '03)

Active sets in optimization

Partial smoothness has classical roots!

Eg: Smooth constraints

$$Q := \{x : g_i(x) \leq 0, \quad \text{for } i = 1, \ldots, m\}$$

where $g_1, \ldots, g_m$ are smooth.

Partial smoothness has classical roots!

Eg: Smooth constraints

$$Q := \{x : g_i(x) \le 0, \quad \text{for } i = 1, \ldots, m\}$$

where $g_1, \ldots, g_m$ are smooth.

Active indices:

$$I(x) = \{i : g_i(x) = 0\}$$

Active sets in optimization

Partial smoothness has classical roots!

Eg: Smooth constraints

$$Q := \{x : g_i(x) \le 0, \quad \text{for } i = 1, \ldots, m\}$$

where $g_1, \ldots, g_m$ are smooth.
Active indices:

$$I(x) = \{i : g_i(x) = 0\}$$

Reasonable conditions $\Longrightarrow$

$$\mathcal{M} := \{x : I(x) = I(\bar{x})\} \quad \text{is a partly smooth manifold near } \bar{x}$$

Active sets in optimization

Partial smoothness has classical roots!

Eg: Smooth constraints

$$Q := \{x : g_i(x) \le 0, \quad \text{for } i = 1, \ldots, m\}$$

where $g_1, \ldots, g_m$ are smooth.

Active indices:

$$I(x) = \{i : g_i(x) = 0\}$$

Reasonable conditions $\Longrightarrow$

$\mathcal{M} := \{x : I(x) = I(\bar{x})\}$ is a partly smooth manifold near $\bar{x}$

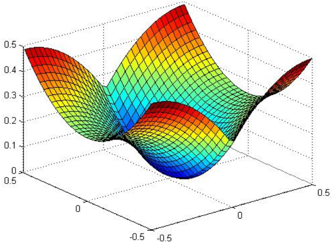


Figure: $Q = \{(x, y, z) : z \ge x(1-x) + y^2,\ z \ge -x(1+x) + y^2\}$

Active sets in optimization

Partial smoothness has classical roots!

Eg: **Sum of perturbed norms**

$$\min_x \; f(x) := \sum_{i=1}^{m} \|F_i(x)\|$$

where $F_1, \ldots, F_m$ are smooth.

Active sets in optimization

Partial smoothness has classical roots!

Eg: **Sum of perturbed norms**

$$\min_x \; f(x) := \sum_{i=1}^{m} \|F_i(x)\|$$

where $F_1, \ldots, F_m$ are smooth.

Active indices:

$$I(x) := \{i : F_i(x) = 0\}$$

Partial smoothness has classical roots!

Eg: Sum of perturbed norms

$$\min_x \; f(x) := \sum_{i=1}^{m} \|F_i(x)\|$$

where $F_1, \ldots, F_m$ are smooth.

Active indices:

$$I(x) := \{i : F_i(x) = 0\}$$

Reasonable conditions $\Longrightarrow$ $\operatorname{epi} f$ is partly smooth relative to

$$\mathcal{M} := \{(x, f(x)) : I(x) = I(\bar{x})\}$$

Partial smoothness has classical roots!

Eg: **Sum of perturbed norms**

$$\min_x \; f(x) := \sum_{i=1}^{m} \|F_i(x)\|$$

where $F_1, \ldots, F_m$ are smooth.

Active indices:

$$I(x) := \{i : F_i(x) = 0\}$$

Reasonable conditions $\Longrightarrow$ epi f is partly smooth relative to

$$\mathcal{M} := \{(x, f(x)) : I(x) = I(\bar{x})\}$$

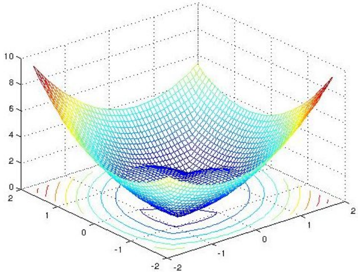


Figure: $f(x, y) := |x^2 + y^2 - 1| + |x - y|$

Why do optimizers care?

Active sets in optimization

Why do optimizers care?

- Many optimization algorithms identify $\mathcal{M}$ in finite time!

Eg: Gradient projection, Newton-like, proximal point.

Active sets in optimization

Why do optimizers care?

- Many optimization algorithms identify $\mathcal{M}$ in finite time!

Eg: Gradient projection, Newton-like, proximal point.

$\Longrightarrow$ Acceleration strategies!

Active sets in optimization

Why do optimizers care?

- Many optimization algorithms identify $\mathcal{M}$ in finite time!

Eg: Gradient projection, Newton-like, proximal point.

$\Longrightarrow$ Acceleration strategies!

Finite Identification: For $\bar{x} \in \mathcal{M}$ and $\bar{v} \in \operatorname{ri} N_Q(\bar{x})$, have

$$\left.\begin{array}{c} x_i \to \bar{x},\, v_i \to \bar{v} \\ v_i \in N_Q(x_i) \end{array}\right\} \Longrightarrow x_i \in \mathcal{M} \text{ for all large } i$$

Active sets in optimization

Why do optimizers care?

- Many optimization algorithms identify $\mathcal{M}$ in finite time!

Eg: Gradient projection, Newton-like, proximal point.

$\Longrightarrow$ Acceleration strategies!

Finite Identification: For $\bar{x} \in \mathcal{M}$ and $\bar{v} \in \operatorname{ri} N_Q(\bar{x})$, have

$$\left.\begin{array}{c} x_i \to \bar{x},\, v_i \to \bar{v} \\ v_i \in N_Q(x_i) \end{array}\right\} \Longrightarrow x_i \in \mathcal{M} \text{ for all large } i$$

finite identification $\iff$ partial smoothness (D-Lewis '13)

Active sets in optimization

Why do optimizers care?

- Many optimization algorithms identify $\mathcal{M}$ in finite time!

Eg: Gradient projection, Newton-like, proximal point.

$\Longrightarrow$ Acceleration strategies!

Finite Identification: For $\bar{x} \in \mathcal{M}$ and $\bar{v} \in \operatorname{ri} N_Q(\bar{x})$, have

$$\left.\begin{array}{c} x_i \to \bar{x},\, v_i \to \bar{v} \\ v_i \in N_Q(x_i) \end{array}\right\} \Longrightarrow x_i \in \mathcal{M} \text{ for all large } i$$

finite identification $\iff$ partial smoothness (D-Lewis '13)

How to see this structure in eigenvalue optimization?

Spectral sets

Consider $\mathbf{S}^n := \{n \times n \text{ symmetric matrices}\}$ and the eigenvalue map

$$X \mapsto (\lambda_1(X), \ldots, \lambda_n(X))$$

where

$$\lambda_1(X) \leq \ldots \leq \lambda_n(X).$$

Spectral sets

Consider $\mathbf{S}^n := \{n \times n \text{ symmetric matrices}\}$ and the eigenvalue map

$$X \mapsto (\lambda_1(X), \ldots, \lambda_n(X))$$

where

$$\lambda_1(X) \leq \ldots \leq \lambda_n(X).$$

Spectral sets:

$$\lambda^{-1}(Q) = \{X \in \mathbf{S}^n : \lambda(X) \in Q\}$$

Spectral sets

Consider $\mathbf{S}^n := \{n \times n \text{ symmetric matrices}\}$ and the eigenvalue map

$$X \mapsto (\lambda_1(X), \ldots, \lambda_n(X))$$

where

$$\lambda_1(X) \leq \ldots \leq \lambda_n(X).$$

Spectral sets:

$$\lambda^{-1}(Q) = \{X \in \mathbf{S}^n : \lambda(X) \in Q\}$$

where Q is permutation-invariant.

Spectral sets

Consider $\mathbf{S}^n := \{n \times n \text{ symmetric matrices}\}$ and the eigenvalue map

$$X \mapsto (\lambda_1(X), \ldots, \lambda_n(X))$$

where

$$\lambda_1(X) \leq \ldots \leq \lambda_n(X).$$

Spectral sets:

$$\lambda^{-1}(Q) = \{X \in \mathbf{S}^n : \lambda(X) \in Q\}$$

where Q is permutation-invariant.

(e.g. $\{X : \lambda_n(X) \leq 1\}$ or $\{X : \operatorname{rank} X \leq r\}$)

Spectral sets

Consider $\mathbf{S}^n := \{n \times n \text{ symmetric matrices}\}$ and the eigenvalue map

$$X \mapsto (\lambda_1(X), \ldots, \lambda_n(X))$$

where

$$\lambda_1(X) \leq \ldots \leq \lambda_n(X).$$

Spectral sets:

$$\lambda^{-1}(Q) = \{X \in \mathbf{S}^n : \lambda(X) \in Q\}$$

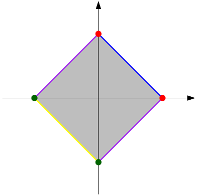
where Q is permutation-invariant.

(e.g. $\{X : \lambda_n(X) \leq 1\}$ or $\{X : \operatorname{rank} X \leq r\}$)

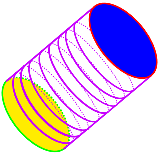
How to describe partly smooth structure of $\lambda^{-1}(Q)$?

Spectral sets

Eg:



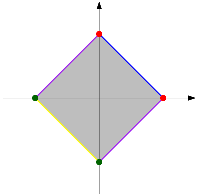
$|x| + |y| \leq 1$



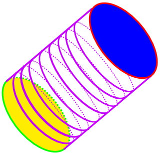
$|\lambda_1(X)| + |\lambda_2(X)| \leq 1$

Spectral sets

Eg:



$|x| + |y| \le 1$



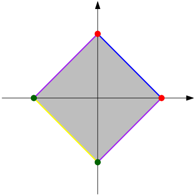
$|\lambda_1(X)| + |\lambda_2(X)| \le 1$

Recognizing partial smoothness (Daniilidis-D-Lewis):

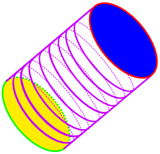
Q partly smooth at $\lambda(\bar{X})$ relative to $\mathcal{M}$

$\Longrightarrow \lambda^{-1}(Q)$ partly smooth at $\bar{X}$ relative to $\lambda^{-1}(\mathcal{M})$.

Eg:



$|x| + |y| \leq 1$



$|\lambda_1(X)| + |\lambda_2(X)| \leq 1$

Recognizing partial smoothness (Daniilidis-D-Lewis):

Q partly smooth at $\lambda(\bar{X})$ relative to $\mathcal{M}$

$\Longrightarrow \lambda^{-1}(Q)$ partly smooth at $\bar{X}$ relative to $\lambda^{-1}(\mathcal{M})$.

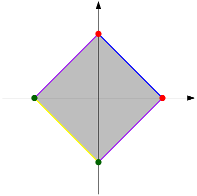
For Q convex polyhedral cone,

$$\lambda^{-1}(Q) \quad \overset{*}{\longleftrightarrow} \quad [\lambda^{-1}(Q)]^*$$

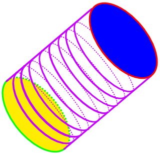
$$\lambda^{-1}(\mathcal{M}) \quad \overset{*}{\longleftrightarrow} \quad \bigcup_{X \in \lambda^{-1}(\mathcal{M})} N_{\lambda^{-1}(Q)}(X)$$

Spectral sets

Eg:



$|x| + |y| \le 1$



$|\lambda_1(X)| + |\lambda_2(X)| \le 1$

Recognizing partial smoothness (Daniilidis-D-Lewis):

Q partly smooth at $\lambda(\bar{X})$ relative to $\mathcal{M}$

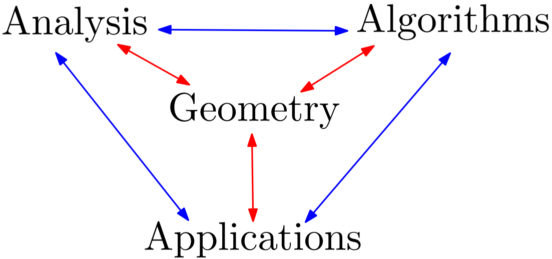
$\Longrightarrow \lambda^{-1}(Q)$ partly smooth at $\bar{X}$ relative to $\lambda^{-1}(\mathcal{M})$.

For Q convex polyhedral cone,

$$\lambda^{-1}(Q) \quad \overset{*}{\longleftrightarrow} \quad [\lambda^{-1}(Q)]^*$$

$$\lambda^{-1}(\mathcal{M}) \quad \overset{*}{\longleftrightarrow} \quad \bigcup_{X \in \lambda^{-1}(\mathcal{M})} N_{\lambda^{-1}(Q)}(X)$$

(Matrix-valued Bessel processes (D-Larsson '13))


Analysis
Algorithms
Geometry
Applications

Thank you.