A Robust Linearization Method for Complementarity Problems

A Detour Through Hyperplane Arrangements

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

July, 16th 2025

Outline

1. Motivation
   - Problem presentation
   - One nonsmooth method
2. Hyperplane Arrangements
   - Computation of $\partial_B H(x)$
   - Some improvements
   - A new approach by duality
   - Some results
3. LM-PNM
   - Presentation of NM and PNM
   - Least-squares over regularity
   - Technical choice of the weights
   - Algorithmical considerations

Outline

Complementarity Problems (CPs)

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

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

Mostly, $F, G : \mathbb{R}^n \mapsto \mathbb{R}^n$ smooth. Often:

- G is the identity ($G(x) = x$), $0 \leqslant F(x) \perp x \geqslant 0$.
- G identity, F affine (LCPs, Cottle Pang Stone [CPS09])

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

Complementarity Problems (CPs)

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

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

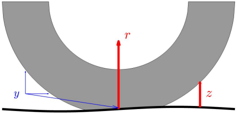
Mostly, $F, G : \mathbb{R}^n \mapsto \mathbb{R}^n$ smooth. Often:

- G is the identity ($G(x) = x$), $0 \leqslant F(x) \perp x \geqslant 0$.
- G identity, F affine (LCPs, Cottle Pang Stone [CPS09])

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

Relevance of CPs

Phenomena in competition, threshold effects [HP90; FP97]

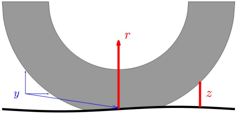


r: reaction, z: height

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

Relevance of CPs

Phenomena in competition, threshold effects [HP90; FP97]



r: reaction, z: height

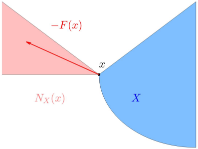
$$\forall \text{ point } y, \begin{cases} r(y) \geqslant 0, \\ z(y) \geqslant 0, \\ r(y)z(y) = 0. \end{cases}$$

constrained optimization

$$\min \ f(x), \quad g(x) \leqslant 0$$

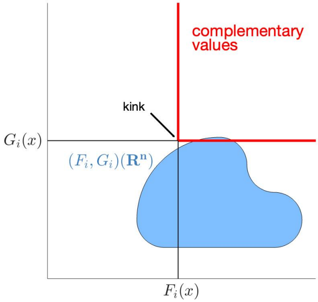
$$\text{KKT} \begin{cases} \nabla f(x) + \nabla g(x)\lambda = 0 \\ 0 \leqslant \lambda \perp (-g(x)) \geqslant 0 \end{cases}$$

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].

Possible reformulation techniques

C-function:

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

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

- $\varphi_{\mathrm{FB}}(a,b) := \sqrt{a^2+b^2} - a - b$ (Fischer [Fis92], [DFK00])
- $\varphi_{\min}(a,b) := \min(a,b)$, $H := H_{\min}$ (Pang [Pan90; Pan91], Qi [Qi93])
- many more ([Gal12; KYF97; FJ00; Alc+20])
- smoothing of the problem [Had09; CNQ00; Vu+21]

Possible reformulation techniques

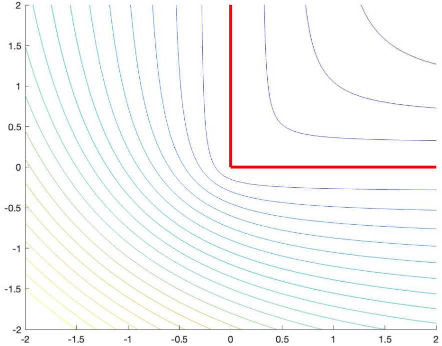
C-function:

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

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

- $\varphi_{\mathrm{FB}}(a,b) := \sqrt{a^2+b^2} - a - b$ (Fischer [Fis92], [DFK00])
- $\varphi_{\min}(a,b) := \min(a,b)$, $H := H_{\min}$ (Pang [Pan90; Pan91], Qi [Qi93])
- many more ([Gal12; KYF97; FJ00; Alc+20])
- smoothing of the problem [Had09; CNQ00; Vu+21]

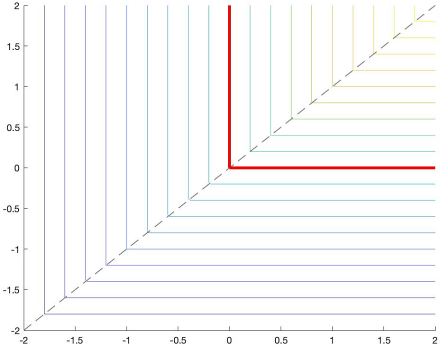
Illustration



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



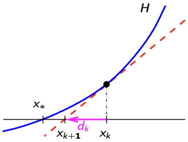
Level sets of the C-function φ_{FB}, nondifferentiable only at the origin.

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

Outline

1. Motivation
   - Problem presentation
   - One nonsmooth method
2. Hyperplane Arrangements
   - Computation of $\partial_B H(x)$
   - Some improvements
   - A new approach by duality
   - Some results
3. LM-PNM
   - Presentation of NM and PNM
   - Least-squares over regularity
   - Technical choice of the weights
   - Algorithmical considerations

Nonlinear (nonsmooth) equations – Newton's method



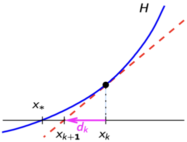
Smooth case

Iteration

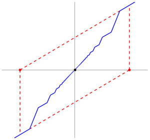
$$x_{k+1} = x_k - H'(x_k)^{-1}H(x_k)$$
$$(= x_k + d_k)$$

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

Nonlinear (nonsmooth) equations – Newton's method



Smooth case



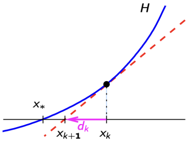
Can cycle if H nonsmooth [Kum88]

Iteration

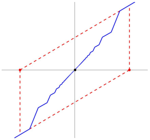
$$x_{k+1} = x_k - H'(x_k)^{-1} H(x_k)$$
$$(= x_k + d_k)$$

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

Nonlinear (nonsmooth) equations – Newton's method



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 $J_k \in \partial_{B|C} H(x_k)$

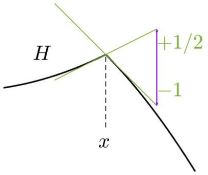
x_0 near x^*, H semismooth,
all $J \in \partial_{B|C} H(x^*)$ nonsingular
$\Rightarrow$ convergence

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

Why the minimum?

FB vs minimum

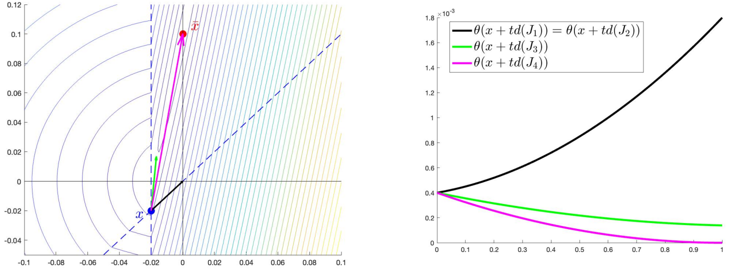
H_{FB}	$H = H_{\min}$
$\partial_B H_{\mathrm{FB}}(x^*)$ is a continuum	$\partial_B H(x^*)$ is a finite set
φ^2_{FB} is smooth (SC^1)	finite convergence if F, G affine
more studied	**less explored**

Globalization

In theory, with a good x_0, any $J \in \partial_B H(x)$ ✓. In practice?
Local $H(x) = 0 \to$ global $\min \theta := ||H(x)||^2/2$.

Globalization

In theory, with a good x_0, any $J \in \partial_B H(x)$ ✓. In practice?
Local $H(x) = 0 \to$ global $\min \theta := ||H(x)||^2/2$.



Left: level sets of θ. Globally, some J decrease θ, some increase it.

Contributions – two main topics

- compute $\partial_B H(x)$ in the affine case:
  - relevant for LCPs and linearizations,
  - link with centered hyperplane arrangements,
  - extension to general arrangements,
  - algorithmic approach;
- globalize $H(x) = 0$ via the minimization of θ:
  - $J \in \partial_B H(x)$ and directions,
  - algorithmic aspects (ongoing work).

Contributions – two main topics

- compute $\partial_B H(x)$ in the affine case:
  - relevant for LCPs and linearizations,
  - link with centered hyperplane arrangements,
  - extension to general arrangements,
  - algorithmic approach;
- globalize $H(x) = 0$ via the minimization of θ:
  - $J \in \partial_B H(x)$ and directions,
  - algorithmic aspects (ongoing work).

Outline

Outline

1. Motivation
   - Problem presentation
   - One nonsmooth method
2. Hyperplane Arrangements
   - Computation of $\partial_B H(x)$
   - Some improvements
   - A new approach by duality
   - Some results
3. LM-PNM
   - Presentation of NM and PNM
   - Least-squares over regularity
   - Technical choice of the weights
   - Algorithmical considerations

Setting

Initial question and objective

$$0 \leqslant Ax + a \perp Bx + b \geqslant 0 \quad \Longleftrightarrow$$
$$H(x) = \min(Ax + a, Bx + b) = 0.$$
$$\partial_B H(x) = \{J \in \mathbb{R}^{n\times n} : \exists\ (x_k)_k \to x, x_k \in \mathcal{D}_H, H'(x_k) \to J\}.$$

Relation between $\partial_B H$ and hyperplane arrangements: first part of [DGP25a], published in Math. Prog. Comp. (Jan. 2025).

Setting

Initial question and objective

$$0 \leqslant Ax + a \perp Bx + b \geqslant 0 \quad \iff$$
$$H(x) = \min(Ax + a, Bx + b) = 0.$$
$$\partial_B H(x) = \{J \in \mathbb{R}^{n\times n} : \exists\ (x_k)_k \to x, x_k \in \mathcal{D}_H, H'(x_k) \to J\}.$$

Relation between $\partial_B H$ and hyperplane arrangements: first part of [DGP25a], published in Math. Prog. Comp. (Jan. 2025).

Computing the B-differential

Summary: (see also [XC11])

- H piecewise affine, "derivative" piecewise constant;
- $H_i(x) \in \{(Ax+a)_i, (By+b)_i\} = \{A_{i,:}x + a_i, B_{i,:}x + b_i\}$, $J_{i,:} \in \{A_{i,:}, B_{i,:}\}\ \forall\ J$;
- 2 possibilities for each i: combinatorial nature;
- H_i nondifferentiable at x iff $A_{i,:}x + a_i = B_{i,:}x + b_i$, $A_{i,:} \neq B_{i,:}$:= **hyperplane** containing x.

Computing the B-differential

Summary: (see also [XC11])

- H piecewise affine, "derivative" piecewise constant;
- $H_i(x) \in \{(Ax+a)_i, (By+b)_i\} = \{A_{i,:}x + a_i, B_{i,:}x + b_i\}$, $J_{i,:} \in \{A_{i,:}, B_{i,:}\}\ \forall\ J$;
- 2 possibilities for each i: combinatorial nature;
- H_i nondifferentiable at x iff $A_{i,:}x + a_i = B_{i,:}x + b_i$, $A_{i,:} \neq B_{i,:}$ $:=$ **hyperplane** containing x.

Computing the B-differential

Summary: (see also [XC11])

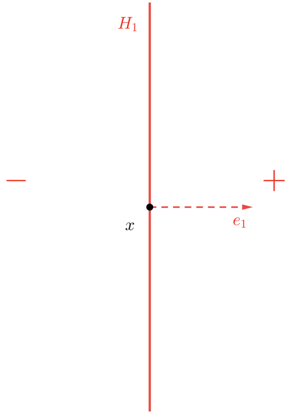
- H piecewise affine, "derivative" piecewise constant;
- $H_i(x) \in \{(Ax+a)_i, (By+b)_i\} = \{A_{i,:}x + a_i, B_{i,:}x + b_i\}$,
  $J_{i,:} \in \{A_{i,:}, B_{i,:}\}\ \forall\ J$;
- 2 possibilities for each i: combinatorial nature;
- H_i nondifferentiable at x iff $A_{i,:}x + a_i = B_{i,:}x + b_i$, $A_{i,:} \neq B_{i,:}$
  $:=$ **hyperplane** containing x.

Computing the B-differential

Summary: (see also [XC11])

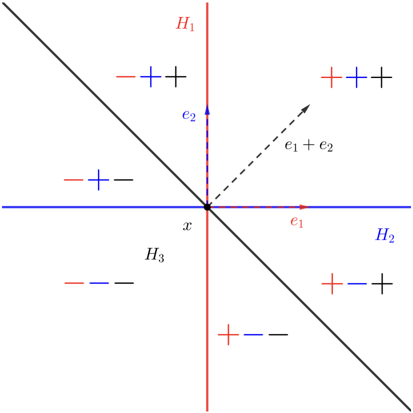
- H piecewise affine, "derivative" piecewise constant;
- $H_i(x) \in \{(Ax+a)_i, (By+b)_i\} = \{A_{i,:}x + a_i, B_{i,:}x + b_i\}$, $J_{i,:} \in \{A_{i,:}, B_{i,:}\}$ $\forall$ J;
- 2 possibilities for each i: combinatorial nature;
- H_i nondifferentiable at x iff $A_{i,:}x + a_i = B_{i,:}x + b_i$, $A_{i,:} \neq B_{i,:}$:= **hyperplane** containing x.

B-differential and hyperplanes



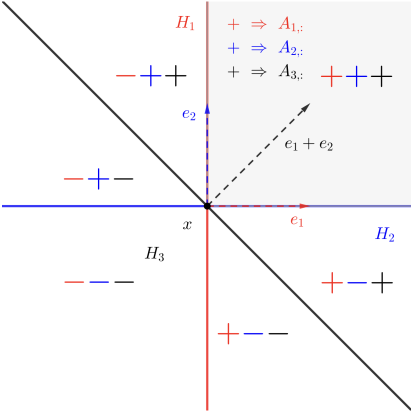
$A_{i,:} \leftrightarrow +$, $B_{i,:} \leftrightarrow -$. In the areas, the J is constant: one needs to determine the nonempty areas.

B-differential and hyperplanes



$A_{i,:} \leftrightarrow +$, $B_{i,:} \leftrightarrow -$. In the areas, the J is constant: one needs to determine the nonempty areas.

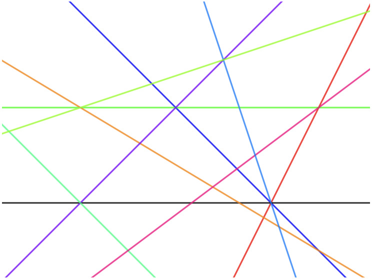
B-differential and hyperplanes



$A_{i,:} \leftrightarrow +$, $B_{i,:} \leftrightarrow -$. In the areas, the J is constant: one needs to determine the nonempty areas.

General case, no intersection

How to determine the chambers?



- Studied for centuries [Ste26; Rob87; Sch50]
- Applications / ... [Zas75; DP22; Sle00; BEK23; Zie07; Oxl11] (Zaslavsky, Ziegler, Oxley...)

Notation

$$H_i := \{x \in \mathbb{R}^n : \overbrace{v_i^\top x = \tau_i}^{v_i \in \mathbb{R}^n, \tau_i \in \mathbb{R}}\}, \quad V = [v_1 \ \dots \ v_p], \quad \tau = [\tau_1; \dots; \tau_p]$$

Notation

$$H_i := \{x \in \mathbb{R}^n : \overbrace{v_i^\mathsf{T} x = \tau_i}^{v_i \in \mathbb{R}^n, \tau_i \in \mathbb{R}}\}, \quad V = [v_1 \ \dots \ v_p], \quad \tau = [\tau_1; \dots; \tau_p]$$

$$H_i^+ := \{x \in \mathbb{R}^n : \underbrace{v_i^\mathsf{T} x > \tau_i}_{+(v_i^\mathsf{T} x - \tau_i) > 0}\}, \quad H_i^- := \{x \in \mathbb{R}^n : \underbrace{v_i^\mathsf{T} x < \tau_i}_{-(v_i^\mathsf{T} x - \tau_i) > 0}\}$$

Which intersections $\bigcap_{i=1}^p (H_i^+ \text{ or } H_i^-)$ are nonempty?

Notation

$$H_i := \{x \in \mathbb{R}^n : \overbrace{v_i^\mathsf{T} x = \tau_i}^{v_i \in \mathbb{R}^n, \tau_i \in \mathbb{R}}\}, \quad V = [v_1 \ \dots \ v_p], \quad \tau = [\tau_1; \dots; \tau_p]$$

$$H_i^+ := \{x \in \mathbb{R}^n : \underbrace{v_i^\mathsf{T} x > \tau_i}_{+(v_i^\mathsf{T} x - \tau_i) > 0}\}, \quad H_i^- := \{x \in \mathbb{R}^n : \underbrace{v_i^\mathsf{T} x < \tau_i}_{-(v_i^\mathsf{T} x - \tau_i) > 0}\}$$

Which intersections $\bigcap_{i=1}^p (H_i^+ \text{ or } H_i^-)$ are nonempty?

Geometric to analytic: sign vectors

$$\text{find } \mathcal{S}(V,\tau) := \{s = (s_1, \dots, s_p) \in \{\pm 1\}^p,$$
$$\text{s.t. } \exists\, x^s \in \mathbb{R}^n, \quad \forall\, i \in [1:p], \quad s_i(v_i^\mathsf{T} x^s - \tau_i) > 0\}$$

Formula (bound from 19th century!) give $|\mathcal{S}(V,\tau)|$, not $\mathcal{S}(V,\tau)$.
Naively: 2^p strict affine feasibility systems (LOPs).

Towards algorithms

Use of linear optimization

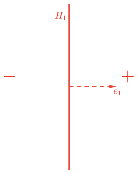
Additional variable α: the > 0 become $\geqslant 0$

$$\alpha_s^* = \min_{x \in \mathbb{R}^n, \alpha \geqslant -1} \quad \alpha \quad \text{s.t.} \quad s_i(v_i^\mathsf{T} x - \tau_i) + \alpha \geqslant 0, \quad \forall\, i \in [1:p] \tag{3}$$

$$s \in \mathcal{S}(V, \tau) \quad \Longleftrightarrow \quad \alpha_s^* < 0.$$

Main algo: Rada & Černý [RČ18]; improvements in [DGP25b].

Illustration of the RČ algorithm



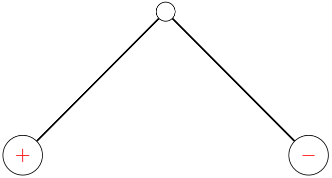
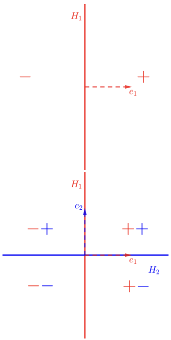


Illustration of the RČ algorithm



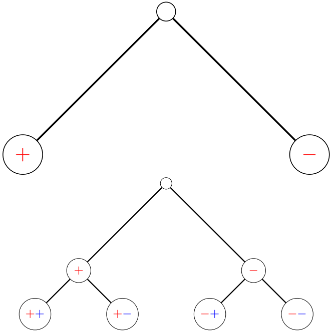
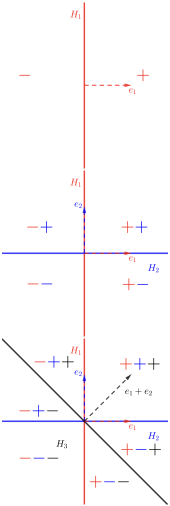
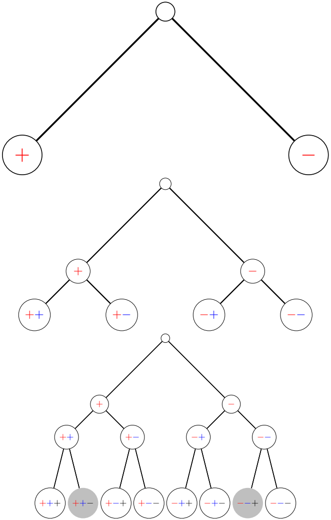


Illustration of the RČ algorithm





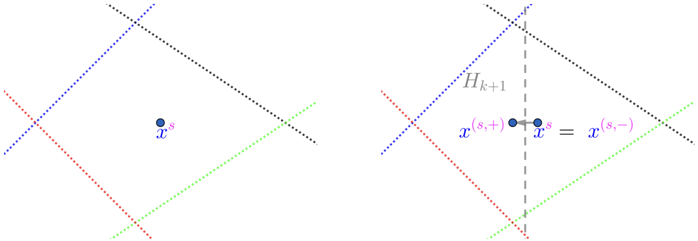
Summary of the algorithm

- at each $s \in \{\pm 1\}^k$, x^s verifies $(s, +)$ or $(s, -)$;
- **one LOP per node** for the other, (3) with $k+1$ signs;
- good theoretical properties ($>$ previous ones [AF92; Sle98], [BN82; EOS86] more general).

Main cost: LOPs – can we solve less?

Outline

Closeness (heuristic "B")

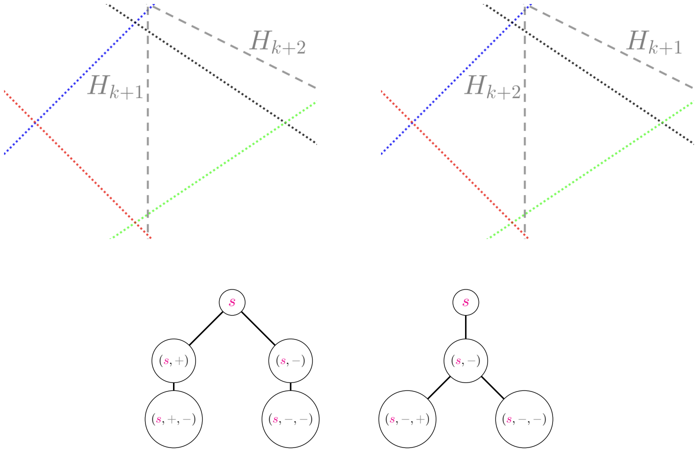


Left: level k. Right: shift of x^s when $x^s \underset{\sim}{\in} H_{k+1}$.

Details

For $s \in \{\pm 1\}^k$ with x^s, if $x^s \underset{\sim}{\in} H_{k+1} \Leftrightarrow v_{k+1}^{\mathsf{T}} x^s - \tau_{k+1} \simeq 0$, $(s, +1)$ and $(s, -1)$ in level $k+1$ without LOP.

Sequencing (heuristic “C”) – which order to choose?



Changes inner levels – level p is always $\mathcal{S}(V,\tau)$.

Outline

1. Motivation
   - Problem presentation
   - One nonsmooth method
2. Hyperplane Arrangements
   - Computation of $\partial_B H(x)$
   - Some improvements
   - A new approach by duality
   - Some results
3. LM-PNM
   - Presentation of NM and PNM
   - Least-squares over regularity
   - Technical choice of the weights
   - Algorithmical considerations

Dual approach: avoid LOPs

- $s \in \{\pm 1\}^p$ is incompatible if $s \notin \mathcal{S}(V, \tau)$ ($s \in \mathcal{S}(V, \tau)^c$):

$$\nexists\, x \in \mathbb{R}^n : \qquad s \bullet (V^\mathsf{T} x - \tau) > 0, \\ \Leftrightarrow s \bullet V^\mathsf{T} x > s \bullet \tau.$$

- With all incompatible s_I: no need for LO in the tree.
- For $s \in \{\pm 1\}^p$ and $I \subseteq [1:p]$, s_I incompatible $\Rightarrow$ s is incompatible (more inequalities).
- For $s \in \{\pm 1\}^p$ incompatible, $\exists$ minimal incompatible s_I.
- Find all the smallest I with s_I incompatible.

Dual approach: avoid LOPs

- $s \in \{\pm 1\}^p$ is incompatible if $s \notin \mathcal{S}(V, \tau)$ ($s \in \mathcal{S}(V, \tau)^c$):

$$\nexists\, x \in \mathbb{R}^n : \qquad s \bullet (V^\mathsf{T} x - \tau) > 0,$$
$$\Leftrightarrow s \bullet V^\mathsf{T} x > s \bullet \tau.$$

- With all incompatible s_I: no need for LO in the tree.
- For $s \in \{\pm 1\}^p$ and $I \subseteq [1:p]$, s_I incompatible $\Rightarrow$ s is incompatible (more inequalities).
- For $s \in \{\pm 1\}^p$ incompatible, $\exists$ minimal incompatible s_I.
- Find all the smallest I with s_I incompatible.

Dual approach: avoid LOPs

- $s \in \{\pm 1\}^p$ is incompatible if $s \notin \mathcal{S}(V, \tau)$ ($s \in \mathcal{S}(V, \tau)^c$):

$$\begin{aligned} \nexists\, x \in \mathbb{R}^n : \quad & s \bullet (V^\top x - \tau) > 0, \\ & \Leftrightarrow s \bullet V^\top x > s \bullet \tau. \end{aligned}$$

- With all incompatible s_I: no need for LO in the tree.
- For $s \in \{\pm 1\}^p$ and $I \subseteq [1:p]$, s_I incompatible $\Rightarrow$ s is incompatible (more inequalities).
- For $s \in \{\pm 1\}^p$ incompatible, $\exists$ minimal incompatible s_I.
- Find all the smallest I with s_I incompatible.

Dual approach: avoid LOPs

- $s \in \{\pm 1\}^p$ is incompatible if $s \notin \mathcal{S}(V, \tau)$ ($s \in \mathcal{S}(V, \tau)^c$):

$$\begin{aligned} \nexists\, x \in \mathbb{R}^n : \qquad & s \bullet (V^\top x - \tau) > 0, \\ & \Leftrightarrow s \bullet V^\top x > s \bullet \tau. \end{aligned}$$

- With all incompatible s_I: no need for LO in the tree.
- For $s \in \{\pm 1\}^p$ and $I \subseteq [1:p]$, s_I incompatible $\Rightarrow$ s is incompatible (more inequalities).
- For $s \in \{\pm 1\}^p$ incompatible, $\exists$ minimal incompatible s_I.
- Find all the smallest I with s_I incompatible.

Dual approach: avoid LOPs

- $s \in \{\pm 1\}^p$ is incompatible if $s \notin \mathcal{S}(V, \tau)$ ($s \in \mathcal{S}(V, \tau)^c$):

$$\begin{aligned} \nexists\, x \in \mathbb{R}^n : \qquad & s \cdot (V^\mathsf{T} x - \tau) > 0, \\ & \Leftrightarrow s \cdot V^\mathsf{T} x > s \cdot \tau. \end{aligned}$$

- With all incompatible s_I: no need for LO in the tree.
- For $s \in \{\pm 1\}^p$ and $I \subseteq [1:p]$, s_I incompatible $\Rightarrow$ s is incompatible (more inequalities).
- For $s \in \{\pm 1\}^p$ incompatible, $\exists$ minimal incompatible s_I.
- Find all the smallest I with s_I incompatible.

Circuits and stem vectors – 1

A convex analysis tool: duality via Motzkin's alternative [Mot36]

$$\nexists\, x : Mx > m \iff \exists\, \alpha \in \mathbb{R}^p_+ \setminus \{0\} : M^\mathsf{T}\alpha = 0,\, m^\mathsf{T}\alpha \geqslant 0.$$

$$s_I \text{ incompatible} \iff \nexists\, x \in \mathbb{R}^n : s_I \bullet V_{:,I}^\mathsf{T} x > s_I \bullet \tau_I$$

$$\iff \exists\, \alpha \in \mathbb{R}^I_+ \setminus \{0\} : V_{:,I}(\underbrace{s_I \bullet \alpha}_{=\eta\in\mathbb{R}^I}) = 0,\ \tau_I^\mathsf{T}(\underbrace{s_I \bullet \alpha}_{=\eta\in\mathbb{R}^I}) \geqslant 0.$$

The η is in $\mathcal{N}(V_{:,I}) \setminus \{0\}$, and oriented: $\tau_I^\mathsf{T}\eta \geqslant 0$ (otherwise: $-\eta$).

Circuits and stem vectors – 1

A convex analysis tool: duality via Motzkin's alternative [Mot36]

$$\nexists\, x : Mx > m \quad \Longleftrightarrow \quad \exists\, \alpha \in \mathbb{R}^p_+ \setminus \{0\} : M^\mathsf{T}\alpha = 0, m^\mathsf{T}\alpha \geqslant 0.$$

$$s_I \text{ incompatible } \Longleftrightarrow \nexists\, x \in \mathbb{R}^n : s_I \bullet V_{:,I}^\mathsf{T} x > s_I \bullet \tau_I$$

$$\Longleftrightarrow \exists\, \alpha \in \mathbb{R}^I_+ \setminus \{0\} : V_{:,I}(\underbrace{s_I \bullet \alpha}_{=\eta\in\mathbb{R}^I}) = 0,\ \tau_I^\mathsf{T}(\underbrace{s_I \bullet \alpha}_{=\eta\in\mathbb{R}^I}) \geqslant 0.$$

The η is in $\mathcal{N}(V_{:,I}) \setminus \{0\}$, and oriented: $\tau_I^\mathsf{T}\eta \geqslant 0$ (otherwise: $-\eta$).

Circuits and stem vectors – 1

A convex analysis tool: duality via Motzkin's alternative [Mot36]

$$\nexists\, x : Mx > m \quad \Longleftrightarrow \quad \exists\, \alpha \in \mathbb{R}^p_+ \setminus \{0\} : M^\mathsf{T}\alpha = 0, m^\mathsf{T}\alpha \geqslant 0.$$

$$s_I \text{ incompatible } \Longleftrightarrow \nexists\, x \in \mathbb{R}^n : s_I \bullet V_{:,I}^\mathsf{T} x > s_I \bullet \tau_I$$

$$\Longleftrightarrow \exists\, \alpha \in \mathbb{R}^I_+ \setminus \{0\} : V_{:,I}(\underbrace{s_I \bullet \alpha}_{=\eta \in \mathbb{R}^I}) = 0,\ \tau_I^\mathsf{T}(\underbrace{s_I \bullet \alpha}_{=\eta \in \mathbb{R}^I}) \geqslant 0.$$

The η is in $\mathcal{N}(V_{:,I}) \setminus \{0\}$, and oriented: $\tau_I^\mathsf{T}\eta \geqslant 0$ (otherwise: $-\eta$).

Circuits and stem vectors – 2

$$s_I \text{ incompatible } \iff \exists\, \eta \in \mathbb{R}^I \setminus \{0\} : V_{:,I} \underbrace{\eta}_{s_I \bullet \alpha} = 0,\ \tau_I^\mathsf{T} \underbrace{\eta}_{s_I \bullet \alpha} \geqslant 0.$$

- Smallest I's, $\eta \in \mathcal{N}(V_{:,I}) \setminus \{0\} \Rightarrow$ *matroid circuits* of V [Oxl11]:

$$\mathcal{C}(V) := \{I \subseteq [1:p] : \underbrace{\operatorname{null}(V_{:,I})}_{\dim(\mathcal{N}(V_{:,I}))} = 1, \operatorname{null}(V_{:,I_0}) = 0\ \forall\ I_0 \subsetneq I\}$$

- *Stem vectors* $\mathfrak{S}(V,\tau) := \{\sigma \in \{\pm 1\}^I : I \in \mathcal{C}(V)$ and

$$\sigma = \operatorname{sgn}(\eta) \text{ for } \eta \in \mathcal{N}(V_{:,I}) \setminus \{0\} \text{ s.t. } \tau_I^\mathsf{T}\eta \geqslant 0\}.$$

Circuits and stem vectors – 2

$$s_I \text{ incompatible } \iff \exists\, \eta \in \mathbb{R}^I \setminus \{0\} : V_{:,I} \underbrace{\eta}_{s_I \bullet \alpha} = 0,\ \tau_I^{\mathsf{T}} \underbrace{\eta}_{s_I \bullet \alpha} \geqslant 0.$$

- Smallest I's, $\eta \in \mathcal{N}(V_{:,I}) \setminus \{0\} \Rightarrow$ *matroid circuits* of V [Oxl11]:

$$\mathcal{C}(V) := \{I \subseteq [1:p] : \underbrace{\operatorname{null}(V_{:,I})}_{\dim(\mathcal{N}(V_{:,I}))} = 1, \operatorname{null}(V_{:,I_0}) = 0\ \forall\ I_0 \subsetneq I\}$$

- *Stem vectors* $\mathfrak{S}(V,\tau) := \{\sigma \in \{\pm 1\}^I : I \in \mathcal{C}(V)$ and

$$\sigma = \operatorname{sgn}(\eta) \text{ for } \eta \in \mathcal{N}(V_{:,I}) \setminus \{0\} \text{ s.t. } \tau_I^{\mathsf{T}}\eta \geqslant 0\}.$$

Circuits and stem vectors – 2

$$s_I \text{ incompatible} \iff \exists\, \eta \in \mathbb{R}^I \setminus \{0\} : V_{:,I} \underbrace{\eta}_{s_I \bullet \alpha} = 0,\ \tau_I^\mathsf{T} \underbrace{\eta}_{s_I \bullet \alpha} \geqslant 0.$$

- Smallest I's, $\eta \in \mathcal{N}(V_{:,I}) \setminus \{0\} \Rightarrow$ *matroid circuits* of V [Oxl11]:

$$\mathcal{C}(V) := \{I \subseteq [1:p] : \underbrace{\mathrm{null}(V_{:,I})}_{\dim(\mathcal{N}(V_{:,I}))} = 1, \mathrm{null}(V_{:,I_0}) = 0\ \forall\ I_0 \subsetneq I\}$$

- *Stem vectors* $\mathfrak{S}(V,\tau) := \{\sigma \in \{\pm 1\}^I : I \in \mathcal{C}(V)$ and

$$\sigma = \mathrm{sgn}(\eta) \text{ for } \eta \in \mathcal{N}(V_{:,I}) \setminus \{0\} \text{ s.t. } \tau_I^\mathsf{T}\eta \geqslant 0\}.$$

Circuits and stem vectors – 3

Covering test

$$s \in \mathcal{S}(V,\tau)^c \quad \Longleftrightarrow \quad s_I \in \mathfrak{S}(V,\tau) \text{ for some } I \subseteq [1:p].$$

$$(\mathrm{sgn}(\eta) = \mathrm{sgn}(s_I \cdot \alpha) = \mathrm{sgn}(s_I) = s_I)$$

Dual algorithm: tree with covering tests

- Compute $\mathfrak{S}(V,\tau)$ (via $\mathcal{C}(V)$).
- Test if $(s, +1)$ covers a stem vector.
- If yes, stop; if no, recursion on $(s, +1)$.
- Same for $(s, -1)$.

Circuits and stem vectors – 3

Covering test

$$s \in \mathcal{S}(V,\tau)^c \quad \Longleftrightarrow \quad s_I \in \mathfrak{S}(V,\tau) \text{ for some } I \subseteq [1:p].$$

$$(\mathrm{sgn}(\eta) = \mathrm{sgn}(s_I \cdot \alpha) = \mathrm{sgn}(s_I) = s_I)$$

Dual algorithm: tree with covering tests

- Compute $\mathfrak{S}(V,\tau)$ (via $\mathcal{C}(V)$).
- Test if $(s,+1)$ covers a stem vector.
- If yes, stop; if no, recursion on $(s,+1)$.
- Same for $(s,-1)$.

Circuits and stem vectors – 3

Covering test

$$s \in \mathcal{S}(V,\tau)^c \quad \Longleftrightarrow \quad s_I \in \mathfrak{S}(V,\tau) \text{ for some } I \subseteq [1:p].$$

$$(\mathrm{sgn}(\eta) = \mathrm{sgn}(s_I \cdot \alpha) = \mathrm{sgn}(s_I) = s_I)$$

Dual algorithm: tree with covering tests

- Compute $\mathfrak{S}(V,\tau)$ (via $\mathcal{C}(V)$).
- Test if $(s,+1)$ covers a stem vector.
- If yes, stop; if no, recursion on $(s,+1)$.
- Same for $(s,-1)$.

Circuits and stem vectors – 3

Covering test

$$s \in \mathcal{S}(V,\tau)^c \quad \Longleftrightarrow \quad s_I \in \mathfrak{S}(V,\tau) \text{ for some } I \subseteq [1:p].$$

$$(\mathrm{sgn}(\eta) = \mathrm{sgn}(s_I \cdot \alpha) = \mathrm{sgn}(s_I) = s_I)$$

Dual algorithm: tree with covering tests

- Compute $\mathfrak{S}(V,\tau)$ (via $\mathcal{C}(V)$).
- Test if $(s,+1)$ covers a stem vector.
- If yes, stop; if no, recursion on $(s,+1)$.
- Same for $(s,-1)$.

Circuits and stem vectors – 3

Covering test

$$s \in \mathcal{S}(V,\tau)^c \quad \iff \quad s_I \in \mathfrak{S}(V,\tau) \text{ for some } I \subseteq [1:p].$$

$$(\mathrm{sgn}(\eta) = \mathrm{sgn}(s_I \cdot \alpha) = \mathrm{sgn}(s_I) = s_I)$$

Dual algorithm: tree with covering tests

- Compute $\mathfrak{S}(V,\tau)$ (via $\mathcal{C}(V)$).
- Test if $(s,+1)$ covers a stem vector.
- If yes, stop; if no, recursion on $(s,+1)$.
- Same for $(s,-1)$.

Comparison

each inner node	Primal	Dual
verification concretely	1 LOP: low-dimension	1-2 covering test(s): array operations

Computing $\mathfrak{S}(V, \tau)$ is a combinatorial problem.
If $|\mathfrak{S}(V, \tau)|$ large, long computation and covering tests.

Intermediate: Primal-Dual, only *some* duality

- launch the primal tree;
- $(s, \pm)$ incompatible $\overset{\text{Motzkin}}{\Longrightarrow}$ stem vector;
- use the *acquired* stem vectors (and still LOPs).

Comparison

each inner node	Primal	Dual
verification concretely	1 LOP: low-dimension	1-2 covering test(s): array operations

Computing $\mathfrak{S}(V,\tau)$ is a combinatorial problem.
If $|\mathfrak{S}(V,\tau)|$ large, long computation and covering tests.

Intermediate: Primal-Dual, only *some* duality

- launch the primal tree;
- $(s,\pm)$ incompatible $\overset{\text{Motzkin}}{\Longrightarrow}$ stem vector;
- use the *acquired* stem vectors (and still LOPs).

Comparison

each inner node	Primal	Dual
verification concretely	1 LOP: low-dimension	1-2 covering test(s): array operations

Computing $\mathfrak{S}(V,\tau)$ is a combinatorial problem.
If $|\mathfrak{S}(V,\tau)|$ large, long computation and covering tests.

Intermediate: Primal-Dual, only *some* duality

- launch the primal tree;
- $(s,\pm)$ incompatible $\overset{\text{Motzkin}}{\Longrightarrow}$ stem vector;
- use the *acquired* stem vectors (and still LOPs).

General improvement: "compaction"

Definition: $s \in \{\pm 1\}^p$ symmetric if $s, -s \in \mathcal{S}(V,\tau)$.
$\mathcal{S}(V,\tau)$ symmetric if $\mathcal{S}(V,\tau) = -\mathcal{S}(V,\tau)$.

General improvement: “compaction”

Definition: $s \in \{\pm 1\}^p$ symmetric if $s, -s \in \mathcal{S}(V, \tau)$.
$\mathcal{S}(V, \tau)$ symmetric if $\mathcal{S}(V, \tau) = -\mathcal{S}(V, \tau)$.

Principle

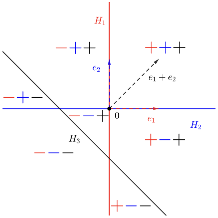
- $\mathcal{S}(V, \tau)$ (and tree) asymmetric, we can “symmetrize”.
- For all variants (RČ, P, D, PD).

General improvement: "compaction"

Definition: $s \in \{\pm 1\}^p$ symmetric if $s, -s \in \mathcal{S}(V,\tau)$.
$\mathcal{S}(V,\tau)$ symmetric if $\mathcal{S}(V,\tau) = -\mathcal{S}(V,\tau)$.

Principle

- $\mathcal{S}(V,\tau)$ (and tree) asymmetric, we can "symmetrize".
- For all variants (RČ, P, D, PD).

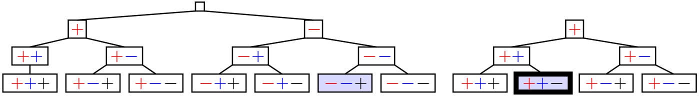


Asymmetric arrangement

$$\mathcal{S}(V,\tau) = \{(+++), (-++), (+-+), (--+), (---), (+--), (-+-)\}$$

except $(--+)$, rest symmetric

Compaction illustrated



Left: classic tree. Right: compact tree.

Blued nodes: asymmetric nodes, correction in the right tree. At the end, the other nodes are multipled by -1 to recover all nodes.

Outline

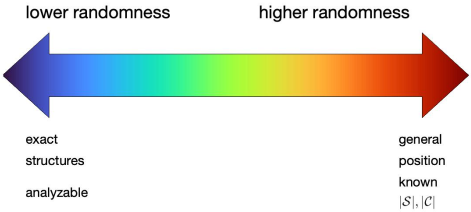
1. Motivation
   - Problem presentation
   - One nonsmooth method
2. Hyperplane Arrangements
   - Computation of $\partial_B H(x)$
   - Some improvements
   - A new approach by duality
   - Some results
3. LM-PNM
   - Presentation of NM and PNM
   - Least-squares over regularity
   - Technical choice of the weights
   - Algorithmical considerations

Algorithms and instances

- Basic: [RČ18] – “RČ” (Rada Černý).
- With heuristics – “P” (Primal).
- Without LOPs, just stem vectors – “D” (Dual).
- LOPs and some stem vectors – “PD” (Primal-Dual).
- Relevance of compaction (/C).

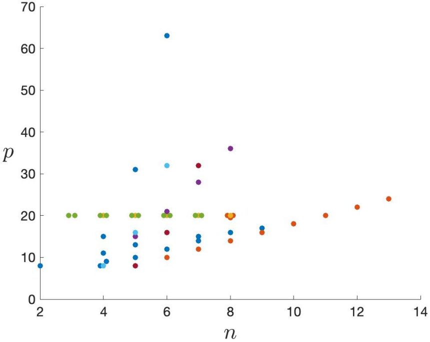
Algorithms and instances

- Basic: [RČ18] – "RČ" (Rada Černý).
- With heuristics – "P" (Primal).
- Without LOPs, just stem vectors – "D" (Dual).
- LOPs and some stem vectors – "PD" (Primal-Dual).
- Relevance of compaction (/C).



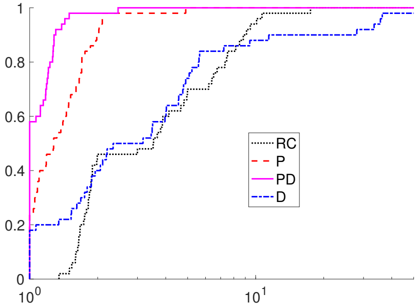
Trade-off between randomness and structure in the 50 tested instances.

Details on instances



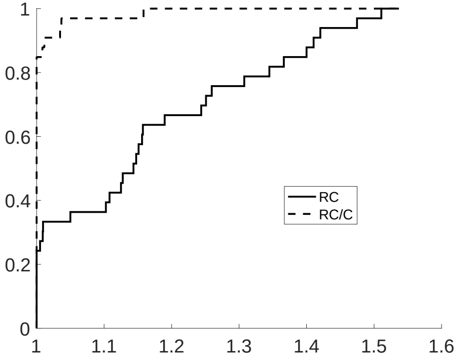
Pairs (n, p) for some linear and affine instances, grouped by colors. Instances up to 10^6 chambers/circuits (to run on a laptop). Example: $n = 7, p = 20$, up to 137980 chambers, 125970 stem vectors.

Comparison of the main variants



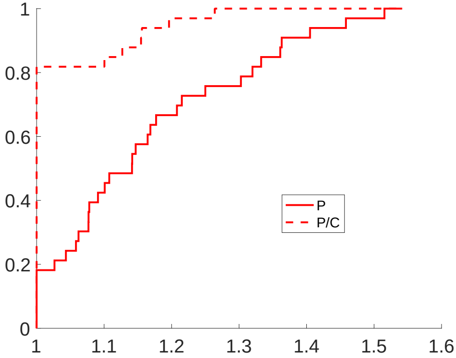
x-axis: relative efficiency (on time), y-axis: % of problems; above/left means being better. One has: primal-dual (PD) > primal (P) on some instances, both > Rada-Černý (RČ) and dual (D), which are quite close.

Variant vs compact variant



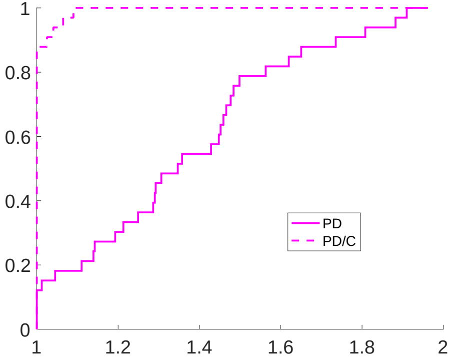
Compaction improves Rada-Černý RC, primal P and primal-dual PD (axis up to 2), but not really dual (D): less tests but more stems.

Variant vs compact variant



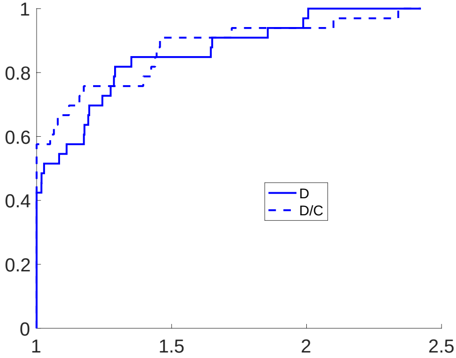
Compaction improves Rada-Černý RC, primal P and primal-dual PD (axis up to 2), but not really dual (D): less tests but more stems.

Variant vs compact variant



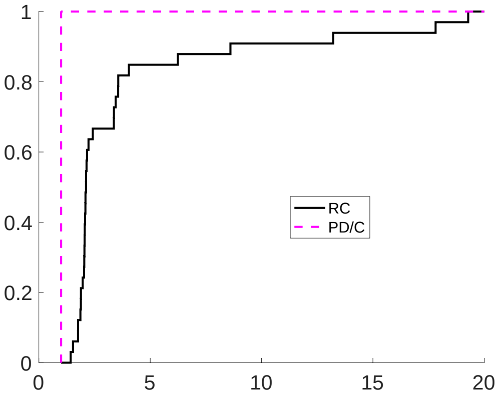
Compaction improves Rada-Černý RC, primal P and primal-dual PD (axis up to 2), but not really dual (D): less tests but more stems.

Variant vs compact variant



Compaction improves Rada-Černý RC, primal P and primal-dual PD (axis up to 2), but not really dual (D): less tests but more stems.

Initial vs best algorithm



Larger x-axis: average $\simeq 4$. Especially better on “structured” instances.

Possible future work

Code: data structures, parallelism; specialized techniques [DP22]?

Possible future work

Code: data structures, parallelism; specialized techniques [DP22]?

Taking symmetry into account (see [BEK23] and [Ram23]):

- either to consider only a part of the tree
- or to obtain / use circuits (stem vectors) much more faster.

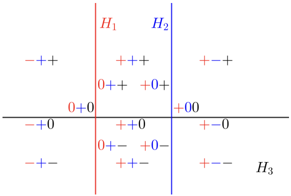
Possible future work

Code: data structures, parallelism; specialized techniques [DP22]?

Taking symmetry into account (see [BEK23] and [Ram23]):

- either to consider only a part of the tree
- or to obtain / use circuits (stem vectors) much more faster.

Also, $\{\pm 1\}^p \to \{-1, 0, +1\}^p$: intersections of halfspaces and/or hyperplanes.



Example with three lines.

Outline

1. Motivation
   - Problem presentation
   - One nonsmooth method
2. Hyperplane Arrangements
   - Computation of $\partial_B H(x)$
   - Some improvements
   - A new approach by duality
   - Some results
3. LM-PNM
   - Presentation of NM and PNM
   - Least-squares over regularity
   - Technical choice of the weights
   - Algorithmical considerations

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 ∂_C);
- requires all $J \in \partial_B^\times H(x^*)$ nonsingular.

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

Index sets

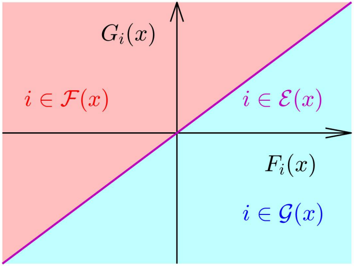
$$\text{if } F_i(x) < G_i(x), \quad H_i(x) = F_i(x), \quad \partial_B H_i(x) = \{F_i'(x)\}$$
$$\text{if } F_i(x) > G_i(x), \quad H_i(x) = G_i(x) \quad \partial_B H_i(x) = \{G_i'(x)\}$$
$$\text{if } F_i(x) = G_i(x), \quad = H_i(x) \quad \partial_B H_i(x) = \{F_i'(x), G_i'(x)\}$$

Index sets

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

Index sets

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



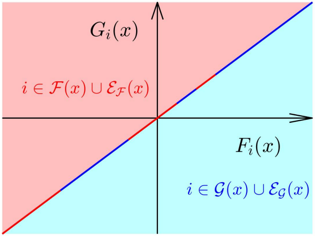
Representation of the index sets.

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?

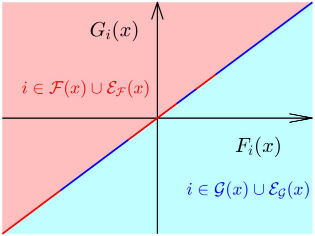
$\rightarrow$ An augmented system ensuring descent (θ decreases).

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.

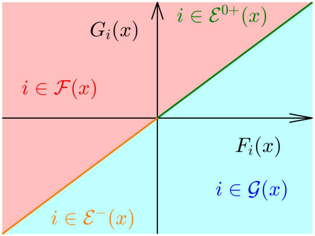


NM index sets.

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

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 $\mathcal{E}^{0+}(x)$:

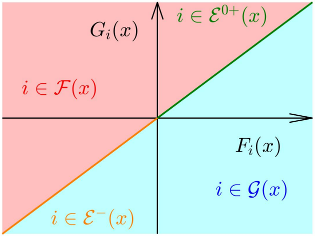


The “positive” and “negative” kinks.

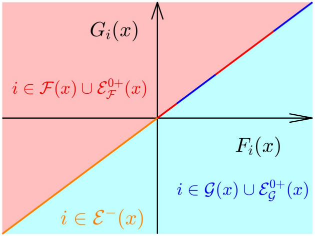
$\mathcal{E}^{0+}(x)$ violates complementarity, $\mathcal{E}^{-}(x)$ also violates $\geqslant 0$.
(See also [PG93; QS94; HPR92].)

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" and "negative" kinks.



Right: type of PNM partitioning.

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

PNM method (simplified)

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

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

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

$$\theta'(x; d) = \underbrace{-2\theta(x)}_{\text{smooth}} + \sum_{i \in \mathcal{E}(x)} \underbrace{H_i(\min(F_i + F_i' d, G_i + G_i' d))}_{\text{due to nonsmooth}} \leqslant -2\theta(x).$$

PNM method (simplified)

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

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

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

$$\theta'(x;d) = \underbrace{-2\theta(x)}_{\text{smooth}} + \sum_{i \in \mathcal{E}(x)} \underbrace{H_i(\min(F_i + F_i' d, G_i + G_i' d))}_{\text{due to nonsmooth}} \leqslant -2\theta(x).$$

PNM method (simplified)

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

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

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

$$\theta'(x; d) = \underbrace{-2\theta(x)}_{\text{smooth}} + \sum_{i \in \mathcal{E}(x)} \underbrace{H_i(\min(F_i + F_i' d, G_i + G_i' d))}_{\text{due to nonsmooth}} \leqslant -2\theta(x).$$

PNM regularity

The polyhedron has $n + |\mathcal{E}^-(x)|$ equations, n variables.
PNM stuck when it is empty.

PNM regularity condition to have a d

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

Hybrid version: if NM works ✓, otherwise PNM (seldom).
Excellent performance on random data / applications.

Now, we try to remove the regularity.

PNM regularity

The polyhedron has $n + |\mathcal{E}^-(x)|$ equations, n variables.
PNM stuck when it is empty.

PNM regularity condition to have a d

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

Hybrid version: if NM works $\checkmark$, otherwise PNM (seldom).
Excellent performance on random data / applications.

Now, we try to remove the regularity.

Outline

What if the polyhedron is empty?

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

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

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 $\mathcal{E}^{0+}(x)$, $\gamma_+ = (\gamma_i)_i \in \{0,1\}^{\mathcal{E}^{0+}(x)}$, $\overline{\gamma}_i := 1 - \gamma_i$ (or $[0,1]$)
$\{0,1\}$: partitionning (F_i or G_i)

What if the polyhedron is empty?

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

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

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 $\mathcal{E}^{0+}(x)$, $\gamma_+ = (\gamma_i)_i \in \{0,1\}^{\mathcal{E}^{0+}(x)}$, $\overline{\gamma}_i := 1 - \gamma_i$ (or $[0,1]$) $\{0,1\}$: partitionning (F_i or G_i)

What if the polyhedron is empty?

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 $\mathcal{E}^{0+}(x)$, $\gamma_+ = (\gamma_i)_i \in \{0,1\}^{\mathcal{E}^{0+}(x)}$, $\overline{\gamma}_i := 1 - \gamma_i$ (or $[0,1]$)
$\{0,1\}$: partitionning (F_i or G_i)

Brief comment on Levenberg-Marquardt (LM)

LM: $\min_d \frac{1}{2} q_x(d) + \frac{\lambda}{2} d^\mathsf{T} S d$, curve $\lambda \mapsto d(\lambda)$.

Brief comment on Levenberg-Marquardt (LM)

LM: $\min_d \frac{1}{2} q_x(d) + \frac{\lambda}{2} d^\top S d$, curve $\lambda \mapsto d(\lambda)$.

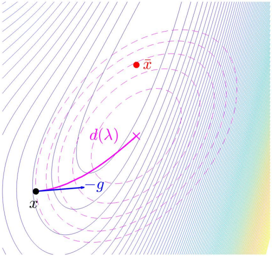


Illustration in the smooth case. When $\lambda \searrow 0$, $d(\lambda) \to$ Newton direction, when $\lambda \to +\infty$: $d(\lambda) \to -S^{-1}\nabla\theta$ the tangent of the curve.

Least-squares formulation

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

$$\min_{d \in \mathbb{R}^n} \psi_x(d) := \frac{1}{2}[q_x(d) + \lambda d^\top S d], \quad \lambda \geqslant 0, S \succ 0 \tag{5}$$

- 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?

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

Least-squares formulation

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

$$\min_{d \in \mathbb{R}^n} \psi_x(d) := \frac{1}{2}[q_x(d) + \lambda d^\mathsf{T} S d], \quad \lambda \geqslant 0, S \succ 0 \tag{5}$$

- 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?

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

Least-squares formulation

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

$$\min_{d \in \mathbb{R}^n} \psi_x(d) := \frac{1}{2}[q_x(d) + \lambda d^\mathsf{T} S d], \quad \lambda \geqslant 0, S \succ 0 \tag{5}$$

- 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?

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

Least-squares formulation

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

$$\min_{d \in \mathbb{R}^n} \psi_x(d) := \frac{1}{2}[q_x(d) + \lambda d^\mathsf{T} S d], \quad \lambda \geqslant 0, S \succ 0 \tag{5}$$

- 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?

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

Outline

1. Motivation
   - Problem presentation
   - One nonsmooth method
2. Hyperplane Arrangements
   - Computation of $\partial_B H(x)$
   - Some improvements
   - A new approach by duality
   - Some results
3. LM-PNM
   - Presentation of NM and PNM
   - Least-squares over regularity
   - Technical choice of the weights
   - Algorithmical considerations

Computations

Goal: decrease the nonsmooth merit function

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

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

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

Choosing correct weights

Lemma: existence of appropriate weights

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

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

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

Choosing correct weights

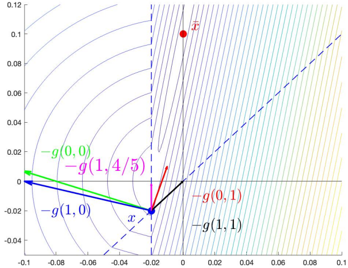
Lemma: existence of appropriate weights

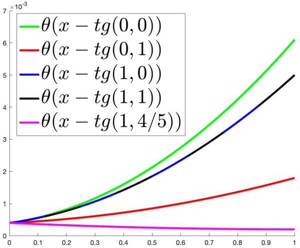
Let γ_+, $\exists\ \gamma_-(\gamma_+)$ s.t. $\rho_{\mathcal{E}^-(x)}(x, \gamma_-(\gamma_+), g(\gamma_+, \gamma_-(\gamma_+))) = 0$:

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

- **way to choose a good piece**,
- possibly multiple $\gamma_-(\gamma_+)$, $\gamma_+ \mapsto g(\gamma_+, \gamma_-(\gamma_+))$ unique;
- if $g(\gamma_+, \gamma_-(\gamma_+)) = 0$, γ_+ irrelevant;
- 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.

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

1. Motivation
   - Problem presentation
   - One nonsmooth method
2. Hyperplane Arrangements
   - Computation of $\partial_B H(x)$
   - Some improvements
   - A new approach by duality
   - Some results
3. LM-PNM
   - Presentation of NM and PNM
   - Least-squares over regularity
   - Technical choice of the weights
   - Algorithmical considerations

Stationarity detection with the weights

Characterizing stationarity with the γ's:

The following properties are equivalent:

1) x is θ-stationary (i.e., $\forall\ d \in \mathbb{R}^n, \theta'(x; d) \geqslant 0$);
2) for **all** $\gamma_+ \in [0,1]^{\mathcal{E}^{0+}(x)}$, $g(\gamma_+, \gamma_-(\gamma_+)) = 0$.

Zonotope := $Z(M) = M[-1,+1]^q$, $M \in \mathbb{R}^{n\times q}$.

Stationarity detection reformulated

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

Stationarity detection with the weights

Characterizing stationarity with the γ's:

The following properties are equivalent:

1) x is θ-stationary (i.e., $\forall\ d \in \mathbb{R}^n, \theta'(x; d) \geqslant 0$);
2) for **all** $\gamma_+ \in [0,1]^{\mathcal{E}^{0+}(x)}$, $g(\gamma_+, \gamma_-(\gamma_+)) = 0$.

Zonotope $:= Z(M) = M[-1,+1]^q$, $M \in \mathbb{R}^{n\times q}$.

Stationarity detection reformulated

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

Some observations

- any γ_+ with $g(\gamma_+, \gamma_-(\gamma_+)) \neq 0$ ✓ (disproves inclusion);
- combinatorial aspect in $|\mathcal{E}^{0+}(x)|$ and $|\mathcal{E}^{-}(x)|$, maybe small;
- some conditions ⇒ polynomial time ([ST19], App. C).
- $\gamma_-(\gamma_+)$ 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 $\lambda^k, \gamma_+, \gamma_-, S_k$
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} and S_{k+1}

No assumption on (F, G), very costly intermediate computation.

Properties

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 $\lambda^k, \gamma_+, \gamma_-, S_k$
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} and S_{k+1}

No assumption on (F, G), very costly intermediate computation.

Properties

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),
- very high cost in practice (γ_+, γ_- and LM steps).

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),
- very high cost in practice (γ_+, γ_- and LM steps).

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),
- very high cost in practice (γ_+, γ_- and LM steps).

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),
- very high cost in practice (γ_+, γ_- and LM steps).

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

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

Conclusion of the thesis

Main take-aways

- H induces an arrangement structure
- duality and heuristics speed-up the computation of chambers
- CPs: γ_+, γ_- clarified by dualized arrangements (zonotopes)

Possible extensions:

- tailoring RČ for specific arrangements;
- for all the arrangement ($\{-1, 0, +1\}^p$);
- better LM convergence theorem;
- refined algorithm.

Thank you for your attention! Any question?

Conclusion of the thesis

Main take-aways

- H induces an arrangement structure
- duality and heuristics speed-up the computation of chambers
- CPs: γ_+, γ_- clarified by dualized arrangements (zonotopes)

Possible extensions:

- tailoring RČ for specific arrangements;
- for all the arrangement ($\{-1, 0, +1\}^p$);
- better LM convergence theorem;
- refined algorithm.

Thank you for your attention! Any question?

Conclusion of the thesis

Main take-aways

- H induces an arrangement structure
- duality and heuristics speed-up the computation of chambers
- CPs: γ_+, γ_- clarified by dualized arrangements (zonotopes)

Possible extensions:

- tailoring RČ for specific arrangements;
- for all the arrangement ($\{-1, 0, +1\}^p$);
- better LM convergence theorem;
- refined algorithm.

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.

[AF92] David Avis and Komei Fukuda. "A Pivoting Algorithm for Convex Hulls and Vertex Enumeration of Arrangements and Polyhedra". In: *Discrete Computational Geometry* 8 (1992), pp. 295–31. ISSN: 0179-5376,1432-0444. DOI: 10.1007/BF02293050.

[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.

[BEK23] Taylor Brysiewicz, Holger Eble, and Lukas Kühne. "Computing Characteristic Polynomials of Hyperplane Arrangements with Symmetries". In: *Discrete & Computational Geometry* 70.4 (Dec. 2023), pp. 1356–1377. ISSN: 0179-5376, 1432-0444. DOI: 10.1007/s00454-023-00557-2.

[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.

[BN82] Hanspeter Bieri and Walter Nef. "A Recursive Sweep-Plane Algorithm, Determining All Cells of a Finite Division of Rˆd". In: *Computing* 28 (1982), pp. 189–198.

[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].

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.

[DGP25a] 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).

[DGP25b] Jean-Pierre Dussault, Jean Charles Gilbert, and Baptiste Plaquevent-Jourdain. "Primal and Dual Approaches for the Chamber Enumeration of Real Hyperplane Arrangements". In: *(submitted)* (2025).

Bibliographic elements III

[DP22] Antoine Deza and Lionel Pournin. "A Linear Optimization Oracle for Zonotope Computation". In: *Computational Geometry* 100 (Jan. 2022), p. 101809. ISSN: 09257721. DOI: 10.1016/j.comgeo.2021.101809.

[DSL06] György Dósa, István Szalkai, and Claude Laflamme. "The Maximum and Minimum Number of Circuits and Bases of Matroids.". In: *Pure Mathematics and Applications*. Mathematics of Optimization 15.4 (Sept. 2006), pp. 383–392.

[EOS86] Herbert Edelsbrunner, Joseph O'ROURKE, and Raimund Seidel. "CONSTRUCTING ARRANGEMENTS OF LINES AND HYPERPLANES WITH APPLICATIONS". In: *SIAM Journal on Computation* 15.2 (1986), pp. 341–363.

[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.

Bibliographic elements IV

[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.

[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].

[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

[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.

[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.

[MM24] Arturo Merino and Torsten Mütze. "Traversing Combinatorial 0/1-Polytopes via Optimization". In: *SIAM Journal on Computing* 53.5 (Oct. 2024), pp. 1257–1292. ISSN: 0097-5397, 1095-7111. DOI: 10.1137/23M1612019.

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

[Oxl11] James G. Oxley. *Matroid Theory*. Second edition. Oxford Graduate Texts in Mathematics 21. Oxford New York, NY: Oxford University Press, 2011. ISBN: 978-0-19-856694-6 978-0-19-960339-8.

[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.

Bibliographic elements VI

[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.

[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.

[Ram23] Jörg Rambau. "Symmetric Lexicographic Subset Reverse Search for the Enumeration of Circuits, Cocircuits, and Triangulations up to Symmetry". In: (2023), pp. 1–41.

[RČ18] Miroslav Rada and Michal Černý. "A New Algorithm for Enumeration of Cells of Hyperplane Arrangements and a Comparison with Avis and Fukuda's Reverse Search". In: *SIAM Journal on Discrete Mathematics* 32.1 (Jan. 2018), pp. 455–473. ISSN: 0895-4801, 1095-7146. DOI: 10.1137/15M1027930.

[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.

Bibliographic elements VII

[Sle00] Nora Helena Sleumer. “Hyperplane Arrangements: Construction, Visualization and Applications”. PhD thesis. Zurich, Switzerland: Swiss Federal Institute of Technology, 2000.

[Sle98] Nora Sleumer. “Output-Sensitive Cell Enumeration in Hyperplane Arrangements”. In: *Algorithm Theory — SWAT’98*. Ed. by Gerhard Goos et al. Vol. 1432. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 300–309. ISBN: 978-3-540-64682-2 978-3-540-69106-8. DOI: 10.1007/BFb0054377.

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

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

[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.

[Win66] Robert Owen Winder. “Partitions of N -Space by Hyperplanes”. In: *SIAM Journal on Applied Mathematics* 14.4 (July 1966), pp. 811–818. ISSN: 0036-1399, 1095-712X. DOI: 10.1137/0114068.

[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.

[Zas75] Thomas Zaslavsky. “Facing up to Arrangements: Face-Count Formulas for Partitions of Space by Hyperplanes”. In: *Memoirs of the American Mathematical Society* 1.154 (1975), 1–109 (?)

Bibliographic elements VIII

[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.

[Zie99] Günter M. Ziegler. *Lectures on 0/1-Polytopes*. Sept. 1999. arXiv: math/9909177.

Number of elements in the differentials

$F(x) = x$, $G(x) = (ax_1 + bx_2, cx_1 + dx_2)$ with $a, b, c, d \in \mathbb{R}$.
At $x^* = 0$, $|\partial_B H(x^*)| \in \{2, 4\}$. What about $\partial_B H_{\mathrm{FB}}(x^*)$?

For $x \neq 0$ (see also Izmailov Solodov [IS14, p.151]):

$$H'_{\mathrm{FB}}(x) = \begin{bmatrix} \frac{x_1 + a(ax_1+bx_2)}{\|(x_1, ax_1+bx_2)\|} - 1 - a & \frac{b(ax_1+bx_2)}{\|(x_1, ax_1+bx_2)\|} - b \\ \frac{c(cx_1+dx_2)}{\|(x_2, cx_1+dx_2)\|} - c & \frac{x_2 + d(cx_1+dx_2)}{\|(x_2, cx_1+dx_2)\|} - 1 - d \end{bmatrix}$$

$\partial_B H_{\mathrm{FB}}(x^*)$ is a continuum; lines interact (like with H).

Number of elements in the differentials

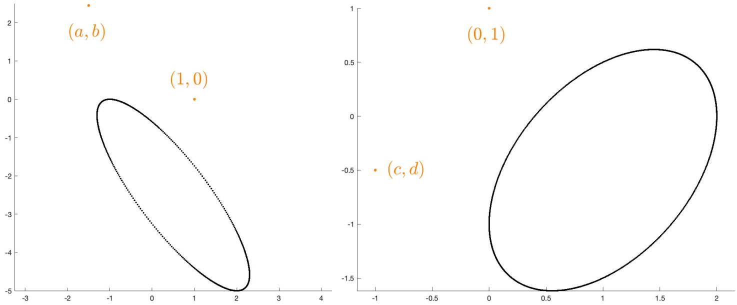
$F(x) = x$, $G(x) = (ax_1 + bx_2, cx_1 + dx_2)$ with $a, b, c, d \in \mathbb{R}$.
At $x^* = 0$, $|\partial_B H(x^*)| \in \{2, 4\}$. What about $\partial_B H_{\mathrm{FB}}(x^*)$?

For $x \neq 0$ (see also Izmailov Solodov [IS14, p.151]):

$$H'_{\mathrm{FB}}(x) = \begin{bmatrix} \frac{x_1 + a(ax_1+bx_2)}{||(x_1, ax_1+bx_2)||} - 1 - a & \frac{b(ax_1+bx_2)}{||(x_1, ax_1+bx_2)||} - b \\ \frac{c(cx_1+dx_2)}{||(x_2, cx_1+dx_2)||} - c & \frac{x_2 + d(cx_1+dx_2)}{||(x_2, cx_1+dx_2)||} - 1 - d \end{bmatrix}$$

$\partial_B H_{\mathrm{FB}}(x^*)$ is a continuum; lines interact (like with H).

Illustration



Left: values for the first line of the Jacobians. Right: values for the second line of the Jacobians. In orange, there are $2 \times 2 = 4$ elements for $\partial_B H(x^*)$ whereas those in $\partial_B H_{\mathrm{FB}}(x^*)$ form a continuum (each dot of one black curve corresponds to one dot on the other).

Summary of some basic methods

		$\partial_?$	reg(ularity) at x^*	$\lvert\partial_?\rvert$	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

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

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

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

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

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

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

Computing the B-differential - multiple indices

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

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

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

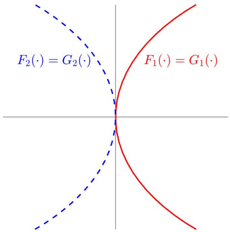
$H_i := (B_{i,:} - A_{i,:})^{\perp} := v_i^{\perp}$, $v_i \neq 0$ ($C2$) ; for ∂_B, $\mathbb{R}^n \backslash (x + \cup_{i=1}^{p} H_i)$

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

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

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): $\to$ 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.

Cardinality properties

We want $\mathcal{S}$, though some properties on $|\mathcal{S}|$ are well-known.

Some formulas

- Bound ($\mathrm{rank}(V) = n$ to simplify) Schläfli [Sch50]:

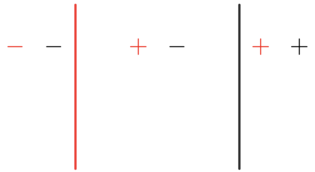
$$|\mathcal{S}(V,\tau)| \leqslant \sum_{i=0}^{n} \binom{p}{i} \quad (\leqslant 2^p);$$

  attained in *general position* ($\simeq$ random V, τ).
- For $|\mathcal{S}|$: Winder, Zaslavsky [Win66; Zas75] (but not $\mathcal{S}$):

$$|\mathcal{S}(V,\tau)| = \sum_{I \subseteq [1:p], \tau_I \in \mathcal{R}(V_{:,I}^\top)} (-1)^{\mathrm{null}(V_{:,I})} = (-1)^n \chi(-1).$$

Illustration of duality

$M = s \cdot V^\top$, $m = s \cdot \tau$: $s \cdot (V^\top x - \tau) > 0 \Leftrightarrow s \cdot V^\top x > s \cdot \tau$



With $V = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $\tau = [-1; 1]$, $\{x : x_1 = -1\}$ and $\{x : x_1 = +1\}$.
No $-+$ since (geometrically) $-$: left to the red hyperplane and $+$ right to the black hyperplane. Algebraically, $-$ means $x_1 < -1$ and $+$ $x_1 > 1$.

$$\alpha = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; (V \cdot [- \;\; +])\alpha = \begin{bmatrix} - & + \\ 0 & 0 \end{bmatrix} \alpha = 0, \; ([- \;\; +] \cdot \tau)\alpha = 2 \geqslant 0$$

About circuits/stem vectors

$$\mathcal{C}(V) := \{I \subseteq [1:p] : \mathrm{null}(V_{:,I}) = 1, \mathrm{null}(V_{:,I_0}) = 0 \ \forall\ I_0 \subsetneq I\}$$

No “good” algo (Rambau [Ram23]); adaptable for symmetries.
Upper bound $\binom{p}{r+1}$ [DSL06], $=$ under general position.
“Double punishment” for fully dual method.

For degenerate arrangements, short circuits so less susbets explored, but maybe lots of circuits (p large).
Ex: parallel hyperplanes – circuits of size 2 (so no larger subsets).

About circuits/stem vectors

$$\mathcal{C}(V) := \{I \subseteq [1:p] : \mathrm{null}(V_{:,I}) = 1, \mathrm{null}(V_{:,I_0}) = 0 \ \forall\ I_0 \subsetneq I\}$$

No "good" algo (Rambau [Ram23]); adaptable for symmetries.
Upper bound $\binom{p}{r+1}$ [DSL06], $=$ under general position.
"Double punishment" for fully dual method.

For degenerate arrangements, short circuits so less susbets explored, but maybe lots of circuits (p large).
Ex: parallel hyperplanes – circuits of size 2 (so no larger subsets).

Affine or linear?

coning/homogeneization/embedding/lifting/...

$$\mathcal{S}\left(\begin{bmatrix} V & 0 \\ \tau & -1 \end{bmatrix}, 0\right) = [\mathcal{S}(V,\tau) \times \{+1\}] \cup [-\mathcal{S}(V,\tau) \times \{-1\}],$$

i.e., "an affine arrangement in dimension n is the upper [or lower] half of a centered arrangement in dimension $n+1$".

Natural way so swap between affine and linear arrangements
$\mathcal{S}(V,\tau) :=$ affine$(n,p) \simeq$ linear$(n+1,p+1)$ (half of);
$\mathcal{S}(V,0) :=$ linear$(n,p) \simeq$ affine$(n-1,p-1)$ (two opposite).

Details on compaction

$$\begin{cases} \mathcal{S}(V,0) & := \{s \in \{\pm 1\}^p : \exists\, x^s \in \mathbb{R}^n : s \bullet V^\mathsf{T} x^s > 0\} \\ \mathcal{S}(V,\tau) & := \{s \in \{\pm 1\}^p : \exists\, x^s \in \mathbb{R}^n : s \bullet (V^\mathsf{T} x^s - \tau) > 0\} \\ \mathcal{S}([V;\tau^\mathsf{T}],0) & := \{s \in \{\pm 1\}^p : \exists\, d^s \in \mathbb{R}^{n+1} : s \bullet [V^\mathsf{T}\ \tau] d^s > 0\} \end{cases}$$

$\mathcal{S}(V,\tau)$ has a *symmetric part* (not perfectly geometrically).

$\mathcal{S}(V,\tau)$ exactly between $\mathcal{S}(V,0)$ and $\mathcal{S}([V;\tau^\mathsf{T}],0)$ (symmetric).

Possible to quantify the difference in # of LOPs.

Compute less than $\mathcal{S}(V,\tau)$ chambers.

Details on compaction

$$\begin{cases} \mathcal{S}(V,0) & := \{s \in \{\pm 1\}^p : \exists\, x^s \in \mathbb{R}^n : s \bullet V^\mathsf{T} x^s > 0\} \\ \mathcal{S}(V,\tau) & := \{s \in \{\pm 1\}^p : \exists\, x^s \in \mathbb{R}^n : s \bullet (V^\mathsf{T} x^s - \tau) > 0\} \\ \mathcal{S}([V;\tau^\mathsf{T}],0) & := \{s \in \{\pm 1\}^p : \exists\, d^s \in \mathbb{R}^{n+1} : s \bullet [V^\mathsf{T}\ \tau] d^s > 0\} \end{cases}$$

$\mathcal{S}(V,\tau)$ has a *symmetric part* (not perfectly geometrically).

$\mathcal{S}(V,\tau)$ exactly between $\mathcal{S}(V,0)$ and $\mathcal{S}([V;\tau^\mathsf{T}],0)$ (symmetric).

Possible to quantify the difference in # of LOPs.

Compute less than $\mathcal{S}(V,\tau)$ chambers.

The unifying method, Merino Mütze [MM24]?

$\{\pm 1\} \to \{0,1\}$, *connected* vertices X of the hypercube.
A priori: the path may not be connected in $\mathbb{R}^n$;
To next chamber: binary variable, not LO

$$\min_{y,z} w^\top (y - x), \quad y_{P_0} = 0, \quad y_{P_1} = 1, \quad (2y-1) \bullet (V^\top z - \tau) > 0?$$

For vertices of $P = \{z : Az \leqslant b\}$ **assumes it is a** $\mathrm{conv}(X)$ from A and b. (Not obvious according to Ziegler [Zie99]?)

For circuits? $x(C)_i := \mathbb{1}(i \in C)$, $x(C) \in \{0,1\}^n$, $C(x) = \cup_{x_j=1}\{j\}$.
No "swaps" (flips) for circuits. The exchange axiom: 3 circuits...

$$\min_{y} w^\top (y - x), \quad y_{P_0} = 0, \quad y_{P_1} = 1, \quad \begin{cases} \mathrm{null}(V_{:,C(y)}) = 1, \\ \mathrm{null}(V_{:,C'}) = 0, C' \subsetneq C(y)? \end{cases}$$

The unifying method, Merino Mütze [MM24]?

$\{\pm 1\} \to \{0,1\}$, *connected* vertices X of the hypercube.
A priori: the path may not be connected in $\mathbb{R}^n$;
To next chamber: binary variable, not LO

$$\min_{y,z} w^\top(y-x), \quad y_{P_0}=0, \quad y_{P_1}=1, \quad (2y-1)\bullet(V^\top z-\tau)>0?$$

For vertices of $P=\{z : Az \leqslant b\}$ **assumes it is a** $\mathrm{conv}(X)$ from A and b. (Not obvious according to Ziegler [Zie99]?)

For circuits? $x(C)_i := \mathbb{1}(i \in C)$, $x(C) \in \{0,1\}^n$, $C(x) = \cup_{x_j=1}\{j\}$.
No "swaps" (flips) for circuits. The exchange axiom: 3 circuits...

$$\min_{y} w^\top(y-x), \quad y_{P_0}=0, \quad y_{P_1}=1, \quad \begin{cases} \mathrm{null}(V_{:,C(y)})=1, \\ \mathrm{null}(V_{:,C'})=0, C' \subsetneq C(y)? \end{cases}$$

The unifying method, Merino Mütze [MM24]?

$\{\pm 1\} \to \{0,1\}$, *connected* vertices X of the hypercube.
A priori: the path may not be connected in $\mathbb{R}^n$;
To next chamber: binary variable, not LO

$$\min_{y,z} w^\mathsf{T}(y-x), \quad y_{P_0} = 0, \quad y_{P_1} = 1, \quad (2y-1) \bullet (V^\mathsf{T} z - \tau) > 0?$$

For vertices of $P = \{z : Az \leqslant b\}$ **assumes it is a** $\operatorname{conv}(X)$ from A and b. (Not obvious according to Ziegler [Zie99]?)

For circuits? $x(C)_i := \mathbb{1}(i \in C)$, $x(C) \in \{0,1\}^n$, $C(x) = \cup_{x_j=1}\{j\}$.
No "swaps" (flips) for circuits. The exchange axiom: 3 circuits...

$$\min_{y} w^\mathsf{T}(y-x), \;\; y_{P_0} = 0, \;\; y_{P_1} = 1, \;\; \begin{cases} \mathrm{null}(V_{:,C(y)}) = 1, \\ \mathrm{null}(V_{:,C'}) = 0, C' \subsetneq C(y)? \end{cases}$$

Full arrangements: not only halfspaces

With $\{-1, 0, +1\}^p$, what changes? $2^p \to 3^p$,
known bounds (general position), RC algorithm with ternary tree.

Some things need to be adapted: especially compaction (relations).
Main issue: equalities ($s_i = 0$) are not maintained if $\tau \neq 0$.

For σ's, no changes? "chamber infeasible has no boundary":
so stem vectors $\sigma \in \{\pm 1\}^I$ mean every "$s^I \in [0, \sigma]$" infeasible too.

Algorithmically? Tree has 1 / 3 descendants (two $\Rightarrow$ third).
Compute chambers, join neighbors for $n-1$ subchambers, $n-2$...
Compute intersections of H_i's and binary trees on them, project in the subspaces...

Full arrangements: not only halfspaces

With $\{-1, 0, +1\}^p$, what changes? $2^p \to 3^p$,
known bounds (general position), RC algorithm with ternary tree.

Some things need to be adapted: especially compaction (relations).
Main issue: equalities ($s_i = 0$) are not maintained if $\tau \neq 0$.

For σ's, no changes? "chamber infeasible has no boundary":
so stem vectors $\sigma \in \{\pm 1\}^I$ mean every "$s^I \in [0, \sigma]$" infeasible too.

Algorithmically? Tree has 1 / 3 descendants (two $\Rightarrow$ third).
Compute chambers, join neighbors for $n-1$ subchambers, $n-2$...
Compute intersections of H_i's and binary trees on them, project in the subspaces...

Full arrangements: not only halfspaces

With $\{-1, 0, +1\}^p$, what changes? $2^p \to 3^p$,
known bounds (general position), RC algorithm with ternary tree.

Some things need to be adapted: especially compaction (relations).
Main issue: equalities ($s_i = 0$) are not maintained if $\tau \neq 0$.

For σ's, no changes? "chamber infeasible has no boundary":
so stem vectors $\sigma \in \{\pm 1\}^I$ mean every "$s^I \in [0, \sigma]$" infeasible too.

Algorithmically? Tree has 1 / 3 descendants (two $\Rightarrow$ third).
Compute chambers, join neighbors for $n-1$ subchambers, $n-2$...
Compute intersections of H_i's and binary trees on them, project in the subspaces...

Full arrangements: not only halfspaces

With $\{-1, 0, +1\}^p$, what changes? $2^p \to 3^p$,
known bounds (general position), RC algorithm with ternary tree.

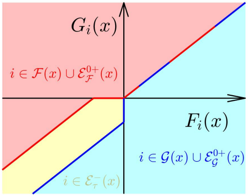
Some things need to be adapted: especially compaction (relations).
Main issue: equalities ($s_i = 0$) are not maintained if $\tau \neq 0$.

For σ's, no changes? "chamber infeasible has no boundary":
so stem vectors $\sigma \in \{\pm 1\}^I$ mean every "$s^I \in [0, \sigma]$" infeasible too.

Algorithmically? Tree has 1 / 3 descendants (two $\Rightarrow$ third).
Compute chambers, join neighbors for $n-1$ subchambers, $n-2\ldots$
Compute intersections of H_i's and binary trees on them, project in the subspaces...

Details on PNM

$\mathcal{E}^-(x)$ becomes, for $\tau > 0$ small
$F_i(x) < 0$, $G_i(x) < 0$
and $|F_i(x) - G_i(x)| < \tau$.



PNM regularity condition to have a d

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

Hybrid NM: most often, d(NM) works ✓; PNM if iterate difficult; spectacular on random data / applications.

Explicit computations

One wants $\theta'(x; -g(\gamma_+, \gamma_-)) \leqslant 0$.
Let $\Gamma_+ = \mathrm{Diag}(\gamma_+), \Gamma_- = \mathrm{Diag}(\gamma_-)$ g

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

$$0 \geqslant \quad + H^\top_{\mathcal{E}0+(x)} \underbrace{[\min(-F'_{\mathcal{E}0+(x)} g(\gamma_+, \gamma_-), -G'_{\mathcal{E}0+(x)} g(\gamma_+, \gamma_-)) + \Gamma_+ F'_{\mathcal{E}0+(x)} g(\gamma_+, \gamma_-) + \overline{\Gamma}_+ G'_{\mathcal{E}0+(x)} g(\gamma_+, \gamma_-)]}_{\leqslant 0}$$

$$0 \leqslant \quad + H^\top_{\mathcal{E}-(x)} \underbrace{[\min(-F'_{\mathcal{E}-(x)} g(\gamma_+, \gamma_-), -G'_{\mathcal{E}-(x)} g(\gamma_+, \gamma_-)) + \Gamma_- F'_{\mathcal{E}-(x)} g(\gamma_+, \gamma_-) + \overline{\Gamma}_- G'_{\mathcal{E}-(x)} g(\gamma_+, \gamma_-)]}_{\leqslant 0}$$

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

Explicit computations

One wants $\theta'(x; -g(\gamma_+, \gamma_-)) \leqslant 0$.

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

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

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

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

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

Explicit computations

One wants $\theta'(x; -g(\gamma_+, \gamma_-)) \leqslant 0$.
Let $\Gamma_+ = \mathrm{Diag}(\gamma_+), \Gamma_- = \mathrm{Diag}(\gamma_-)$ g

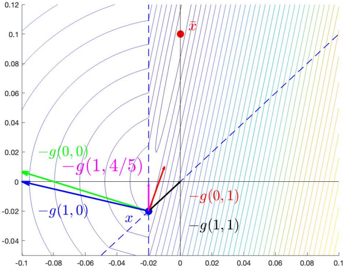
$$\theta'(x; -g(\gamma_+, \gamma_-)) = -||g(\gamma_+, \gamma_-)||^2$$

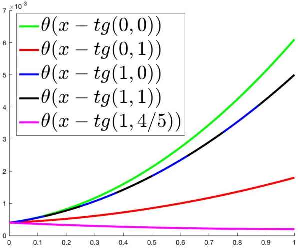
$$0 \geqslant \ + H^{\mathsf{T}}_{\mathcal{E}^{0+}(x)}\underbrace{[\min(-F'_{\mathcal{E}^{0+}(x)}g(\gamma_+, \gamma_-), -G'_{\mathcal{E}^{0+}(x)}g(\gamma_+, \gamma_-)) + \Gamma_+ F'_{\mathcal{E}^{0+}(x)}g(\gamma_+, \gamma_-) + \overline{\Gamma}_+ G'_{\mathcal{E}^{0+}(x)}g(\gamma_+, \gamma_-)]}_{\leqslant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

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

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

Example of weights computation – 1





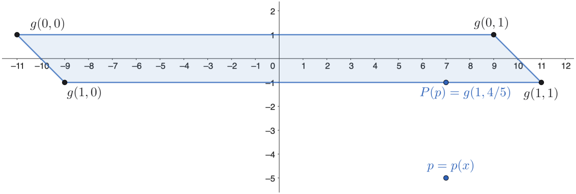
$$M = \begin{pmatrix} 1/2 & 1/2 \\ -5 & 1 \end{pmatrix}, \quad q = \begin{pmatrix} 0 \\ -1/10 \end{pmatrix}, \quad x = \begin{pmatrix} -1/50 \\ -1/50 \end{pmatrix}$$

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 $g(1,4/5) = 4/5g(1,1) + 1/5g(1,0)$

Computation of the weights

Goal

Finding the best γ_+ such that $g(\gamma_+, \gamma_-(\gamma_+)) \neq 0$ (or stationarity).

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

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

The outer max is a convex function (distance) on a hypercube: maximized on a vertex, combinatorial nature $\rightsquigarrow \{0, 1\}^{\mathcal{E}^{0+}(x)}$, partitions of $\mathcal{E}^{0+}(x)$, inclusion by vertices (strict local maxima also work).

Computation of the weights

Goal

Finding the best γ_+ such that $g(\gamma_+, \gamma_-(\gamma_+)) \neq 0$ (or stationarity).

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

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

The outer max is a convex function (distance) on a hypercube: maximized on a vertex, combinatorial nature $\rightsquigarrow \{0,1\}^{\mathcal{E}^{0+}(x)}$, partitions of $\mathcal{E}^{0+}(x)$, inclusion by vertices (strict local maxima also work).

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.

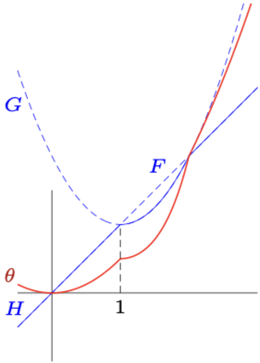
1) The sequence $(\theta(x_k))_k$ decreases thus converges.
2) If $(F'(x_k), G'(x_k), \lambda_k S_k)_{k \in \mathcal{K}}$ for a subsequence $\mathcal{K}$ is bounded, then $g_k \to 0$.

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

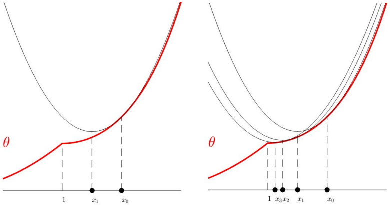


Counter-example

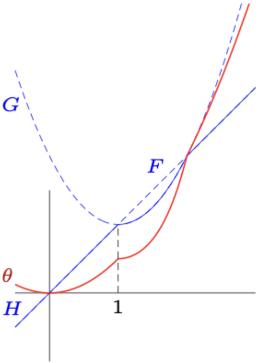
"Concave kinks" – difficult (bad) limit points.

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