A robust method for complementarity problems using a geometric viewpoint

Baptiste Plaquevent-Jourdain, with
Jean-Pierre Dussault, Université de Sherbrooke
Jean Charles Gilbert, INRIA Paris

October, 7th 2025

Outline

Complementarity Problems (CPs)

General form: (Facchinei Pang [FP03], Acary Brogliato [AB08]. . .)

Let $F, G : \mathbb{R}^n \mapsto \mathbb{R}^n$ be smooth,

$$\begin{aligned} \text{find } x \in \mathbb{R}^n \text{ s.t.} \quad & F(x) \geqslant 0,\ G(x) \geqslant 0,\ F(x)^{\mathsf{T}} G(x) = 0 \\ \iff \quad & 0 \leqslant F(x) \perp G(x) \geqslant 0. \end{aligned} \tag{1}$$

- NCP(F): $0 \leqslant x \perp F(x) \geqslant 0$;
- LCP(M, q): (Cottle Pang Stone [CPS09])

$$0 \leqslant x \perp (Mx + q) \geqslant 0. \tag{2}$$

Complementarity Problems (CPs)

General form: (Facchinei Pang [FP03], Acary Brogliato [AB08]...)

Let $F, G : \mathbb{R}^n \mapsto \mathbb{R}^n$ be smooth,

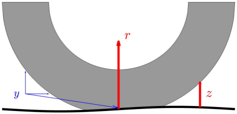
$$\begin{aligned} \text{find } x \in \mathbb{R}^n \text{ s.t.} \quad & F(x) \geqslant 0, \; G(x) \geqslant 0, \; F(x)^\mathsf{T} G(x) = 0 \\ \iff \quad & 0 \leqslant F(x) \perp G(x) \geqslant 0. \end{aligned} \tag{1}$$

- NCP(F): $0 \leqslant x \perp F(x) \geqslant 0$;
- LCP(M, q): (Cottle Pang Stone [CPS09])

$$0 \leqslant x \perp (Mx + q) \geqslant 0. \tag{2}$$

Relevance of CPs

Phenomena in competition, threshold effects [HP90; FP97]

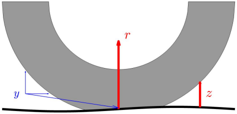


r: reaction, z: height

$$\forall \text{ point } y, \begin{cases} r(y) \geqslant 0, \\ z(y) \geqslant 0, \\ r(y)z(y) = 0. \end{cases}$$

Relevance of CPs

Phenomena in competition, threshold effects [HP90; FP97]



r: reaction, z: height

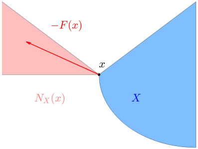
$$\forall \text{ point } y, \begin{cases} r(y) \geqslant 0, \\ z(y) \geqslant 0, \\ r(y)z(y) = 0. \end{cases}$$

Constrained optimization:

$$\min \; f(x) \quad \text{s.t.} \quad g(x) \leqslant 0,$$

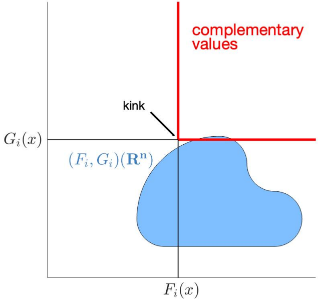
$$\text{KKT} \begin{cases} \nabla f(x) + \nabla g(x)\lambda = 0, \\ 0 \leqslant \lambda \perp (-g(x)) \geqslant 0. \end{cases}$$

Various other problems (Robinson [Rob80], [FP03]...).



An essential difficulty

At a solution x, $F_i(x) = 0$ or $G_i(x) = 0$ for all $i \in [1 : n]$:
2^n possibilities, **combinatorial aspect**.



NP-hard even for LCPs [Chu89], [Koj+91].

Some techniques

Pivoting [Lem65]: for LCPs, pivots between $x_i = 0$ or $(Mx + q)_i = 0$. But can visit all 2^n combinations [Mur78; Fat79].

Linearization [Jos79]: NCP $\rightarrow$ sequence of LCPs.

Interior points, smoothing [Koj+91; Had09; CNQ00; Vu+21]: $F_i(x) \geqslant 0$, $G_i(x) \geqslant 0$, $F_i(x)G_i(x) = 0 \rightarrow F_i(x)G_i(x) = \mu > 0$. Solution curve becomes smooth; different forms/approaches exist.

Some techniques

Pivoting [Lem65]: for LCPs, pivots between $x_i = 0$ or $(Mx+q)_i = 0$. But can visit all 2^n combinations [Mur78; Fat79].

Linearization [Jos79]: NCP $\to$ sequence of LCPs.

Interior points, smoothing [Koj+91; Had09; CNQ00; Vu+21]: $F_i(x) \geqslant 0$, $G_i(x) \geqslant 0$, $F_i(x)G_i(x) = 0 \to F_i(x)G_i(x) = \mu > 0$. Solution curve becomes smooth; different forms/approaches exist.

Some techniques

Pivoting [Lem65]: for LCPs, pivots between $x_i = 0$ or $(Mx + q)_i = 0$. But can visit all 2^n combinations [Mur78; Fat79].

Linearization [Jos79]: NCP $\rightarrow$ sequence of LCPs.

Interior points, smoothing [Koj+91; Had09; CNQ00; Vu+21]: $F_i(x) \geqslant 0$, $G_i(x) \geqslant 0$, $F_i(x)G_i(x) = 0 \rightarrow F_i(x)G_i(x) = \mu > 0$. Solution curve becomes smooth; different forms/approaches exist.

Outline

Possible reformulation techniques

C-function:

$$\varphi : \mathbb{R}^2 \to \mathbb{R}, \text{ s.t. } \quad \varphi(a,b) = 0 \quad \begin{aligned} &\iff \quad a \geqslant 0, b \geqslant 0, ab = 0, \\ &\iff \quad 0 \leqslant a \perp b \geqslant 0. \end{aligned}$$

$$0 \leqslant F(x) \perp G(x) \geqslant 0 \;\Leftrightarrow\; H_\varphi(x) := (\varphi(F_i(x), G_i(x)))_{i\in[1:n]} = 0$$

- $\varphi_{\mathrm{FB}}(a,b) := \sqrt{a^2+b^2} - a - b$ ([Fis92])
- $\varphi_{\min}(a,b) := \min(a,b)$, $H := H_{\min}$ ([Pan90; Pan91; Qi93], see also [Kos76; Wie82])
- Framework of [LT97; KYF97], surveys [FJ00; Alc+20; Gal12].
- See also [Fuk92; MS93]

Possible reformulation techniques

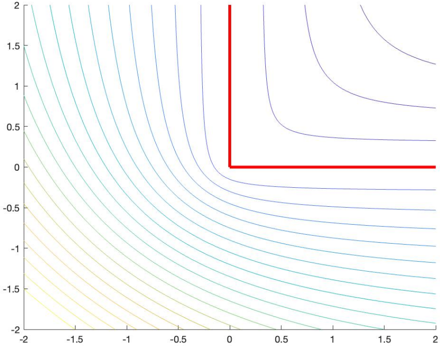
C-function:

$$\varphi : \mathbb{R}^2 \to \mathbb{R}, \text{ s.t. } \quad \varphi(a,b) = 0 \quad \begin{array}{cl} \iff & a \geqslant 0, b \geqslant 0, ab = 0, \\ \iff & 0 \leqslant a \perp b \geqslant 0. \end{array}$$

$$0 \leqslant F(x) \perp G(x) \geqslant 0 \;\Leftrightarrow\; H_\varphi(x) := (\varphi(F_i(x), G_i(x)))_{i \in [1:n]} = 0$$

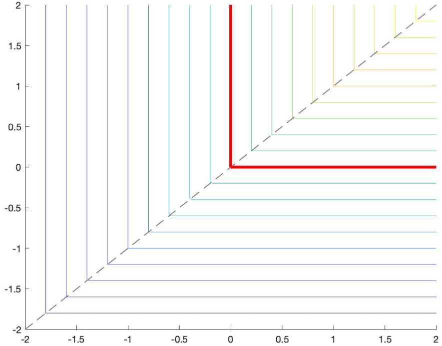
- $\varphi_{\mathrm{FB}}(a,b) := \sqrt{a^2 + b^2} - a - b$ ([Fis92])
- $\varphi_{\min}(a,b) := \min(a,b)$, $H := H_{\min}$ ([Pan90; Pan91; Qi93], see also [Kos76; Wie82])
- Framework of [LT97; KYF97], surveys [FJ00; Alc+20; Gal12].
- See also [Fuk92; MS93]

Illustration



Level sets of the C-function φ_{FB}, nondifferentiable only at the origin.
Level sets of the C-function $\varphi_{\min}$, nondifferentiable at the dashed line.

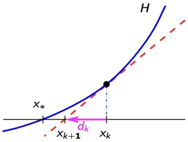
Illustration



Level sets of the C-function φ_{FB}, nondifferentiable only at the origin.

Level sets of the C-function $\varphi_{\min}$, nondifferentiable at the dashed line.

Systems of (nonsmooth) equations



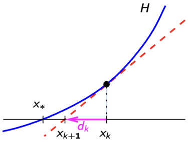
Smooth case

Iteration

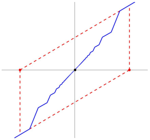
$$x_{k+1} = x_k - H'(x_k)^{-1} H(x_k)$$
$$(= x_k + d_k)$$

x_0 near x^*, $H \in \mathcal{C}^{1,1}$,
$H'(x^*)$ nonsingular
$\Rightarrow$ convergence

Systems of (nonsmooth) equations



Smooth case



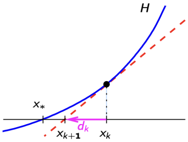
Can cycle if H nonsmooth [Kum88]

Iteration

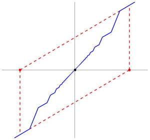
$$x_{k+1} = x_k - H'(x_k)^{-1} H(x_k)$$
$$\left(= x_k + d_k \right)$$

x_0 near x^*, $H \in \mathcal{C}^{1,1}$,
$H'(x^*)$ nonsingular
$\Rightarrow$ convergence

Systems of (nonsmooth) equations



Smooth case



Can cycle if H nonsmooth [Kum88]

Iteration

$$x_{k+1} = x_k - H'(x_k)^{-1} H(x_k)$$
$$(= x_k + d_k)$$

x_0 near x^*, $H \in \mathcal{C}^{1,1}$,
$H'(x^*)$ nonsingular
$\Rightarrow$ convergence

Semismooth Newton iteration

$$x_{k+1} = x_k - J_k^{-1} H(x_k)$$
any $J_k \in \partial_{B|C} H(x_k)$

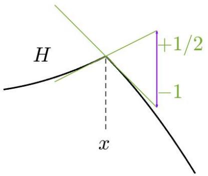
x_0 near x^*, H semismooth,
all $J \in \partial_{B|C} H(x^*)$ nonsingular
$\Rightarrow$ convergence

Generalizing derivatives

Two main differentials: [Cla90] (Clarke)

$$\partial_B H(x) := \{J \in \mathbb{R}^{n\times n} : \exists\ (x_k)_k \to x, x_k \in \overset{\text{domain}}{\mathcal{D}_H}, H'(x_k) \to J\},$$
$$\partial_C H(x) := \mathrm{conv}(\partial_B H(x)).$$

Bouligand, Clarke (= generalizes the convex subdifferential)



1D example where $\partial_B H(x) = \{-1, +1/2\}$, $\partial_C H(x) = [-1, +1/2] \ni 0$!

Smooth C-functions?

Yes! Mangasarian's framework [Man76]

Family of C-functions and easy condition to get smooth ones:

$$\begin{aligned} &\rho : \mathbb{R} \to \mathbb{R}, \quad \rho(0) = 0, \quad \rho \nearrow, \\ &\rho(|F_i(x) - G_i(x)|) - \rho(F_i(x)) - \rho(G_i(x)) = 0 \ \forall\ i \in [1:n]. \end{aligned} \tag{3}$$

But if x^* is not strictly complementary, then $\nabla H_\varphi(x^*)$ is singular: no fast local convergence [FP03, prop. 9.1.1, pp. 794-795].

Algorithm with nondegeneracy assumption in [Sub93].

Smooth C-functions?

Yes! Mangasarian's framework [Man76]

Family of C-functions and easy condition to get smooth ones:

$$\begin{aligned} &\rho : \mathbb{R} \to \mathbb{R}, \quad \rho(0) = 0, \quad \rho \nearrow, \\ &\rho(|F_i(x) - G_i(x)|) - \rho(F_i(x)) - \rho(G_i(x)) = 0 \ \forall\ i \in [1 : n]. \end{aligned} \tag{3}$$

But if x^* is not strictly complementary, then $\nabla H_\varphi(x^*)$ is singular: no fast local convergence [FP03, prop. 9.1.1, pp. 794-795].

Algorithm with nondegeneracy assumption in [Sub93].

Outline

Globalization

How to get x_0 near x^*? $H(x) = 0 \to \min \theta := \frac{1}{2}||H(x)||^2$

Case of FB (and many others)

$\theta_{\mathrm{FB}} := \frac{1}{2}||H_{\mathrm{FB}}||^2 = \frac{1}{2}\sum_{i=1}^{n} \varphi_{\mathrm{FB}}^2(\ldots_i)$ is smooth.

Concretely, one can use $-\nabla\theta_{\mathrm{FB}}$ as a descent direction...
But $-\nabla$ is limited: often, a last resort (admittedly practical).

Globalization

How to get x_0 near x^*? $H(x) = 0 \to \min \theta := \frac{1}{2}||H(x)||^2$

Case of FB (and many others)

$\theta_{\mathrm{FB}} := \frac{1}{2}||H_{\mathrm{FB}}||^2 = \frac{1}{2}\sum_{i=1}^n \varphi_{\mathrm{FB}}^2(\ldots_i)$ is smooth.

Concretely, one can use $-\nabla\theta_{\mathrm{FB}}$ as a descent direction...
But $-\nabla$ is limited: often, a last resort (admittedly practical).

Globalization

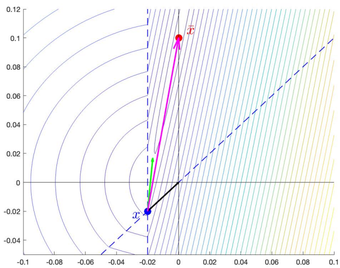
How to get x_0 near x^*? $H(x) = 0 \to \min \theta := \frac{1}{2}||H(x)||^2$

Case of FB (and many others)

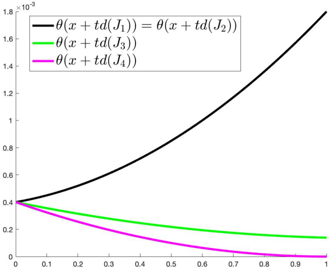
$\theta_{\mathrm{FB}} := \frac{1}{2}||H_{\mathrm{FB}}||^2 = \frac{1}{2}\sum_{i=1}^{n} \varphi_{\mathrm{FB}}^2(\dots_i)$ is smooth.

Concretely, one can use $-\nabla\theta_{\mathrm{FB}}$ as a descent direction...
But $-\nabla$ is limited: often, a last resort (admittedly practical).

Globalization: example hindered by nonsmoothness



(a) Level sets of θ, nonsmooth at the dashed lines; the arrows are directions related to the J's.



(b) Depending on the J's, increase/decrease of θ (see also [Ben12]).

(Similar difficulty in MPECs, for instance [HK09, rem. 5.4 p. 888].)

FB vs minimum

- x_0 / globalization, FB because φ_{FB}^2 is smooth [Mun+01].
- J / ∂_B, minimum better: $\partial_B H(x)$ smaller [IS14, p.151] + finite termination if F and G affine.

- Numerically, no clear winner [DFK96; DFK00; PGP03];
- Combinations between both [KK98];
- adjust φ_{FB} [MS93; SQ99; CCK00] [FP03, fig. 9.1 p. 797]

How to choose J? Differentials are generally unknown.
Either just one J (suffices), or $\partial_B H(x) \subseteq$ [simple].

FB vs minimum

- x_0 / globalization, FB because φ_{FB}^2 is smooth [Mun+01].
- J / ∂_B, minimum better: $\partial_B H(x)$ smaller [IS14, p.151] + finite termination if F and G affine.

- Numerically, no clear winner [DFK96; DFK00; PGP03];
- Combinations between both [KK98];
- adjust φ_{FB} [MS93; SQ99; CCK00] [FP03, fig. 9.1 p. 797]

How to choose J? Differentials are generally unknown.
Either just one J (suffices), or $\partial_B H(x) \subseteq$ [simple].

FB vs minimum

- x_0 / globalization, FB because φ_{FB}^2 is smooth [Mun+01].
- J / ∂_B, minimum better: $\partial_B H(x)$ smaller [IS14, p.151] + finite termination if F and G affine.

- Numerically, no clear winner [DFK96; DFK00; PGP03];
- Combinations between both [KK98];
- adjust φ_{FB} [MS93; SQ99; CCK00] [FP03, fig. 9.1 p. 797]

How to choose J? Differentials are generally unknown.
Either just one J (suffices), or $\partial_B H(x) \subseteq$ [simple].

One simple case

For LCP / linearizations of F, G

$$0 \leqslant Ax + a \perp Bx + b \geqslant 0 \;\Leftrightarrow\; H(x) = \min(Ax + a, Bx + b) = 0$$

$$\partial_B H(x) = \{J \in \mathbb{R}^{n \times n} : \exists\ (x_k)_k \to x, x_k \in \mathcal{D}_H, H'(x_k) \to J\}$$

- Related to hyperplane arrangements: [DGP25].
- Basic topic from combinatorial geometry, studied since early 19th century [Ste26; Rob87; Sch50].

One simple case

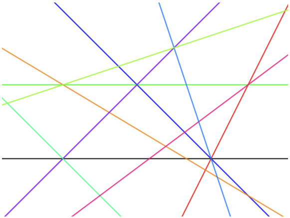
For LCP / linearizations of F, G

$$0 \leqslant Ax + a \perp Bx + b \geqslant 0 \;\Leftrightarrow\; H(x) = \min(Ax + a, Bx + b) = 0$$

$$\partial_B H(x) = \{J \in \mathbb{R}^{n \times n} : \exists\ (x_k)_k \to x, x_k \in \mathcal{D}_H, H'(x_k) \to J\}$$

- Related to hyperplane arrangements: [DGP25].
- Basic topic from combinatorial geometry, studied since early 19th century [Ste26; Rob87; Sch50].

Quick comments



- $|\partial_B H(x)|$ is #P-hard; $\partial_B H(x)$ is at least as difficult.
- Related to: matroids, zonotopes, graphs, geometric / combinatorial notions posets... Families of arrangements...
- Heuristics and "primal-dual" technique (convex duality): state-of-the-art algorithm becomes $\simeq 8$ times faster.

Outline

Recall of Newton-min (NM)

Problem and Newton-min algorithm ([Pan91; Qi93])

$$H(x) := \min(F(x), G(x)) = 0$$

- take a "good" x_0;
- $x_+ = x + d = x - J^{-1}H(x)$, $J \in \partial_B^\times H(x)$ (or $\partial_C^\times H(x)$);
- requires all $J \in \partial_B^\times H(x^*)$ nonsingular.

$$\partial_B^\times H(x) := \partial_B H_1(x) \times \cdots \times \partial_B H_n(x)$$

$\partial_B H(x) \subseteq \partial_B^\times H(x) =$ simple; $\partial_{B(|C)} H$ not really necessary.

Recall of Newton-min (NM)

Problem and Newton-min algorithm ([Pan91; Qi93])

$$H(x) := \min(F(x), G(x)) = 0$$

- take a "good" x_0;
- $x_+ = x + d = x - J^{-1}H(x)$, $J \in \partial_B^\times H(x)$ (or $\partial_C^\times H(x)$);
- requires all $J \in \partial_B^\times H(x^*)$ nonsingular.

$$\partial_B^\times H(x) := \partial_B H_1(x) \times \cdots \times \partial_B H_n(x)$$

$\partial_B H(x) \subseteq \partial_B^\times H(x)$ = simple; $\partial_{B(|C)} H$ not really necessary.

Index sets

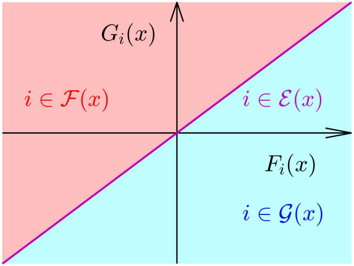
$$\begin{aligned}
&\text{if } F_i(x) < G_i(x), \quad H_i(x) = F_i(x), \quad \partial_B H_i(x) = \{F_i'(x)\},\\
&\text{if } F_i(x) > G_i(x), \quad H_i(x) = G_i(x), \quad \partial_B H_i(x) = \{G_i'(x)\},\\
&\text{if } F_i(x) = G_i(x), \quad = H_i(x), \quad \partial_B H_i(x) = \{F_i'(x), G_i'(x)\}.
\end{aligned}$$

Index sets

$$
\begin{aligned}
&\text{if } F_i(x) < G_i(x), \quad H_i(x) = F_i(x), \quad \partial_B H_i(x) = \{F_i'(x)\},\\
&\text{if } F_i(x) > G_i(x), \quad H_i(x) = G_i(x), \quad \partial_B H_i(x) = \{G_i'(x)\},\\
&\text{if } F_i(x) = G_i(x), \quad = H_i(x), \quad \partial_B H_i(x) = \{F_i'(x), G_i'(x)\}.
\end{aligned}
$$

Index sets

$$
\begin{aligned}
&\text{if } F_i(x) < G_i(x), && H_i(x) = F_i(x), && \partial_B H_i(x) = \{F_i'(x)\},\\
&\text{if } F_i(x) > G_i(x), && H_i(x) = G_i(x), && \partial_B H_i(x) = \{G_i'(x)\},\\
&\text{if } F_i(x) = G_i(x), && = H_i(x), && \partial_B H_i(x) = \{F_i'(x), G_i'(x)\}.
\end{aligned}
$$



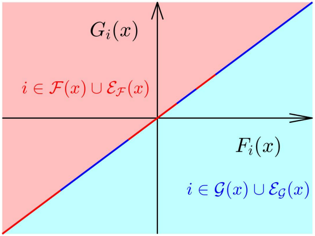
Representation of the index sets.

Newton-min system

NM splits $\mathcal{E}(x) = \mathcal{E}_{\mathcal{F}}(x) \cup \mathcal{E}_{\mathcal{G}}(x)$:

$$\begin{cases} (F(x) + F'(x)d)_{\mathcal{F}(x) \,\cup\, \mathcal{E}_{\mathcal{F}}(x)} & = 0, \\ (G(x) + G'(x)d)_{\mathcal{G}(x) \,\cup\, \mathcal{E}_{\mathcal{G}}(x)} & = 0, \end{cases}$$

simple, good local convergence.



NM index sets.

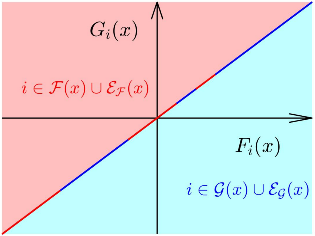
But which partition far from solutions?
$\to$ An augmented system ensuring descent (θ decreases).

Newton-min system

NM splits $\mathcal{E}(x) = \mathcal{E}_{\mathcal{F}}(x) \cup \mathcal{E}_{\mathcal{G}}(x)$:

$$\begin{cases} (F(x) + F'(x)d)_{\mathcal{F}(x)\,\cup\,\mathcal{E}_{\mathcal{F}}(x)} & = 0, \\ (G(x) + G'(x)d)_{\mathcal{G}(x)\,\cup\,\mathcal{E}_{\mathcal{G}}(x)} & = 0, \end{cases}$$

simple, good local convergence.



NM index sets.

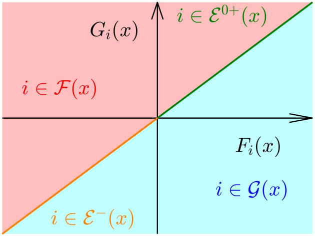
But which partition far from solutions?
$\rightarrow$ An augmented system ensuring descent (θ decreases).

Outline

1. Introduction
2. C-functions
   - Principle and examples
   - Beyond the local method
3. Newton-type methods and beyond
   - Newton-min
   - Polyhedral Newton-Min (PNM)
   - Least-squares over regularity
   - Choice of the weights
   - Complexity and algorithm

Polyhedral Newton-min (PNM) [DFG25]

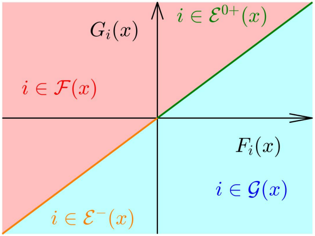
Let $\mathcal{E}^{0+}(x) := \{i : F_i = G_i \geqslant 0\}$ and $\mathcal{E}^{-}(x) := \{i : F_i = G_i < 0\}$,



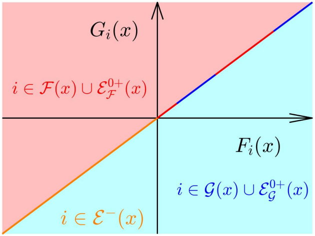
The "positive"/"negative" kinks.

Polyhedral Newton-min (PNM) [DFG25]

Let $\mathcal{E}^{0+}(x) := \{i : F_i = G_i \geqslant 0\}$ and $\mathcal{E}^{-}(x) := \{i : F_i = G_i < 0\}$, $\mathcal{E}^{0+}_{\mathcal{F}}(x) \cup \mathcal{E}^{0+}_{\mathcal{G}}(x)$ be a partition of $\mathcal{E}^{0+}(x)$:



The “positive”/“negative” kinks.



Type of PNM partitioning.

$\mathcal{E}^{0+}(x)$ violates complementarity, $\mathcal{E}^{-}(x)$ also violates $\geqslant 0$.
(See also [PG93; QS94; HPR92].)

Basic PNM method

For $\mathcal{E}_{\mathcal{F}}^{0+}(x)$ and $\mathcal{E}_{\mathcal{G}}^{0+}(x)$, one equality, for $\mathcal{E}^{-}(x)$, **2** inequalities.

$$\text{polyhedron in } d : \begin{cases} F_i + F_i' d = 0 & i \in \mathcal{F}(x) \cup \mathcal{E}_{\mathcal{F}}^{0+}(x), \\ G_i + G_i' d = 0 & i \in \mathcal{G}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x), \\ F_i + F_i' d \geqslant 0 & i \in \mathcal{E}^{-}(x), \\ G_i + G_i' d \geqslant 0 & i \in \mathcal{E}^{-}(x). \end{cases}$$

1st-order info $\Rightarrow \theta'(x; d) \leqslant -2\theta(x) < 0$, θ **decreases** along d.

For a $\mathcal{E}_{\mathcal{F}}^{0+}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x)$, is there a d?

Basic PNM method

For $\mathcal{E}_{\mathcal{F}}^{0+}(x)$ and $\mathcal{E}_{\mathcal{G}}^{0+}(x)$, one equality, for $\mathcal{E}^{-}(x)$, **2** inequalities.

$$\text{polyhedron in } d : \begin{cases} F_i + F_i' d = 0 & i \in \mathcal{F}(x) \cup \mathcal{E}_{\mathcal{F}}^{0+}(x), \\ G_i + G_i' d = 0 & i \in \mathcal{G}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x), \\ F_i + F_i' d \geqslant 0 & i \in \mathcal{E}^{-}(x), \\ G_i + G_i' d \geqslant 0 & i \in \mathcal{E}^{-}(x). \end{cases}$$

1st-order info $\Rightarrow \theta'(x; d) \leqslant -2\theta(x) < 0$, θ **decreases** along d.

For a $\mathcal{E}_{\mathcal{F}}^{0+}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x)$, is there a d?

Basic PNM method

For $\mathcal{E}^{0+}_{\mathcal{F}}(x)$ and $\mathcal{E}^{0+}_{\mathcal{G}}(x)$, one equality, for $\mathcal{E}^{-}(x)$, **2** inequalities.

$$\text{polyhedron in } d : \begin{cases} F_i + F_i' d = 0 & i \in \mathcal{F}(x) \cup \mathcal{E}^{0+}_{\mathcal{F}}(x), \\ G_i + G_i' d = 0 & i \in \mathcal{G}(x) \cup \mathcal{E}^{0+}_{\mathcal{G}}(x), \\ F_i + F_i' d \geqslant 0 & i \in \mathcal{E}^{-}(x), \\ G_i + G_i' d \geqslant 0 & i \in \mathcal{E}^{-}(x). \end{cases}$$

1st-order info $\Rightarrow \theta'(x; d) \leqslant -2\theta(x) < 0$, θ **decreases** along d.

For a $\mathcal{E}^{0+}_{\mathcal{F}}(x) \cup \mathcal{E}^{0+}_{\mathcal{G}}(x)$, is there a d?

PNM regularity

PNM regularity condition to have a d

At a solution x^*, for all partitions of $\mathcal{E}^{0+}(x)$ with x near x^*, a Mangasarian-Fromovitz condition holds at x^* with each partition.

Verified for instance if $M \in \mathbf{P}$

- Hybrid: if NM "works well" ✓ (most often), otherwise PNM.
- Excellent performance on random data / applications.

PNM regularity

PNM regularity condition to have a d

At a solution x^*, for all partitions of $\mathcal{E}^{0+}(x)$ with x near x^*, a Mangasarian-Fromovitz condition holds at x^* with each partition.

Verified for instance if $M \in \mathbf{P}$

- Hybrid: if NM "works well" ✓ (most often), otherwise PNM.
- Excellent performance on random data / applications.

Outline

What if the polyhedron is empty?

The problem is:

- nonsmooth,
- nonconvex,
- without (major) assumption.

Method based on the structure of min.
Combinatorial nature, but no hypothesis on F and G.

What if the polyhedron is empty?

The problem is:

- nonsmooth,
- nonconvex,
- without (major) assumption.

Method based on the structure of min.
Combinatorial nature, but no hypothesis on F and G.

From the PNM system

Least-squares to find "a best possible d", $\min ||(4)||^2/2$:

$$\begin{cases} F_i + F_i'd = 0 & i \in \mathcal{F}(x) \\ G_i + G_i'd = 0 & i \in \mathcal{G}(x) \\ F_i + F_i'd = 0 & i \in \mathcal{E}^{0+}_{\mathcal{F}}(x) & \gamma_i = 1, \overline{\gamma}_i = 0 \\ G_i + G_i'd = 0 & i \in \mathcal{E}^{0+}_{\mathcal{G}}(x) & \gamma_i = 0, \overline{\gamma}_i = 1 \\ F_i + F_i'd \geqslant 0 & i \in \mathcal{E}^{-}(x) & \times \gamma_i \\ G_i + G_i'd \geqslant 0 & i \in \mathcal{E}^{-}(x) & \times \overline{\gamma}_i \end{cases} \tag{4}$$

Indices of $\mathcal{E}^-(x)$ appear twice: **convex weights to balance**, $\gamma_- = (\gamma_i)_i \in [0,1]^{\mathcal{E}^-(x)}$ and $\overline{\gamma}_i := 1 - \gamma_i$.

Same for $\mathcal{E}^{0+}(x)$, $\gamma_+ = (\gamma_i)_i \in \{0,1\}^{\mathcal{E}^{0+}(x)}$, $\overline{\gamma}_i := 1 - \gamma_i$ $\{0,1\}$ for the partition, F_i or G_i.

From the PNM system

Least-squares to find "a best possible d", $\min ||(4)||^2/2$:

$$\begin{cases} F_i + F_i' d = 0 & i \in \mathcal{F}(x) \\ G_i + G_i' d = 0 & i \in \mathcal{G}(x) \\ F_i + F_i' d = 0 & i \in \mathcal{E}^{0+}_{\mathcal{F}}(x) & \gamma_i = 1, \overline{\gamma}_i = 0 \\ G_i + G_i' d = 0 & i \in \mathcal{E}^{0+}_{\mathcal{G}}(x) & \gamma_i = 0, \overline{\gamma}_i = 1 \\ F_i + F_i' d \geqslant 0 & i \in \mathcal{E}^{-}(x) & \times \gamma_i \\ G_i + G_i' d \geqslant 0 & i \in \mathcal{E}^{-}(x) & \times \overline{\gamma}_i \end{cases} \quad (4)$$

Indices of $\mathcal{E}^{-}(x)$ appear twice: **convex weights to balance**, $\gamma_- = (\gamma_i)_i \in [0,1]^{\mathcal{E}^{-}(x)}$ and $\overline{\gamma}_i := 1 - \gamma_i$.

Same for $\mathcal{E}^{0+}(x)$, $\gamma_+ = (\gamma_i)_i \in \{0,1\}^{\mathcal{E}^{0+}(x)}$, $\overline{\gamma}_i := 1 - \gamma_i$
$\{0,1\}$ for the partition, F_i or G_i.

From the PNM system

Least-squares to find "a best possible d", $\min ||(4)||^2/2$:

$$\begin{cases} F_i + F_i' d = 0 & i \in \mathcal{F}(x) & \\ G_i + G_i' d = 0 & i \in \mathcal{G}(x) & \\ F_i + F_i' d = 0 & i \in \mathcal{E}^{0+}_{\mathcal{F}}(x) & \gamma_i = 1, \overline{\gamma}_i = 0 \\ G_i + G_i' d = 0 & i \in \mathcal{E}^{0+}_{\mathcal{G}}(x) & \gamma_i = 0, \overline{\gamma}_i = 1 \\ F_i + F_i' d \geqslant 0 & i \in \mathcal{E}^{-}(x) & \times \gamma_i \\ G_i + G_i' d \geqslant 0 & i \in \mathcal{E}^{-}(x) & \times \overline{\gamma}_i \end{cases} \tag{4}$$

Indices of $\mathcal{E}^-(x)$ appear twice: **convex weights to balance**, $\gamma_- = (\gamma_i)_i \in [0,1]^{\mathcal{E}^-(x)}$ and $\overline{\gamma}_i := 1 - \gamma_i$.

Same for $\mathcal{E}^{0+}(x)$, $\gamma_+ = (\gamma_i)_i \in \{0,1\}^{\mathcal{E}^{0+}(x)}$, $\overline{\gamma}_i := 1 - \gamma_i$ $\{0,1\}$ for the partition, F_i or G_i.

Brief comment on Levenberg-Marquardt (LM)

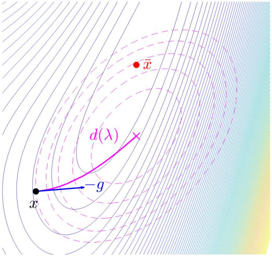
$$\min q(x) := \frac{1}{2}||r(x)||^2, \to \min_d \frac{1}{2}||r(x) + r'(x)d||^2 = q_x(d).$$

Brief comment on Levenberg-Marquardt (LM)

$\min q(x) := \frac{1}{2}||r(x)||^2$, $\to$ $\min_d \frac{1}{2}||r(x) + r'(x)d||^2 = q_x(d)$.
LM: $\min_d q_x(d) + \frac{\lambda}{2}d^\mathsf{T}d$, curve $\lambda \mapsto d(\lambda)$. (matrix $= I$ here)

Brief comment on Levenberg-Marquardt (LM)

$\min q(x) := \frac{1}{2}||r(x)||^2, \to \min_d \frac{1}{2}||r(x) + r'(x)d||^2 = q_x(d)$.
LM: $\min_d q_x(d) + \frac{\lambda}{2} d^\mathsf{T} d$, curve $\lambda \mapsto d(\lambda)$. (matrix $= I$ here)



Smooth example. When $\lambda \searrow 0$, $d(\lambda) \to$ Newton direction, when $\lambda \to +\infty$: $d(\lambda)/||d(\lambda)|| \to -\nabla q(x)/||\nabla q(x)||$ the tangent of the curve.

Least-squares formulation

$q_x(d)/2 := ||(4)||^2/2$ with γ_+, γ_- + Levenberg-Marquardt

$$\min_{d \in \mathbb{R}^n} \psi_x(d) := q_x(d) + \frac{\lambda}{2} d^\mathsf{T} d, \quad \lambda \geqslant 0 \tag{5}$$

- ψ_x piecewise quadratic model of θ at x
- ψ_x **always has a minimizer** d even with empty polyhedron
- $g(\gamma_+, \gamma_-) := \nabla \psi_x(d = 0)$ has a descent property for θ?

$$g(\gamma_+, \gamma_-) = g_0(x) + M_+ \gamma_+ + M_- \gamma_-$$

Do γ_+, γ_- generate a good curve?

Least-squares formulation

$q_x(d)/2 := ||(4)||^2/2$ with γ_+, γ_- + Levenberg-Marquardt

$$\min_{d \in \mathbb{R}^n} \psi_x(d) := q_x(d) + \frac{\lambda}{2} d^\mathsf{T} d, \quad \lambda \geqslant 0 \tag{5}$$

- ψ_x piecewise quadratic model of θ at x
- ψ_x **always has a minimizer** d even with empty polyhedron
- $g(\gamma_+, \gamma_-) := \nabla \psi_x(d = 0)$ has a descent property for θ?

$$g(\gamma_+, \gamma_-) = g_0(x) + M_+ \gamma_+ + M_- \gamma_-$$

Do γ_+, γ_- generate a good curve?

Least-squares formulation

$q_x(d)/2 := ||(4)||^2/2$ with γ_+, γ_- + Levenberg-Marquardt

$$\min_{d \in \mathbb{R}^n} \psi_x(d) := q_x(d) + \frac{\lambda}{2} d^\mathsf{T} d, \quad \lambda \geqslant 0 \tag{5}$$

- ψ_x piecewise quadratic model of θ at x
- ψ_x **always has a minimizer** d even with empty polyhedron
- $g(\gamma_+, \gamma_-) := \nabla\psi_x(d = 0)$ has a descent property for θ?

$$g(\gamma_+, \gamma_-) = g_0(x) + M_+\gamma_+ + M_-\gamma_-$$

Do γ_+, γ_- generate a good curve?

Least-squares formulation

$q_x(d)/2 := ||(4)||^2/2$ with γ_+, γ_- + Levenberg-Marquardt

$$\min_{d \in \mathbb{R}^n} \psi_x(d) := q_x(d) + \frac{\lambda}{2} d^T d, \quad \lambda \geqslant 0 \tag{5}$$

- ψ_x piecewise quadratic model of θ at x
- ψ_x **always has a minimizer** d even with empty polyhedron
- $g(\gamma_+, \gamma_-) := \nabla \psi_x(d = 0)$ has a descent property for θ?

$$g(\gamma_+, \gamma_-) = g_0(x) + M_+ \gamma_+ + M_- \gamma_-$$

Do γ_+, γ_- generate a good curve?

Outline

Computations

Goal: decrease the nonsmooth merit function

Can we choose γ_+, γ_- to have $\theta'(x; -g(\gamma_+, \gamma_-)) \leqslant 0$?

$$\begin{aligned}\theta'(x; -g(\gamma_+, \gamma_-)) = \;& -\|g(\gamma_+, \gamma_-)\|^2 && \} \leqslant 0\\ & +\rho_{\mathcal{E}^{0+}(x)}(x, \gamma_+, g(\gamma_+, \gamma_-)) && \} \leqslant 0\\ & +\rho_{\mathcal{E}^{-}(x)}(x, \gamma_-, g(\gamma_+, \gamma_-)) && \} \geqslant 0\end{aligned}$$

$\rho_{\mathcal{E}^{0+}(x)}$, $\rho_{\mathcal{E}^{-}(x)}$ nonsmooth, quadratic in γ_+, γ_- respectively.

Choosing appropriate weights

Lemma:

Let γ_+, $\exists\ \gamma_-(\gamma_+)$ s.t. $\rho_{\mathcal{E}^-(x)}(x, \gamma_-(\gamma_+), g(\gamma_+, \gamma_-(\gamma_+))) = 0$:

$$\theta'(x; -g(\gamma_+, \gamma_-(\gamma_+))) = -\left\|g(\gamma_+, \gamma_-(\gamma_+))\right\|^2 + \rho_{\mathcal{E}^{0+}(x)}(x, \gamma_+, g(\gamma_+, \gamma_-(\gamma_+))) + 0 \leqslant 0.$$

- Possibly multiple $\gamma_-(\gamma_+)$, $\gamma_+ \mapsto g(\gamma_+, \gamma_-(\gamma_+))$ unique;
- if $g(\gamma_+, \gamma_-(\gamma_+)) = 0$, γ_+ irrelevant;
- **way to choose a good piece**,
- wrong γ_- can make $\theta'(x; -g(\gamma_+, \gamma_-)) \geqslant 0$.

Choosing appropriate weights

Lemma:

Let γ_+, $\exists\ \gamma_-(\gamma_+)$ s.t. $\rho_{\mathcal{E}^-(x)}(x, \gamma_-(\gamma_+), g(\gamma_+, \gamma_-(\gamma_+))) = 0$:

$$\theta'(x; -g(\gamma_+, \gamma_-(\gamma_+))) = -\left\|g(\gamma_+, \gamma_-(\gamma_+))\right\|^2 + \rho_{\mathcal{E}^{0+}(x)}(x, \gamma_+, g(\gamma_+, \gamma_-(\gamma_+))) + 0 \leqslant 0.$$

- Possibly multiple $\gamma_-(\gamma_+)$, $\gamma_+ \mapsto g(\gamma_+, \gamma_-(\gamma_+))$ unique;
- if $g(\gamma_+, \gamma_-(\gamma_+)) = 0$, γ_+ irrelevant;
- **way to choose a good piece**,
- wrong γ_- can make $\theta'(x; -g(\gamma_+, \gamma_-)) \geqslant 0$.

Choosing appropriate weights

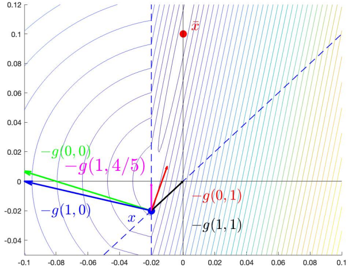
Lemma:

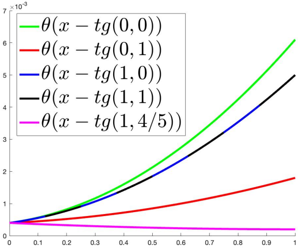
Let γ_+, $\exists\ \gamma_-(\gamma_+)$ s.t. $\rho_{\mathcal{E}^-(x)}(x, \gamma_-(\gamma_+), g(\gamma_+, \gamma_-(\gamma_+))) = 0$:

$$\theta'(x; -g(\gamma_+, \gamma_-(\gamma_+))) = -\|g(\gamma_+, \gamma_-(\gamma_+))\|^2 + \rho_{\mathcal{E}^{0+}(x)}(x, \gamma_+, g(\gamma_+, \gamma_-(\gamma_+))) + 0 \leqslant 0.$$

- Possibly multiple $\gamma_-(\gamma_+)$, $\gamma_+ \mapsto g(\gamma_+, \gamma_-(\gamma_+))$ unique;
- if $g(\gamma_+, \gamma_-(\gamma_+)) = 0$, γ_+ irrelevant;
- **way to choose a good piece**,
- wrong γ_- can make $\theta'(x; -g(\gamma_+, \gamma_-)) \geqslant 0$.

Illustration of the lemma





All partitions ($\{0,1\}^2$) increase θ. Correct weights yield descent.

Relation with PNM

PNM: assumption $\rightarrow$ every $\mathcal{E}_{\mathcal{F}}^{0+}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x)$ (γ_+) is convenient.
LM: one γ_- per γ_+ ($\mathcal{E}_{\mathcal{F}}^{0+}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x)$), bad if $g(\gamma_+, \gamma_-(\gamma_+)) = 0$.

Important link

For a given $\mathcal{E}_{\mathcal{F}}^{0+}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x)$, if there exists a d(PNM), there is NO γ_- s.t. $g = 0$ (every γ_- ✓).

Link based on convex duality [Mot36], weights "generalize" PNM.
Not an equivalence: possible to have no d(PNM) but weights work.

Relation with PNM

PNM: assumption $\to$ every $\mathcal{E}_{\mathcal{F}}^{0+}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x)$ (γ_+) is convenient.
LM: one γ_- per γ_+ $(\mathcal{E}_{\mathcal{F}}^{0+}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x))$, bad if $g(\gamma_+, \gamma_-(\gamma_+)) = 0$.

Important link

For a given $\mathcal{E}_{\mathcal{F}}^{0+}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x)$, if there exists a d(PNM), there is NO γ_- s.t. $g = 0$ (every $\gamma_- \checkmark$).

Link based on convex duality [Mot36], weights "generalize" PNM.
Not an equivalence: possible to have no d(PNM) but weights work.

Relation with PNM

PNM: assumption $\rightarrow$ every $\mathcal{E}_{\mathcal{F}}^{0+}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x)$ (γ_+) is convenient.
LM: one γ_- per γ_+ ($\mathcal{E}_{\mathcal{F}}^{0+}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x)$), bad if $g(\gamma_+, \gamma_-(\gamma_+)) = 0$.

Important link

For a given $\mathcal{E}_{\mathcal{F}}^{0+}(x) \cup \mathcal{E}_{\mathcal{G}}^{0+}(x)$, if there exists a d(PNM), there is NO γ_- s.t. $g = 0$ (every γ_-✓).

Link based on convex duality [Mot36], weights “generalize” PNM.
Not an equivalence: possible to have no d(PNM) but weights work.

Summary

- Robustification of the polyhedral system;
- "control" role for $\mathcal{E}^{0+}(x)$;
- convex weights for $\mathcal{E}^{-}(x)$.

Remaining questions:

- finding γ_+ such that $g(\gamma_+, \gamma_-(\gamma_+)) \neq 0$;
- convergence and algorithmic aspect.

Summary

- Robustification of the polyhedral system;
- "control" role for $\mathcal{E}^{0+}(x)$;
- convex weights for $\mathcal{E}^{-}(x)$.

Remaining questions:

- finding γ_+ such that $g(\gamma_+, \gamma_-(\gamma_+)) \neq 0$;
- convergence and algorithmic aspect.

Outline

Stationarity detection with the weights

We say x is strongly (Dini) stationary if $\forall\ d \in \mathbb{R}^n, \theta'(x;d) \geqslant 0$: 1[st]-order order local optimality.

Stationarity with the γ's:

The following properties are equivalent:

1) x is strongly stationary;
2) for **all** $\gamma_+ \in [0,1]^{\mathcal{E}^{0+}(x)}$, $g(\gamma_+, \gamma_-(\gamma_+)) = 0$.

- For the LCP, if $M \in \mathbf{P}$ (or slightly more specific)
- caracterization of solutions
- conditions with only one (γ_+, γ_-),

Stationarity detection with the weights

We say x is strongly (Dini) stationary if $\forall\ d \in \mathbb{R}^n, \theta'(x; d) \geqslant 0$: 1^{st}-order order local optimality.

Stationarity with the γ's:

The following properties are equivalent:

1) x is strongly stationary;
2) for **all** $\gamma_+ \in [0,1]^{\mathcal{E}^{0+}(x)}$, $g(\gamma_+, \gamma_-(\gamma_+)) = 0$.

- For the LCP, if $M \in \mathbf{P}$ (or slightly more specific)
- caracterization of solutions
- conditions with only one (γ_+, γ_-),

Stationarity detection with the weights

We say x is strongly (Dini) stationary if $\forall\, d \in \mathbb{R}^n, \theta'(x; d) \geqslant 0$:
1st-order order local optimality.

Stationarity with the γ's:

The following properties are equivalent:

1) x is strongly stationary;
2) for **all** $\gamma_+ \in [0,1]^{\mathcal{E}^{0+}(x)}$, $g(\gamma_+, \gamma_-(\gamma_+)) = 0$.

- For the LCP, if $M \in \mathbf{P}$ (or slightly more specific)
- caracterization of solutions
- conditions with only one (γ_+, γ_-),

Complexity estimate via geometry

$M \in \mathbb{R}^{n\times q}$, $Z(M) := M[0,+1]^q = \{M\eta : 0_q \leqslant \eta \leqslant +1_q\} \subseteq \mathbb{R}^n$.
Centrally symmetric polytope, matrix $\times$ hypercube.
Combinatorics, control... [McM71; Zie07; Alt22; ST19]

2) is equivalent to verifying an inclusion between two zonotopes, which is co-NP-complete (in general) [KA21].

- any γ_+ with $g(\gamma_+, \gamma_-(\gamma_+)) \neq 0$ ✓: descent;
- combinatorial aspect in $|\mathcal{E}^{0+}(x)|$ and $|\mathcal{E}^{-}(x)|$, maybe small;
- some conditions $\Rightarrow$ polynomial time.
- $\gamma_-(\gamma_+)$ is a projection on a zonotope.

Complexity estimate via geometry

$M \in \mathbb{R}^{n \times q}$, $Z(M) := M[0, +1]^q = \{M\eta : 0_q \leqslant \eta \leqslant +1_q\} \subseteq \mathbb{R}^n$.
Centrally symmetric polytope, matrix $\times$ hypercube.
Combinatorics, control… [McM71; Zie07; Alt22; ST19]

2) is equivalent to verifying an inclusion between two zonotopes, which is co-NP-complete (in general) [KA21].

- any γ_+ with $g(\gamma_+, \gamma_-(\gamma_+)) \neq 0$ ✓: descent;
- combinatorial aspect in $|\mathcal{E}^{0+}(x)|$ and $|\mathcal{E}^{-}(x)|$, maybe small;
- some conditions $\Rightarrow$ polynomial time.
- $\gamma_-(\gamma_+)$ is a projection on a zonotope.

Complexity estimate via geometry

$M \in \mathbb{R}^{n\times q}$, $Z(M) := M[0,+1]^q = \{M\eta : 0_q \leqslant \eta \leqslant +1_q\} \subseteq \mathbb{R}^n$.
Centrally symmetric polytope, matrix $\times$ hypercube.
Combinatorics, control... [McM71; Zie07; Alt22; ST19]

2) is equivalent to verifying an inclusion between two zonotopes, which is co-NP-complete (in general) [KA21].

- any γ_+ with $g(\gamma_+, \gamma_-(\gamma_+)) \neq 0$ ✓: descent;
- combinatorial aspect in $|\mathcal{E}^{0+}(x)|$ and $|\mathcal{E}^{-}(x)|$, maybe small;
- some conditions $\Rightarrow$ polynomial time.
- $\gamma_-(\gamma_+)$ is a projection on a zonotope.

An iteration of the algorithm

Algorithm Compute γ then usual LM

1: Obtain a suitable pair $\gamma_+, \gamma_-(\gamma_+)$ (or x^k stationary)
2: Get $d(\lambda^k) = \arg\min_d \psi_{x^k}(d)$ using $\lambda^k, \gamma_+, \gamma_-$
3: **while** $d(\lambda^k)$ not suitable **do**
4: increase λ^k and recompute $d(\lambda^k)$
5: **end while**
6: Update $x^{k+1} = x^k + d(\lambda^k)$ then λ^{k+1}

Convergence property

With LM assumptions, $(\theta(x_k))_k$ decreases, $g(x_k) \to 0$.

An iteration of the algorithm

Algorithm Compute γ then usual LM

1: Obtain a suitable pair $\gamma_+, \gamma_-(\gamma_+)$ (or x^k stationary)
2: Get $d(\lambda^k) = \arg\min_d \psi_{x^k}(d)$ using $\lambda^k, \gamma_+, \gamma_-$
3: **while** $d(\lambda^k)$ not suitable **do**
4: increase λ^k and recompute $d(\lambda^k)$
5: **end while**
6: Update $x^{k+1} = x^k + d(\lambda^k)$ then λ^{k+1}

Convergence property

With LM assumptions, $(\theta(x_k))_k$ decreases, $g(x_k) \to 0$.

Into perspective

- good $\gamma_+ \Rightarrow g(\gamma_+, \gamma_-(\gamma_+)) \neq 0$ (for algo),
- optimal $\gamma_+ \Rightarrow g(\gamma_+, \gamma_-(\gamma_+)) \in \partial_C \theta(x) \setminus \{0\}$,
- by ∂_C, if $x_k \to x^*$, $0 \in \partial_C \theta(x^*)$ (weak stationarity).

Paradigm (see [BH20; Car+20; Car+21; JLZ22])

Detecting a stationarity point is co-NP-complete,
obtaining one is much more difficult.

Into perspective

- good $\gamma_+ \Rightarrow g(\gamma_+, \gamma_-(\gamma_+)) \neq 0$ (for algo),
- optimal $\gamma_+ \Rightarrow g(\gamma_+, \gamma_-(\gamma_+)) \in \partial_C \theta(x) \setminus \{0\}$,
- by ∂_C, if $x_k \to x^*$, $0 \in \partial_C \theta(x^*)$ (weak stationarity).

Paradigm (see [BH20; Car+20; Car+21; JLZ22])

Detecting a stationarity point is co-NP-complete,
obtaining one is much more difficult.

Into perspective

- good $\gamma_+ \Rightarrow g(\gamma_+, \gamma_-(\gamma_+)) \neq 0$ (for algo),
- optimal $\gamma_+ \Rightarrow g(\gamma_+, \gamma_-(\gamma_+)) \in \partial_C \theta(x) \setminus \{0\}$,
- by ∂_C, if $x_k \to x^*$, $0 \in \partial_C \theta(x^*)$ (weak stationarity).

Paradigm (see [BH20; Car+20; Car+21; JLZ22])

Detecting a stationarity point is co-NP-complete,
obtaining one is much more difficult.

Into perspective

- good $\gamma_+ \Rightarrow g(\gamma_+, \gamma_-(\gamma_+)) \neq 0$ (for algo),
- optimal $\gamma_+ \Rightarrow g(\gamma_+, \gamma_-(\gamma_+)) \in \partial_C \theta(x) \setminus \{0\}$,
- by ∂_C, if $x_k \to x^*$, $0 \in \partial_C \theta(x^*)$ (weak stationarity).

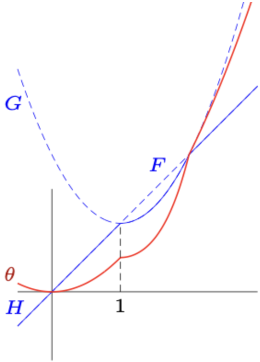
Paradigm (see [BH20; Car+20; Car+21; JLZ22])

Detecting a stationarity point is co-NP-complete,
obtaining one is much more difficult.

"Concave kinks" – difficult (bad) limit points.

Consider a simple example with $F(x) = x,\ G(x) = 1 + (x-1)^2$.

$x = 1$ is weakly stationary but not strongly (even less a solution)



Counter-example

Conclusion

Main take-aways

- H induces a geometric / combinatorial structure (zonotopes)
- explicitation of some results/properties.

Natural pursuits:

- numerical implementation;
- improved LM convergence theorem;
- $|F_i(x) - G_i(x)| < \tau$ instead of $F_i(x) = G_i(x)$.

Thank you for your attention! Any question?

Conclusion

Main take-aways

- H induces a geometric / combinatorial structure (zonotopes)
- explicitation of some results/properties.

Natural pursuits:

- numerical implementation;
- improved LM convergence theorem;
- $|F_i(x) - G_i(x)| < \tau$ instead of $F_i(x) = G_i(x)$.

Thank you for your attention! Any question?

Conclusion

Main take-aways

- H induces a geometric / combinatorial structure (zonotopes)
- explicitation of some results/properties.

Natural pursuits:

- numerical implementation;
- improved LM convergence theorem;
- $|F_i(x) - G_i(x)| < \tau$ instead of $F_i(x) = G_i(x)$.

Thank you for your attention! Any question?

Bibliographic elements I

[AB08] Vincent Acary and Bernard Brogliato. *Numerical Methods for Nonsmooth Dynamical Systems: Applications in Mechanics and Electronics*. Lecture Notes in Applied and Computational Mechanics Ser v.v. 35. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2008. ISBN: 978-3-540-75391-9 978-3-540-75392-6.

[Alc+20] Jan Harold Alcantara et al. "On Construction of New NCP Functions". In: *Operations Research Letters* 48.2 (Mar. 2020), pp. 115–121. ISSN: 01676377. DOI: 10.1016/j.orl.2020.01.002.

[Alt22] Matthias Althoff. *On Computing the Minkowski Difference of Zonotopes*. Aug. 2022. arXiv: 1512.02794 [cs].

[Ben12] Ibtihel Ben Gharbia. "Résolution de Problèmes de Complémentarité.: Application à Un Écoulement Diphasique Dans Un Milieu Poreux". PhD thesis. Université Paris Dauphine Paris IX, 2012.

[BH20] Amir Beck and Nadav Hallak. "On the Convergence to Stationary Points of Deterministic and Randomized Feasible Descent Directions Methods". In: *SIAM Journal on Optimization* 30.1 (Jan. 2020), pp. 56–79. ISSN: 1052-6234, 1095-7189. DOI: 10.1137/18M1217760.

[Car+20] Yair Carmon et al. "Lower Bounds for Finding Stationary Points I". In: *Mathematical Programming* 184.1-2 (Nov. 2020), pp. 71–120. ISSN: 0025-5610, 1436-4646. DOI: 10.1007/s10107-019-01406-y.

[Car+21] Yair Carmon et al. "Lower Bounds for Finding Stationary Points II: First-Order Methods". In: *Mathematical Programming* 185.1-2 (Feb. 2021), pp. 315–355. ISSN: 0025-5610. DOI: 10.48550/arXiv.1711.00841. arXiv: 1711.00841 [math].

[CCK00] Bintong Chen, Xiaojun Chen, and Christian Kanzow. "A Penalized Fischer-Burmeister NCP-function: Theoretical Investigation and Numerical Results". In: *Mathematical Programming* 88.1 (June 2000), pp. 211–216. ISSN: 0025-5610. DOI: 10.1007/PL00011375.

Bibliographic elements II

[Chu89] Sung Jin Chung. "NP-Completeness of the Linear Complementarity Problem". In: *Journal of Optimization Theory and Applications* 60.3 (Mar. 1989), pp. 393–399. ISSN: 0022-3239, 1573-2878. DOI: 10.1007/BF00940344.

[Cla90] Frank H. Clarke. *Optimization and Nonsmooth Analysis*. second. Vol. 1. Classics in Applied Mathematics. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics (SIAM), 1990. ISBN: 0-89871-256-4.

[CNQ00] Xiaojun Chen, Zuhair Nashed, and Liqun Qi. "Smoothing Methods and Semismooth Methods for Nondifferentiable Operator Equations". In: *SIAM Journal on Numerical Analysis* 38.4 (Jan. 2000), pp. 1200–1216. ISSN: 0036-1429, 1095-7170. DOI: 10.1137/S0036142999356719.

[CPS09] Richard Warren Cottle, Jong-Shi Pang, and Richard E. Stone. *The Linear Complementarity Problem*. SIAM. Classics in Applied Mathematics 60. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics (SIAM), 2009. ISBN: 978-0-89871-686-3.

[DFG25] Jean-Pierre Dussault, Mathieu Frappier, and Jean Charles Gilbert. "Polyhedral Newton-min Algorithms for Complementarity Problems". In: *Mathematical Programming* (2025).

[DFK00] Tecla De Luca, Francisco Facchinei, and Christian Kanzow. "A Theoretical and Numerical Comparison of Some Semismooth Algorithms for Complementarity Problems". In: *Computational Optimization and Applications* 16 (Jan. 2000), pp. 173–205. DOI: 10.1023/A:1008705425484.

[DFK96] Tecla De Luca, Francisco Facchinei, and Christian Kanzow. "A Semismooth Equation Approach to the Solution of Nonlinear Complementarity Problems". In: *Mathematical Programming* 75.3 (Dec. 1996), pp. 407–439. ISSN: 0025-5610, 1436-4646. DOI: 10.1007/BF02592192.

[DGP25] Jean-Pierre Dussault, Jean Charles Gilbert, and Baptiste Plaquevent-Jourdain. "On the B-differential of the Componentwise Minimum of Two Affine Vector Functions". In: *Mathematical Programming Computation* (2025).

Bibliographic elements III

[Fat79] Yahya Fathi. "Computational Complexity of LCPs Associated with Positive Definite Symmetric Matrices". In: *Mathematical Programming* 17.1 (Dec. 1979), pp. 335–344. ISSN: 0025-5610, 1436-4646. DOI: 10.1007/BF01588254.

[Fis92] Andreas Fischer. "A Special Newton-type Optimization Method". In: *Optimization* 24 (Jan. 1992), pp. 269–284. DOI: 10.1080/02331939208843795.

[FJ00] Andreas Fischer and Houyuan Jiang. "Merit Functions for Complementarity and Related Problems: A Survey". In: *Computational Optimization and Applications* 17 (Dec. 2000), pp. 159–182. DOI: 10.1023/A:1026598214921.

[FP03] Francisco Facchinei and Jong-Shi Pang, eds. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer Series in Operations Research and Financial Engineering. New York, NY: Springer-Verlag New York, Inc, 2003. ISBN: 978-0-387-95580-3 978-0-387-21814-4. DOI: 10.1007/b97543.

[FP97] Michael C. Ferris and Jong-Shi Pang. "Engineering and Economic Applications of Complementarity Problems". In: *SIAM Review* 39.4 (Jan. 1997), pp. 669–713. ISSN: 0036-1445, 1095-7200. DOI: 10.1137/S0036144595285963.

[Fuk92] Masao Fukushima. "Equivalent Differentiable Optimization Problems and Descent Methods for Asymmetric Variational Inequality Problems". In: *Mathematical Programming* 53.1-3 (Jan. 1992), pp. 99–110. ISSN: 0025-5610, 1436-4646. DOI: 10.1007/BF01585696.

[Gal12] Aurél Galántai. "Properties and Construction of NCP Functions". In: *Computational Optimization and Applications* 52.3 (July 2012), pp. 805–824. ISSN: 0926-6003, 1573-2894. DOI: 10.1007/s10589-011-9428-9.

[Had09] Mounir Haddou. "A New Class of Smoothing Methods for Mathematical Programs with Equilibrium Constraints". In: *Pacific Journal of Optimization* 5 (2009), pp. 87–95.

Bibliographic elements IV

[HK09] Michael Hintermüller and Ian Kopacka. "Mathematical Programs with Complementarity Constraints in Function Space: C- and Strong Stationarity and a Path-Following Algorithm". In: *SIAM Journal on Optimization* 20.2 (Jan. 2009), pp. 868–902. ISSN: 1052-6234, 1095-7189. DOI: 10.1137/080720681.

[HP90] Patrick T. Harker and Jong-Shi Pang. "Finite-Dimensional Variational Inequality and Nonlinear Complementarity Problems: A Survey of Theory, Algorithms and Applications". In: *Mathematical Programming* 48.1-3 (Mar. 1990), pp. 161–220. ISSN: 0025-5610, 1436-4646. DOI: 10.1007/BF01582255.

[HPR92] Shih-Ping Han, Jong-Shi Pang, and Narayan Rangaraj. "Globally Convergent Newton Methods for Nonsmooth Equations". In: *Mathematics of Operations Research* 17.3 (Aug. 1992), pp. 586–607. ISSN: 0364-765X, 1526-5471. DOI: 10.1287/moor.17.3.586.

[IS14] Alexey F. Izmailov and Mikhail V. Solodov. *Newton-Type Methods for Optimization and Variational Problems*. Springer Series in Operations Research and Financial Engineering. Cham: Springer International Publishing, 2014. ISBN: 978-3-319-04246-6 978-3-319-04247-3. DOI: 10.1007/978-3-319-04247-3.

[JLZ22] Michael I. Jordan, Tianyi Lin, and Manolis Zampetakis. *On the Complexity of Deterministic Nonsmooth and Nonconvex Optimization*. Nov. 2022. DOI: 10.48550/arXiv.2209.12463. arXiv: 2209.12463 [math].

[Jos79] Norman H. Josephy. *Newton's Method for Generalized Equations*. Tech. rep. ADA077096. wisconsin: University of Madison, 1979, p. 37.

[KA21] Adrian Kulmburg and Matthias Althoff. "On the Co-NP-completeness of the Zonotope Containment Problem". In: *European Journal of Control* 62 (Nov. 2021), pp. 84–91. ISSN: 09473580. DOI: 10.1016/j.ejcon.2021.06.028.

Bibliographic elements V

[KK98] Christian Kanzow and Helmut Kleinmichel. "A New Class of Semismooth Newton-Type Methods for Nonlinear Complementarity Problems". In: *Computational Optimization and Applications* 11 (1998), pp. 227–251.

[Koj+91] Masakazu Kojima et al. *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*. Lecture Notes in Computer Science 538. Berlin: Springer, Jan. 1991. ISBN: 978-3-540-54509-5.

[Kos76] Michael Martin Kostreva. "Direct Algorithms for Complementarity Problems". PhD thesis. Ann Arbor, MI: Rensselaer Polytechnic Institute, 1976.

[Kum88] Bernd Kummer. "NEWTON's METHOD FOR NON-DIFFERENTIABLE FUNCTIONS". In: *Advances in Mathematical Optimization*. Ed. by J. Guddat Et Al. De Gruyter, Dec. 1988, pp. 114–125. ISBN: 978-3-11-247992-6. DOI: 10.1515/9783112479926-011.

[KYF97] Christian Kanzow, Nobuo Yamashita, and Masao Fukushima. "New NCP-Functions and Their Properties". In: *Journal of Optimization Theory and Applications* 94.1 (July 1997), pp. 115–135. ISSN: 0022-3239. DOI: 10.1023/A:1022659603268.

[Lem65] Carlton E. Lemke. "Bimatrix Equilibrium Points and Mathematical Programming". In: *Management Science* 11.7 (May 1965), pp. 681–689. ISSN: 0025-1909, 1526-5501. DOI: 10.1287/mnsc.11.7.681.

[LT97] Zhi-Quan Luo and Paul Tseng. "A New Class of Merit Functions for the Nonlinear Complementarity Problem". In: *Complementarity and Variational Problems (Baltimore, MD, 1995)*. SIAM, Philadelphia, PA, 1997, pp. 204–225. ISBN: 0-89871-391-9.

[Man76] Olvi Leon Mangasarian. "Equivalence of the Complementarity Problem to a System of Nonlinear Equations". In: *SIAM Journal on Applied Mathematics* 31.1 (July 1976), pp. 89–92. ISSN: 0036-1399, 1095-712X. DOI: 10.1137/0131009.

Bibliographic elements VI

[McM71] Peter McMullen. “On Zonotopes”. In: *Transactions of the American Mathematical Society* 159 (Sept. 1971), pp. 91–109. DOI: 10.2307/1996000.

[Mot36] Theodore S. Motzkin. *Beiträge zur Theorie der linearen Ungleichungen*. Tech. rep. Jerusalem, Israel: University Basel, 1936.

[MS93] Olvi Leon Mangasarian and Mikhail V. Solodov. “Nonlinear Complementarity as Unconstrained and Constrained Minimization”. In: *Mathematical Programming* 62.1-3 (Feb. 1993), pp. 277–297. ISSN: 0025-5610. DOI: 10.1007/BF01585171.

[Mun+01] Todd S. Munson et al. “The Semismooth Algorithm for Large Scale Complementarity Problems”. In: *INFORMS Journal on Computing* 13.4 (Nov. 2001), pp. 294–311. ISSN: 1091-9856, 1526-5528. DOI: 10.1287/ijoc.13.4.294.9734.

[Mur78] Katta G. Murty. “Computational Complexity of Complementary Pivot Methods”. In: *Complementarity and Fixed Point Problems* 7 (1978), pp. 61–73. DOI: 10.1007/BFb0120782.

[Pan90] Jong-Shi Pang. “Newton’s Method for B-Differentiable Equations”. In: *Mathematics of Operations Research* 15.2 (May 1990), pp. 311–341. ISSN: 0364-765X, 1526-5471. DOI: 10.1287/moor.15.2.311.

[Pan91] Jong-Shi Pang. “A B-differentiable Equation-Based, Globally and Locally Quadratically Convergent Algorithm for Nonlinear Programs, Complementarity and Variational Inequality Problems”. In: *Mathematical Programming* 51.1-3 (July 1991), pp. 101–131. ISSN: 0025-5610, 1436-4646. DOI: 10.1007/BF01586928.

[PG93] Jong-Shi Pang and Steven A. Gabriel. “NE/SQP: A Robust Algorithm for the Nonlinear Complementarity Problem”. In: *Mathematical Programming* 60.1-3 (June 1993), pp. 295–337. ISSN: 0025-5610, 1436-4646. DOI: 10.1007/BF01580617.

Bibliographic elements VII

[PGP03] Sandra Pieraccini, Maria Grazia Gasparo, and Aldo Pasquali. "Global Newton-type Methods and Semismooth Reformulations for NCP". In: *Applied Numerical Mathematics* 44.3 (Feb. 2003), pp. 367–384. ISSN: 01689274. DOI: 10.1016/S0168-9274(02)00169-1.

[Qi93] Liqun Qi. "Convergence Analysis of Some Algorithms for Solving Nonsmooth Equations". In: *Mathematics of Operations Research* 18.1 (1993), pp. 227–244.

[QS94] Liqun Qi and Jie Sun. "A Trust Region Algorithm for Minimization of Locally Lipschitzian Functions". In: *Mathematical Programming* 66.1-3 (Aug. 1994), pp. 25–43. ISSN: 0025-5610, 1436-4646. DOI: 10.1007/BF01581136.

[Rad19] Hans Rademacher. "Über partielle und totale Differenzierbarkeit von Funktionen mehrerer Variablen und über die Transformation der Doppelintegrale". In: *Matematische Annalen* 79 (Jan. 1919). DOI: 10.1007/BF01498415.

[Rob80] Stephen M. Robinson. "Strongly Regular Generalized Equations". In: *Mathematics of Operations Research* 5.1 (Feb. 1980), pp. 43–62. ISSN: 0364-765X, 1526-5471. DOI: 10.1287/moor.5.1.43.

[Rob87] Samuel Roberts. "On the Figures Formed by the Intercepts of a System of Straight Lines in a, Plane, and on Analogous Relations in Space of Three Dimensions". In: *Proceedings of the London Mathematical Society* s1-19.1 (Nov. 1887), pp. 405–422. ISSN: 00246115. DOI: 10.1112/plms/s1-19.1.405.

[Sch50] Ludwig Schläfli. *Gesammelte mathematische Abhandlugen*. Springer, Basel: Birkhäuser, 1950.

[SQ99] Defeng Sun and Liqun Qi. "On NCP-functions". In: *Computational Optimization and Applications* 13 (1999), pp. 201–220. DOI: 10.1023/A:1008669226453.

[ST19] Sadra Sadraddini and Russ Tedrake. *Linear Encodings for Polytope Containment Problems*. Mar. 2019. arXiv: 1903.05214 [math].

Bibliographic elements VIII

[Ste26] Jakob Steiner. "Einige Gesetze über die Theilung der Ebene und des Raumes.". In: *J. Reine Angew. Math* (1826), pp. 349–364.

[Sub93] Pudukkottai K. Subramanian. "Gauss-Newton Methods for the Complementarity Problem". In: *Journal of Optimization Theory and Applications* 77.3 (June 1993), pp. 467–482. ISSN: 0022-3239, 1573-2878. DOI: 10.1007/BF00940445.

[Vu+21] Duc Thach Son Vu et al. "A New Approach for Solving Nonlinear Algebraic Systems with Complementarity Conditions. Application to Compositional Multiphase Equilibrium Problems". In: *Mathematics and Computers in Simulation* 190 (Dec. 2021), pp. 1243–1274. ISSN: 03784754. DOI: 10.1016/j.matcom.2021.07.015.

[Wie82] Andrzej P. Wierzbicki. "Note on the Equivalence of Kuhn-Tucker Complementarity Conditions to an Equation". In: *Journal of Optimization Theory and Applications* 37.3 (July 1982), pp. 401–405. ISSN: 0022-3239, 1573-2878. DOI: 10.1007/BF00935279.

[XC11] Shuhuang Xiang and Xiaojun Chen. "Computation of Generalized Differentials in Nonlinear Complementarity Problems". In: *Computational Optimization and Applications* 50.2 (Oct. 2011), pp. 403–423. ISSN: 0926-6003, 1573-2894. DOI: 10.1007/s10589-010-9349-z.

[Zie07] Günter M. Ziegler. *Lectures on Polytopes*. 7th. Vol. 152. Graduate Texts in Mathematics. New York, NY: Springer New York, 2007. ISBN: 978-0-387-94365-7 978-1-4613-8431-1. DOI: 10.1007/978-1-4613-8431-1.

Summary of some basic methods

		$\partial_?$	reg(ularity) at x^*	$\|\partial_?\|$	differentiable?
	NM	$\partial_B^\times H$	b-reg,	2^p J	piecewise
SNM	φ_{FB}	$\partial_C^{(\times)} H_{\mathrm{FB}}$	CD(FB)-reg,	© (“ball”)	SC^1
		$\partial_B^{(\times)} H_{\mathrm{FB}}$	BD(FB)-reg,	© (“sphere”)	
	min	$\partial_C^\times H$	CD(min)-reg,	© (“cube”)	piecewise
		$\partial_B H$	smaller b-reg,	$\leqslant 2^p$ J	piecewise

Table: Properties of some nonsmooth methods. SNM: Semismooth Newton Method; NM: Newton-Min; © : continuum

$$(\partial_C^\times H(x) := \prod_{i=1}^{n} \partial_C H_i(x))$$

Computing the B-differential - multiple indices

$$H \text{ non} - \text{diff in } x_k \Leftrightarrow \exists\, i : (Ax_k + a)_i \overset{C1}{=} (Bx_k + b)_i, A_{i,:} \overset{C2}{\neq} B_{i,:}$$

$$I(x) := \{i \in [1:n] : A_{i,:}x + a_i = B_{i,:}x + b_i, A_{i,:} \neq B_{i,:}\}; |I(x)| = p$$

$$(Ax_k + a)_i \overset{C1}{=} (Bx_k + b)_i \overset{x_k = x + d_k}{\Leftrightarrow} (Ax + a)_i + A_{i,:}d_k = (Bx + b)_i + B_{i,:}d_k$$

$$\Leftrightarrow A_{i,:}d_k = B_{i,:}d_k \Leftrightarrow d_k \in v_i^{\perp}$$

$H_i := (B_{i,:} - A_{i,:})^{\perp} := v_i^{\perp}$, $v_i \neq 0$ ($C2$) ; for ∂_B, $\mathbb{R}^n \backslash (x + \cup_{i=1}^{p} H_i)$

$$\mathbb{R}^n = H_i^- \cup H_i \cup H_i^+, \quad \begin{cases} H_i^- = \{x \in \mathbb{R}^n : v_i^T x < 0\} \\ H_i^+ = \{x \in \mathbb{R}^n : v_i^T x > 0\} \end{cases}$$

$$H_i^- \Leftrightarrow B_{i,:}d - A_{i,:}d < 0 \Leftrightarrow \min(\dots) = (B\dots)_i \Leftrightarrow J_{i,:} = B_{i,:}$$
$$H_i^+ \Leftrightarrow B_{i,:}d - A_{i,:}d > 0 \Leftrightarrow \min(\dots) = (A\dots)_i \Leftrightarrow J_{i,:} = A_{i,:}$$

Computing the B-differential - multiple indices

$$H \text{ non} - \text{diff in } x_k \Leftrightarrow \exists\ i : (Ax_k + a)_i \overset{C1}{=} (Bx_k + b)_i, A_{i,:} \overset{C2}{\neq} B_{i,:}$$

$$I(x) := \{i \in [1:n] : A_{i,:}x + a_i = B_{i,:}x + b_i, A_{i,:} \neq B_{i,:}\}; |I(x)| = p$$

$$\begin{aligned}(Ax_k + a)_i \overset{C1}{=} (Bx_k + b)_i \overset{x_k = x + d_k}{\Leftrightarrow} & (Ax + a)_i + A_{i,:}d_k \\ & = (Bx + b)_i + B_{i,:}d_k \\ & \Leftrightarrow\ A_{i,:}d_k = B_{i,:}d_k \Leftrightarrow d_k \in v_i^{\perp}\end{aligned}$$

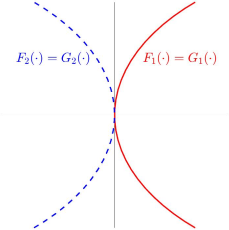
$H_i := (B_{i,:} - A_{i,:})^{\perp} := v_i^{\perp}$, $v_i \neq 0$ ($C2$) ; for ∂_B, $\mathbb{R}^n \backslash (x + \cup_{i=1}^{p} H_i)$

$$\mathbb{R}^n = H_i^- \cup H_i \cup H_i^+, \quad \begin{cases} H_i^- = \{x \in \mathbb{R}^n : v_i^T x < 0\} \\ H_i^+ = \{x \in \mathbb{R}^n : v_i^T x > 0\} \end{cases}$$

$$H_i^- \Leftrightarrow B_{i,:}d - A_{i,:}d < 0 \Leftrightarrow \min(\dots) = (B\dots)_i \Leftrightarrow J_{i,:} = B_{i,:}$$
$$H_i^+ \Leftrightarrow B_{i,:}d - A_{i,:}d > 0 \Leftrightarrow \min(\dots) = (A\dots)_i \Leftrightarrow J_{i,:} = A_{i,:}$$

If functions are not affine

- Sets $\{y \text{ near } x : F_i(y) = G_i(y)\}$ are not hyperplanes.
- First order $F_i(y) = F_i(x) + F_i'(x)(y-x)$ (and G): $\rightarrow$ affine case with hyperplanes and yields a subset of the real $\partial_B H(x)$.
- Much harder due to “higher order”...



Here, linearizations yield the same hyperplane: two chambers, but there is a third from points between the two curves.

PNM computations

$$\theta'(x;d) = \sum_{i\in\mathcal{F}(x)} F_i(x)F_i'(x)d + \sum_{i\in\mathcal{G}(x)} G_i(x)G_i'(x)d$$

$$+ \sum_{i\in\mathcal{E}(x)} H_i(x)\min(F_i'(x)d, G_i'(x)d)$$

$$[d \text{ s.t.} \dots] = -2\theta(x) \quad [\text{smooth}] \qquad [\downarrow \text{nonsmooth} \downarrow]$$

$$+ \sum_{i\in\mathcal{E}^{0+}(x)} \overbrace{H_i(x)}^{\geqslant 0} \overbrace{\min(F_i(x)+F_i'(x)d, G_i(x)+G_i'(x)d)}^{\text{one is } 0, \text{ so } \min \leqslant 0}$$

$$+ \sum_{i\in\mathcal{E}^{-}(x)} \overbrace{H_i(x)}^{<0} \overbrace{\min(F_i(x)+F_i'(x)d, G_i(x)+G_i'(x)d)}^{\text{both are } \geqslant 0, \text{ so } \min \geqslant 0}$$

$$= -2\theta(x) - \dots - \dots \leqslant -2\theta(x) \quad (< 0)$$

PNM computations

$$\theta'(x;d) = \sum_{i\in\mathcal{F}(x)} F_i(x)F_i'(x)d + \sum_{i\in\mathcal{G}(x)} G_i(x)G_i'(x)d + \sum_{i\in\mathcal{E}(x)} H_i(x)\min(F_i'(x)d, G_i'(x)d)$$

$$[d \text{ s.t.} \ldots] = -2\theta(x) \quad [\text{smooth}] \qquad [\downarrow \text{nonsmooth} \downarrow]$$

$$+ \sum_{i\in\mathcal{E}^{0+}(x)} \overbrace{H_i(x)}^{\geqslant 0} \overbrace{\min(F_i(x)+F_i'(x)d, G_i(x)+G_i'(x)d)}^{\text{one is } 0, \text{ so min } \leqslant 0}$$

$$+ \sum_{i\in\mathcal{E}^{-}(x)} \overbrace{H_i(x)}^{<0} \overbrace{\min(F_i(x)+F_i'(x)d, G_i(x)+G_i'(x)d)}^{\text{both are } \geqslant 0, \text{ so min } \geqslant 0}$$

$$= -2\theta(x) - \cdots - \cdots \leqslant -2\theta(x) \quad (<0)$$

PNM computations

$$\theta'(x;d) = \sum_{i\in\mathcal{F}(x)} F_i(x)F_i'(x)d + \sum_{i\in\mathcal{G}(x)} G_i(x)G_i'(x)d$$

$$+ \sum_{i\in\mathcal{E}(x)} H_i(x)\min(F_i'(x)d, G_i'(x)d)$$

$$[d \text{ s.t.} \dots] = -2\theta(x) \quad [\text{smooth}] \qquad [\downarrow \text{nonsmooth} \downarrow]$$

$$+ \sum_{i\in\mathcal{E}^{0+}(x)} \overbrace{H_i(x)}^{\geqslant 0} \overbrace{\min(F_i(x)+F_i'(x)d, G_i(x)+G_i'(x)d)}^{\text{one is } 0, \text{ so min } \leqslant 0}$$

$$+ \sum_{i\in\mathcal{E}^{-}(x)} \overbrace{H_i(x)}^{<0} \overbrace{\min(F_i(x)+F_i'(x)d, G_i(x)+G_i'(x)d)}^{\text{both are } \geqslant 0, \text{ so min } \geqslant 0}$$

$$= -2\theta(x) - \dots - \dots \leqslant -2\theta(x) \quad (<0)$$

PNM computations

$$\theta'(x; d) = \sum_{i \in \mathcal{F}(x)} F_i(x) F_i'(x) d + \sum_{i \in \mathcal{G}(x)} G_i(x) G_i'(x) d$$

$$+ \sum_{i \in \mathcal{E}(x)} H_i(x) \min(F_i'(x) d, G_i'(x) d)$$

$$[d \text{ s.t.} \dots] = \; -2\theta(x) \quad [\text{smooth}] \qquad [\downarrow \text{nonsmooth} \downarrow]$$

$$+ \sum_{i \in \mathcal{E}^{0+}(x)} \overbrace{H_i(x)}^{\geqslant 0} \overbrace{\min(F_i(x) + F_i'(x) d, G_i(x) + G_i'(x) d)}^{\text{one is } 0, \text{ so min } \leqslant 0}$$

$$+ \sum_{i \in \mathcal{E}^{-}(x)} \overbrace{H_i(x)}^{< 0} \overbrace{\min(F_i(x) + F_i'(x) d, G_i(x) + G_i'(x) d)}^{\text{both are } \geqslant 0, \text{ so min } \geqslant 0}$$

$$= -2\theta(x) - \dots - \dots \leqslant -2\theta(x) \quad (< 0)$$

Explicit computations

Let $\Gamma_+ = \mathrm{Diag}(\gamma_+), \Gamma_- = \mathrm{Diag}(\gamma_-)$

Explicit computations

Let $\Gamma_+ = \mathrm{Diag}(\gamma_+), \Gamma_- = \mathrm{Diag}(\gamma_-)$

$$\theta'(x; -g(\gamma_+, \gamma_-)) = -||g(\gamma_+, \gamma_-)||^2$$
$$+ H^{\mathsf{T}}_{\mathcal{E}^{0+}(x)} \underbrace{[\min(-F'_{\mathcal{E}^{0+}(x)} g(\gamma_+, \gamma_-), -G'_{\mathcal{E}^{0+}(x)} g(\gamma_+, \gamma_-)) + \Gamma_+ F'_{\mathcal{E}^{0+}(x)} g(\gamma_+, \gamma_-) + \overline{\Gamma}_+ G'_{\mathcal{E}^{0+}(x)} g(\gamma_+, \gamma_-)]}$$
$$+ H^{\mathsf{T}}_{\mathcal{E}^{-}(x)} \underbrace{[\min(-F'_{\mathcal{E}^{-}(x)} g(\gamma_+, \gamma_-), -G'_{\mathcal{E}^{-}(x)} g(\gamma_+, \gamma_-)) + \Gamma_- F'_{\mathcal{E}^{-}(x)} g(\gamma_+, \gamma_-) + \overline{\Gamma}_- G'_{\mathcal{E}^{-}(x)} g(\gamma_+, \gamma_-)]}$$

Explicit computations

Let $\Gamma_+ = \mathrm{Diag}(\gamma_+), \Gamma_- = \mathrm{Diag}(\gamma_-)$

$$\theta'(x; -g(\gamma_+, \gamma_-)) = -||g(\gamma_+, \gamma_-)||^2$$

$$0 \geqslant \ + H^{\mathsf{T}}_{\mathcal{E}^{0+}(x)}\underbrace{[\min(-F'_{\mathcal{E}^{0+}(x)}g(\gamma_+, \gamma_-), -G'_{\mathcal{E}^{0+}(x)}g(\gamma_+, \gamma_-)) + \Gamma_+ F'_{\mathcal{E}^{0+}(x)}g(\gamma_+, \gamma_-) + \overline{\Gamma}_+ G'_{\mathcal{E}^{0+}(x)}g(\gamma_+, \gamma_-)]}_{\leqslant 0}$$

$$0 \leqslant \ + H^{\mathsf{T}}_{\mathcal{E}^{-}(x)}\underbrace{[\min(-F'_{\mathcal{E}^{-}(x)}g(\gamma_+, \gamma_-), -G'_{\mathcal{E}^{-}(x)}g(\gamma_+, \gamma_-)) + \Gamma_- F'_{\mathcal{E}^{-}(x)}g(\gamma_+, \gamma_-) + \overline{\Gamma}_- G'_{\mathcal{E}^{-}(x)}g(\gamma_+, \gamma_-)]}_{\leqslant 0}$$

$$\min(-a, -b) + \gamma a + \overline{\gamma} b = \begin{cases} \overline{\gamma}(b-a) \leqslant 0 & \text{if } a \geqslant b \\ \gamma(a-b) \leqslant 0 & \text{if } a \leqslant b \end{cases}.$$

Computation of the weights

Goal: get a nonzero g

$$\max_{\gamma_+ \in [0,1]^{\mathcal{E}^{0+}(x)}} \min_{\gamma_- \in [0,1]^{\mathcal{E}^{-}(x)}} ||g(\gamma_+, \gamma_-)||^2/2$$

where $g(\gamma_+, \gamma_-) = g_0 + M_+\gamma_+ + M_-\gamma_-$.

min: $\gamma_-(\gamma_+)$ to descend; max: distance (convex) on hypercube: $\rightsquigarrow \{0,1\}^{\mathcal{E}^{0+}(x)}$ (partitions); inclusion of polytopes: vertices suffice.

Computation of the weights

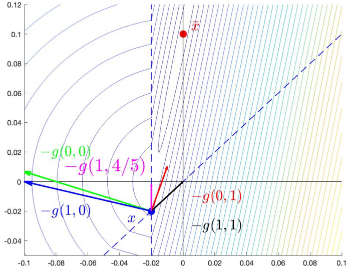
Goal: get a nonzero g

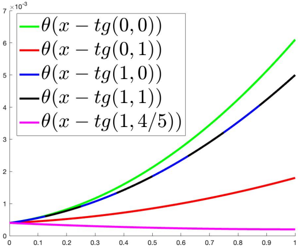
$$\max_{\gamma_+ \in [0,1]^{\mathcal{E}^{0+}(x)}} \min_{\gamma_- \in [0,1]^{\mathcal{E}^{-}(x)}} ||g(\gamma_+, \gamma_-)||^2/2$$

where $g(\gamma_+, \gamma_-) = g_0 + M_+\gamma_+ + M_-\gamma_-$.

min: $\gamma_-(\gamma_+)$ to descend; max: distance (convex) on hypercube: $\rightsquigarrow \{0,1\}^{\mathcal{E}^{0+}(x)}$ (partitions); inclusion of polytopes: vertices suffice.

Example of weights computation – 1





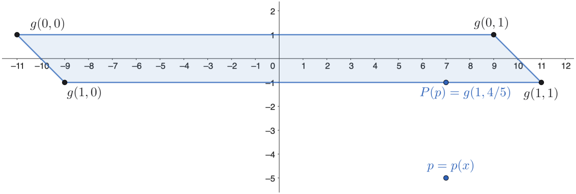
$$M = \begin{pmatrix} 1/2 & 1/2 \\ -5 & 1 \end{pmatrix}, \quad q = \begin{pmatrix} 0 \\ -1/10 \end{pmatrix}, \quad x = \begin{pmatrix} -1/50 \\ -1/50 \end{pmatrix}$$

Here, $\mathcal{F}(x) = \varnothing = \mathcal{G}(x)$, $\mathcal{E}(x) = \{1, 2\}$. After computations. . .

Example of weights computation – 2

[up to factor $1/200$] The point to project is $(7; -5)(= g_1(x))$.

The zonotope to project on is $\begin{bmatrix} 1 & 10 \\ -1 & 0 \end{bmatrix} \times [-1, +1]^2$



The projection is $(7; -1) = 4/5 \times (11; -1) + 1/5 \times (-9; -1)$, and also $g(1, 4/5) = 4/5 g(1, 1) + 1/5 g(1, 0)$

Some precisions for LCPs (1)

$$0 \leqslant F(x) = x \perp Mx + q = G(x) \geqslant 0$$
$$A = \mathcal{F}(x), \quad E = \mathcal{E}(x), \quad I = \mathcal{G}(x)$$

$$\mathcal{C}_{\text{ND}} \begin{cases} M_{I,I} \text{ nonsingular and} \\ M_{E,E} - M_{E,I} M_{I,I}^{-1} M_{I,E} \in \boldsymbol{ND} \end{cases}$$

Weighted framework for LCPs with $\mathcal{C}_{\text{ND}}$

The following properties are equivalent:

1) x is a solution,
2) for any (γ_+, γ_-), $g(\gamma_+, \gamma_-) = 0$,
3) for some $(\gamma_+, \gamma_-) \in \{0,1\}^{\mathcal{E}(x)}$, $g(\gamma_+, \gamma_-) = 0$.

Some precisions for LCPs (1)

$$0 \leqslant F(x) = x \perp Mx + q = G(x) \geqslant 0$$
$$A = \mathcal{F}(x), \quad E = \mathcal{E}(x), \quad I = \mathcal{G}(x)$$

$$\mathcal{C}_{\text{ND}} \begin{cases} M_{I,I} \text{ nonsingular and} \\ M_{E,E} - M_{E,I} M_{I,I}^{-1} M_{I,E} \in \boldsymbol{ND} \end{cases}$$

Weighted framework for LCPs with $\mathcal{C}_{\text{ND}}$

The following properties are equivalent:

1) x is a solution,
2) for any (γ_+, γ_-), $g(\gamma_+, \gamma_-) = 0$,
3) for some $(\gamma_+, \gamma_-) \in \{0, 1\}^{\mathcal{E}(x)}$, $g(\gamma_+, \gamma_-) = 0$.

Some precisions for LCPs (2)

$$\mathcal{C}_{\mathrm{P}} \begin{cases} M_{I,I} \text{ nonsingular and} \\ M_{E,E} - M_{E,I}M_{I,I}^{-1}M_{I,E} \in \boldsymbol{P} \end{cases}$$

Weighted framework for LCPs with $\mathcal{C}_{\mathrm{P}}$

The following properties are equivalent:

1) x is a solution,
2) for any (γ_+, γ_-), $g(\gamma_+, \gamma_-) = 0$,
3) for some (γ_+, γ_-), $g(\gamma_+, \gamma_-) = 0$.

- For M a $\boldsymbol{P}$-matrix, $\mathcal{C}_{\mathrm{ND}}$ and $\mathcal{C}_{\mathrm{P}}$ ✓, solution ⇔ stationary,
- conditions with only one (γ_+, γ_-),
- a priori, still need to $\gamma_-(\gamma_+)$,
- outside of solutions $\neg$ 1), any γ_+ suffices $\neg$ 2)

Some precisions for LCPs (2)

$$\mathcal{C}_{\mathrm{P}} \begin{cases} M_{I,I} \text{ nonsingular and} \\ M_{E,E} - M_{E,I} M_{I,I}^{-1} M_{I,E} \in \boldsymbol{P} \end{cases}$$

Weighted framework for LCPs with $\mathcal{C}_{\mathrm{P}}$

The following properties are equivalent:

1) x is a solution,
2) for any (γ_+, γ_-), $g(\gamma_+, \gamma_-) = 0$,
3) for some (γ_+, γ_-), $g(\gamma_+, \gamma_-) = 0$.

- For M a $\boldsymbol{P}$-matrix, $\mathcal{C}_{\mathrm{ND}}$ and $\mathcal{C}_{\mathrm{P}}$ ✓, solution $\Leftrightarrow$ stationary,
- conditions with only one (γ_+, γ_-),
- a priori, still need to $\gamma_-(\gamma_+)$,
- outside of solutions $\neg$ 1), any γ_+ suffices $\neg$ 2)

Algorithm properties

Sufficient decrease

1) $d_k(\lambda)/||d_k(\lambda)|| \underset{\lambda \to +\infty}{\to} -S_k^{-1}g_k/||S_k^{-1}g_k||$
2) for λ large enough, a descent formula holds

Convergence

Let (x_k, λ_k, S_k) be a sequence generated by algorithm 1.

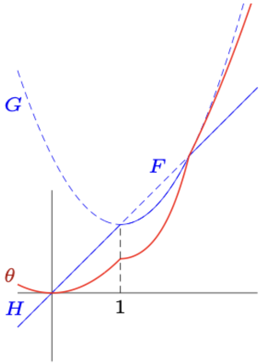
1) The sequence $(\theta(x_k))_k$ decreases thus converges.
2) If $(F'(x_k), G'(x_k), \lambda_k S_k)_{k \in \mathcal{K}}$ for a subsequence $\mathcal{K}$ is bounded, then $g_k \to 0$.

In particular, "good behavior" of algorithm is assumed: $\lambda_k \nrightarrow +\infty$.

"Concave kinks" – difficult (bad) limit points.

Consider a simple example with
$n = 1$, $F(x) = x$,
$G(x) = 1 + (x - 1)^2$, $x_0 = 3/2$.

For $x \in (1, 2)$, $F(x) \neq G(x)$.

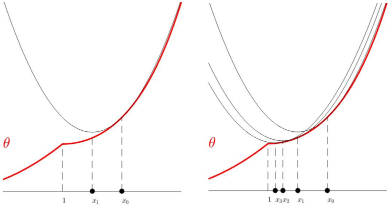


Counter-example

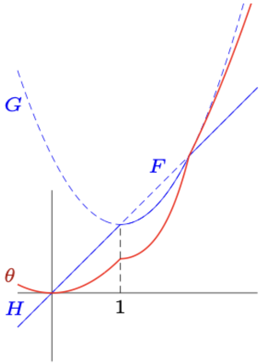
"Concave kinks" – difficult (bad) limit points.

Consider a simple example with
$n = 1$, $F(x) = x$,
$G(x) = 1 + (x-1)^2$, $x_0 = 3/2$.

For $x \in (1,2)$, $F(x) \neq G(x)$.



First iterates, convergence to $x = 1$. The black curves are the quadratic models ψ_x.



Counter-example