Polyhedral Newton-min algorithms for complementarity problems [28]

Jean Charles Gilbert (Inria-Paris & Université de Sherbrooke)

Joint work with

Jean-Pierre Dussault (Université de Sherbrooke)
Mathieu Frappier (Université de Sherbrooke)

June 7, 2023

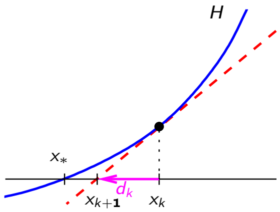
Outline

Outline

Local Newton's method for a smooth function

- Let $H : \mathbb{E} \to \mathbb{F}$ be a smooth function ($\mathbb{E}$ a vector space).
- Find $x_* \in \mathbb{E}$ such that $H(x_*) = 0$?
- Local Newton's algorithm:

$$\begin{cases} H(x_k) + H'(x_k)d_k = 0 \\ x_{k+1} := x_k + d_k. \end{cases}$$



- 3 conditions for quadratic convergence
  - x_0 close to x_*,
  - $H \in \mathcal{C}^{1,1}$,
  - $H'(x_*)$ nonsingular.

Globalization of Newton's method for a smooth function: miracle or mirage?

- Let $(\mathbb{F}, \langle\cdot,\cdot\rangle)$ be a Euclidean space; associated norm $\|\cdot\| = \langle\cdot,\cdot\rangle^{1/2}$.
- Consider the least-square merit function: $\theta : \mathbb{E} \to \mathbb{R}$ defined at $x \in \mathbb{E}$ by

$$\theta(x) := \frac{1}{2}\|H(x)\|^2.$$

- Miracle: the Newton's direction $d := -H'(x)^{-1}H(x)$ is a descent direction of θ:

$$\theta'(x)d = \langle H(x), H'(x)d\rangle = -\|H(x)\|^2 = -2\theta(x) < 0 \qquad [\text{if } d \text{ exists and } H(x) \neq 0]$$

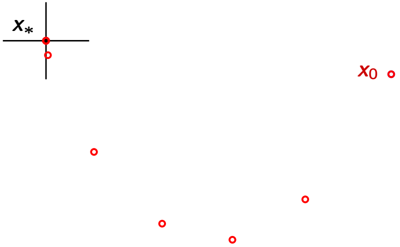
- Globalization by linesearch: $x_{k+1} := x_k + \alpha_k d_k$ with $\alpha_k > 0$ *not too small* such that

$$\theta(x_k + \alpha_k d_k) \leqslant \theta(x_k) + \omega\alpha_k\theta'(x_k)d_k \qquad [\omega \simeq 10^{-4}].$$

- Mirage: If $\bar{x}$ is a limit point of $\{x_k\}$, that is regular ($F'(\bar{x})$ nonsingular), then $F(\bar{x}) = 0$.

But there may be no such limit point!

Success of the globalization of Newton's algorithm with LS



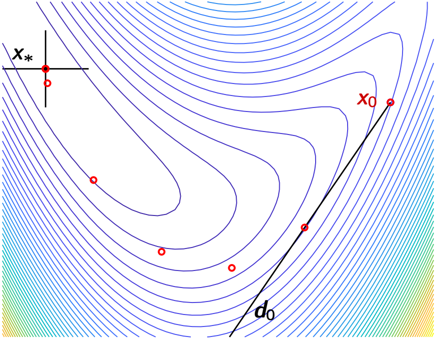
$$F(x) = \begin{pmatrix} x_1 \\ -(x_1-2)^2 + x_2 + 4 \end{pmatrix}$$

$$F'(x) = \begin{pmatrix} 1 & 0 \\ -2(x_1-2) & 1 \end{pmatrix}.$$

Success of the globalization of Newton's algorithm with LS

$$F(x) = \begin{pmatrix} x_1 \\ -(x_1-2)^2 + x_2 + 4 \end{pmatrix}$$

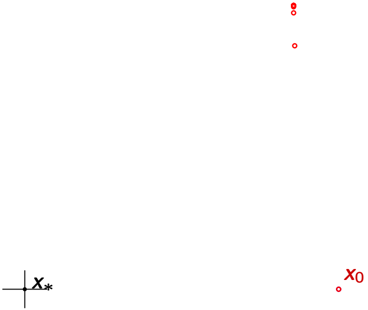
$$F'(x) = \begin{pmatrix} 1 & 0 \\ -2(x_1-2) & 1 \end{pmatrix}.$$



Failure of the globalization of Newton's algorithm with LS I

$$F(x) = \begin{pmatrix} x_1 \\ -(x_1-2)^2 + (x_2-1)^2 + 3 \end{pmatrix},$$

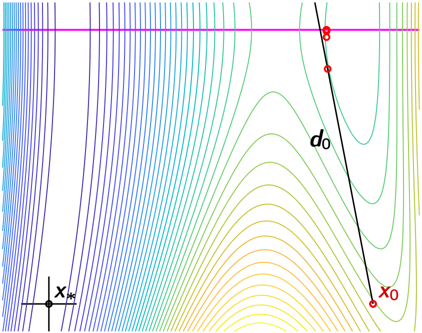
$$F'(x) = \begin{pmatrix} 1 & 0 \\ -2(x_1-2) & 2(x_2-1) \end{pmatrix}.$$



Failure of the globalization of Newton's algorithm with LS I

$$F(x) = \begin{pmatrix} x_1 \\ -(x_1-2)^2 + (x_2-1)^2 + 3 \end{pmatrix},$$

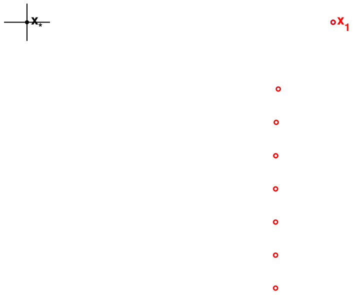
$$F'(x) = \begin{pmatrix} 1 & 0 \\ -2(x_1-2) & 2(x_2-1) \end{pmatrix}.$$



Failure of the globalization of Newton's algorithm with LS II

$$F(x) = \begin{pmatrix} x_1 \\ -(x_1 - 2)^2 + e^{x_2} + 3 \end{pmatrix},$$

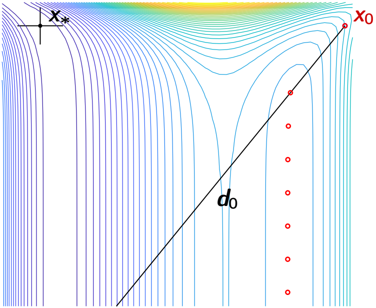
$$F'(x) = \begin{pmatrix} 1 & 0 \\ -2(x_1 - 2) & e^{x_2} \end{pmatrix}.$$



Failure of the globalization of Newton's algorithm with LS II

$$F(x) = \begin{pmatrix} x_1 \\ -(x_1-2)^2 + e^{x_2} + 3 \end{pmatrix},$$

$$F'(x) = \begin{pmatrix} 1 & 0 \\ -2(x_1-2) & e^{x_2} \end{pmatrix}.$$



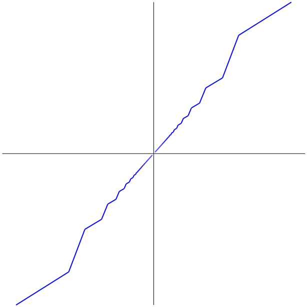
Failure of the globalization of Newton's algorithm with LS III

Conclusion

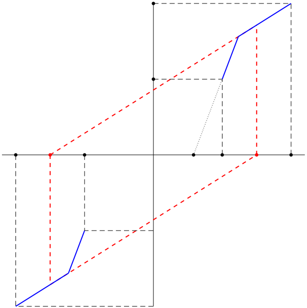
- A "global" convergence result of the kind "any regular limit point of the generated sequence is a solution" must be taken with caution, since the generated sequence may have no regular limit point.
- Such a "global" convergence result is just a means to improve algorithms.

Local Newton's method for a nonsmooth function may fail

Newton's method may cycle, regardless of the proximity of x_0 and x_*.
Example, Kummer's function [49; 1988] (differentiable at 0, $\partial_C H(0) = [1/2, 2] \not\ni 0$)



Kummer's function



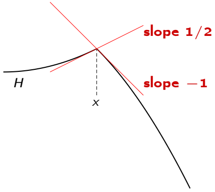
Cycling of Newton's algorithm

B-differential and C-differential

- Let $\mathbb{E}$ and $\mathbb{F}$ be two vector spaces of finite dimensions $n := \dim \mathbb{E}$ and $m := \dim \mathbb{F}$.
- Let $H : \mathbb{E} \to \mathbb{F}$ be a function.
- The B-differential (B for Bouligand) of H at $x \in \mathbb{E}$ is denoted and defined by

$$\partial_B H(x) := \{J \in \mathcal{L}(\mathbb{E}, \mathbb{F}) : H'(x_k) \to J \text{ for a sequence } \{x_k\} \subseteq \mathcal{D}_H \text{ converging to } x\},$$

where $\mathcal{L}(\mathbb{E}, \mathbb{F})$ is the set of linear (continuous) maps from $\mathbb{E}$ to $\mathbb{F}$ and $\mathcal{D}_H$ is the set of points at which H is differentiable.



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

- The C-differential (C for Clarke [19]) of H at $x \in \mathbb{E}$ is denoted and defined by

$$\partial_C H(x) := \operatorname{co} \partial_B H(x),$$

where $\operatorname{co} S$ denotes the convex hull of a set S.

- H locally Lipschitz near $x \implies \partial_B H(x)$ and $\partial_C H(x)$ nonempty and bounded.

Semismoothness definition [61, 60; 1993]

- Let $\mathbb{E}$ and $\mathbb{F}$ be two normed spaces and Ω be an open set of $\mathbb{E}$.
- Let $H : \Omega \to \mathbb{F}$ be a function and $x \in \Omega$.
- The function H is said to be semismooth at x if the following three conditions hold:

  (SS1) H is Lipschitz near x,
  (SS2) H has directional derivatives at x in all directions,
  (SS3) when $h \to 0$ in $\mathbb{E}$, one has

$$\sup_{J \in \partial_C H(x+h)} \|H(x+h) - H(x) - Jh\| = o(\|h\|). \tag{1a}$$

- The function H is said to be strongly semismooth at x if it is semismooth at x with (SS3) strengthened into

  (SS3') for h near 0, one has

$$\sup_{J \in \partial_C H(x+h)} \|H(x+h) - H(x) - Jh\| = O(\|h\|^2). \tag{1b}$$

- The function $H : \Omega \to \mathbb{F}$ is said to be semismooth (resp. strongly semismooth) on a part P of Ω if it is semismooth (resp. strongly semismooth) at all points of P.

Semismoothness properties

- Semismooth Newton's method [61, 60; 1993]
  - Choose some nonsingular $J_k \in \partial_B H(x_k)$, if any,
  - $x_{k+1} := x_k - J_k^{-1} H(x_k)$.
- Local quadratic convergence of semismooth Newton's method if
  - x_0 is close to x_*,
  - H is strongly semismooth,
  - all $J \in \partial_B H(x_*)$ is nonsingular.
- Nice properties
  - H continuously differentiable at $x \Rightarrow H$ semismooth at x.
  - H_1 semismooth at x, H_2 semismooth at $H_1(x) \Rightarrow H_2 \circ H_1$ semismooth at x.
    - H_1, H_2 semismooth at $x \Rightarrow H_1 + H_2$ semismooth at x.
    - H_1, H_2 semismooth at $x \Leftrightarrow (H_1, H_2)$ semismooth at x.
    - H_1, H_2 semismooth at $x \Rightarrow \langle H_1, H_2 \rangle$ semismooth at x.
  - H_1, H_2 semismooth at $x \Rightarrow \min(H_1, H_2)$ semismooth at x.

Globalization of Newton's method for a nonsmooth function

No general technique.

Reason: $d_k = -J_k^{-1} H(x_k)$ may not be a descent direction of $\theta : x \mapsto \frac{1}{2}\|H(x)\|^2$. Often, it depends on the choice of $J_k \in \partial_B H(x_k)$.

Outline

Nonlinear complementarity problem

A complementarity problem consists in finding $x \in \Omega$ (open subset of $\mathbb{R}^n$) such that

$$F(x) \geqslant 0, \qquad G(x) \geqslant 0, \qquad \text{and} \qquad F(x)^{\mathsf{T}} G(x) = 0, \tag{2a}$$

where $F : \Omega \to \mathbb{R}^n$ and $G : \Omega \to \mathbb{R}^n$ are smooth. This is written compactly as follows:

$$\text{(NLCP)} \qquad 0 \leqslant F(x) \perp G(x) \geqslant 0. \tag{2b}$$

Linear complementarity problem

Sometimes, we shall refer to the *linear* complementarity problem [22]: this is (2) with $F(x) = Mx + q$ and $G(x) = x$:

$$\text{(LCP)} \qquad 0 \leqslant (Mx + q) \perp x \geqslant 0, \tag{3}$$

where $M \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$.

P-matrix

$$M \in \mathbf{P} \iff \det M_{I,I} > 0 \text{ for all } I \subseteq [1:n]$$
$$\iff \text{(3) has a unique solution for all } q \in \mathbb{R}^n.$$

We are interested in *linearization numerical methods* to solve these problems.

Nonlinear complementarity problem

A complementarity problem consists in finding $x \in \Omega$ (open subset of $\mathbb{R}^n$) such that

$$F(x) \geqslant 0, \qquad G(x) \geqslant 0, \qquad \text{and} \qquad F(x)^{\mathsf{T}} G(x) = 0, \tag{2a}$$

where $F : \Omega \to \mathbb{R}^n$ and $G : \Omega \to \mathbb{R}^n$ are smooth. This is written compactly as follows:

$$\text{(NLCP)} \qquad 0 \leqslant F(x) \perp G(x) \geqslant 0. \tag{2b}$$

Linear complementarity problem

Sometimes, we shall refer to the *linear* complementarity problem [22]: this is (2) with $F(x) = Mx + q$ and $G(x) = x$:

$$\text{(LCP)} \qquad 0 \leqslant (Mx + q) \perp x \geqslant 0, \tag{3}$$

where $M \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$.

P-matrix

$$\begin{aligned} M \in \mathbf{P} \quad & \Longleftrightarrow \quad \det M_{I,I} > 0 \text{ for all } I \subseteq [1:n] \\ & \Longleftrightarrow \quad \text{(3) has a unique solution for all } q \in \mathbb{R}^n. \end{aligned}$$

We are interested in *linearization numerical methods* to solve these problems.

Comments on the problem

- It is a set of nonlinear inequalities and one equation, so it may look like an easy problem to solve.

- Mangasarian-Fromovitz does not hold $\Longrightarrow$ instability for small perturbations.

- By the inequalities $F(x) \geqslant 0$ and $G(x) \geqslant 0$, the equation $F(x)^{\mathsf{T}} G(x) = 0$ also reads

$$\forall\, i \in [1:n]: \quad F_i(x) G_i(x) = 0.$$

  There are 2^n ways of realizing these complementarity conditions. Hence a huge combinatorial aspect.

- Even the LCP (3) is NP-hard in general [18, 47]. Depends on M:
  - at most n iterations if M is an **M**-matrix (Newton-min) [2],
  - ??? if M is a **P**-matrix (Lemke exponential [54], Newton-min cycles [9, 10, 11]),
  - ??? if M is a nondegenerate matrix,
  - NP-hard if M is a $\mathbf{P}_0$-matrix [47],
  - $O((1+\kappa)n^\alpha \log \varepsilon^{-1})$ iterations if M is a $\mathbf{P}_*(\kappa)$-matrix (interior points) [47, 59], but κ may be exponential in the length L of the data [24].

Comments on the problem

- It is a set of nonlinear inequalities and one equation, so it may look like an easy problem to solve.

- Mangasarian-Fromovitz does not hold $\Longrightarrow$ instability for small perturbations.

- By the inequalities $F(x) \geqslant 0$ and $G(x) \geqslant 0$, the equation $F(x)^\mathsf{T} G(x) = 0$ also reads

$$\forall\, i \in [1:n]: \quad F_i(x) G_i(x) = 0.$$

  There are 2^n ways of realizing these complementarity conditions. Hence a huge combinatorial aspect.

- Even the LCP (3) is NP-hard in general [18, 47]. Depends on M:
  - at most n iterations if M is an **M**-matrix (Newton-min) [2],
  - ??? if M is a **P**-matrix (Lemke exponential [54], Newton-min cycles [9, 10, 11]),
  - ??? if M is a nondegenerate matrix,
  - NP-hard if M is a $\mathbf{P}_0$-matrix [47],
  - $O((1+\kappa)n^\alpha \log \varepsilon^{-1})$ iterations if M is a $\mathbf{P}_*(\kappa)$-matrix (interior points) [47, 59], but κ may be exponential in the length L of the data [24].

Comments on the problem

- It is a set of nonlinear inequalities and one equation, so it may look like an easy problem to solve.
- Mangasarian-Fromovitz does not hold $\Longrightarrow$ instability for small perturbations.
- By the inequalities $F(x) \geqslant 0$ and $G(x) \geqslant 0$, the equation $F(x)^{\mathsf{T}} G(x) = 0$ also reads

$$\forall\, i \in [1:n]: \quad F_i(x) G_i(x) = 0.$$

  There are 2^n ways of realizing these complementarity conditions. Hence a huge combinatorial aspect.

- Even the LCP (3) is NP-hard in general [18, 47]. Depends on M:
  - at most n iterations if M is an **M**-matrix (Newton-min) [2],
  - ??? if M is a **P**-matrix (Lemke exponential [54], Newton-min cycles [9, 10, 11]),
  - ??? if M is a nondegenerate matrix,
  - NP-hard if M is a $\mathbf{P}_0$-matrix [47],
  - $O((1+\kappa)n^{\alpha} \log \varepsilon^{-1})$ iterations if M is a $\mathbf{P}_*(\kappa)$-matrix (interior points) [47, 59], but κ may be exponential in the length L of the data [24].

Comments on the problem

- It is a set of nonlinear inequalities and one equation, so it may look like an easy problem to solve.
- Mangasarian-Fromovitz does not hold $\Longrightarrow$ instability for small perturbations.
- By the inequalities $F(x) \geqslant 0$ and $G(x) \geqslant 0$, the equation $F(x)^{\mathsf{T}} G(x) = 0$ also reads

$$\forall\, i \in [1:n]: \quad F_i(x) G_i(x) = 0.$$

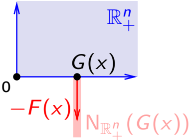
  There are 2^n ways of realizing these complementarity conditions. Hence a huge combinatorial aspect.

- Even the LCP (3) is NP-hard in general [18, 47]. Depends on M:
  - at most n iterations if M is an **M**-matrix (Newton-min) [2],
  - ??? if M is a **P**-matrix (Lemke exponential [54], Newton-min cycles [9, 10, 11]),
  - ??? if M is a nondegenerate matrix,
  - NP-hard if M is a $\mathbf{P}_0$-matrix [47],
  - $O((1+\kappa)n^\alpha \log \varepsilon^{-1})$ iterations if M is a $\mathbf{P}_*(\kappa)$-matrix (interior points) [47, 59], but κ may be exponential in the length L of the data [24].

Link with other problems

- It is a particular case of functional inclusion problem

$$F(x) + (\mathrm{N}_{\mathbb{R}^n_+} \circ G)(x) \ni 0.$$



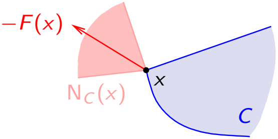
- First order optimality conditions of the optimization problem "$\min\{f(x) : c(x) \leqslant 0\}$":

$$\text{Find } (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m \text{ s.t. } \left\{ \begin{array}{ll} \nabla f(x) + c'(x)^{\mathsf{T}} \lambda = 0 & (n \text{ equations}) \\ 0 \leqslant \lambda \perp -c(x) \geqslant 0 & (m \text{ "conditions"}). \end{array} \right. \quad (4)$$

- The LCP was introduced and analyzed in the linear case by Cottle in his PhD thesis [20, 21; 1964], as an extension of the linear optimization problem.

- The related variational inequality problem

$$\left\{ \begin{array}{l} x \in C \text{ (a convex set)} \\ \langle F(x), y - x \rangle \geqslant 0, \quad \forall y \in C. \end{array} \right.$$



was introduced by Hartman and Stampacchia [44; 1966] for an EDP.

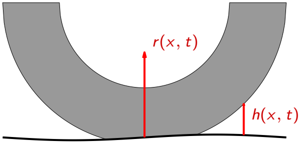
Examples of use

- **General principle**. Useful for systems in competition with threshold effects:

$$\text{If the threshold } F_i(x) \text{ is inactive } (> 0) \quad \Longrightarrow \quad G_i(x) = 0.$$

- Examples in
  - nonsmooth mechanics and dynamics, contact problems [1, 14, 3],

Tire/road contact
in (space,time)



$$\begin{cases} r(x,t) \geqslant 0 \\ h(x,t) \geqslant 0 \\ r(x,t)h(x,t) = 0. \end{cases}$$

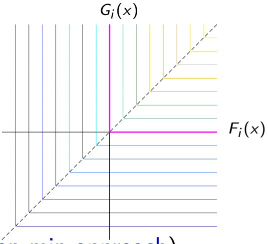
  - phase transition problem in multiphase flows [52, 53, 7, 4, 6, 5, 16, 23],
  - precipitation-dissolution problems in chemistry [15, 48],
  - portfolio management in finance [41],
  - computer graphics [31],
  - free boundary problems, meteorology simulation, economic equilibrium, ...

- More examples of applications in [42, 45, 57, 37, 32].

- Pivoting (Lemke) for LCP.
- Interior points.
- Nonsmooth equation reformulation and pseudo-linearization. $\longleftarrow$
- Smoothing nonsmooth reformulations.
- Other methods . . .

Outline

Equation reformulation of NLCP (I)



Solve the following nonsmooth reformulation of (NLCP):

$$H(x) = 0, \qquad (5a)$$

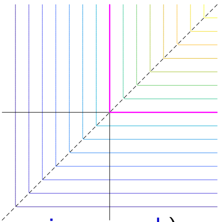
where $H : \Omega \to \mathbb{R}^n$ is the function defined at $x \in \Omega$ by

$$H(x) := \min(F(x), G(x)). \qquad (5b)$$

Compute a direction d by a pseudo-linearization of H ($\equiv$ Newton-min approach).

- (5) is equivalent to (NLCP) since $\min(a, b) = 0$ iff $a \geqslant 0$, $b \geqslant 0$ and $ab = 0$.
- H has directional derivatives and is semismooth (if F and G are smooth).
- There are other equation reformulations, like the one using the Fischer function $\varphi_{\mathrm{F}}(a, b) = \sqrt{a^2 + b^2} - (a + b)$ [38, 34, 51, 25, 58].
- The function "min" reformulation is a choice guided by
  - scientific curiosity (there are still possibilities of improvement),
  - efficiency of the approach ("min" is more linear, although less differentiable than φ_{F}),
  - can give better local convergence result than with φ_{F} [32],
  - can give finite termination for LCP [39].

Equation reformulation of NLCP (I)



Solve the following nonsmooth reformulation of (NLCP):

$$H(x) = 0, \tag{5a}$$

where $H : \Omega \to \mathbb{R}^n$ is the function defined at $x \in \Omega$ by

$$H(x) := \min(F(x), G(x)). \tag{5b}$$

Compute a direction d by a pseudo-linearization of H ($\equiv$ Newton-min approach).

- (5) is equivalent to (NLCP) since $\min(a, b) = 0$ iff $a \geqslant 0$, $b \geqslant 0$ and $ab = 0$.
- H has directional derivatives and is semismooth (if F and G are smooth).
- There are other equation reformulations, like the one using the Fischer function $\varphi_{\mathrm{F}}(a, b) = \sqrt{a^2 + b^2} - (a + b)$ [38, 34, 51, 25, 58].
- The function "min" reformulation is a choice guided by
  - scientific curiosity (there are still possibilities of improvement),
  - efficiency of the approach ("min" is more linear, although less differentiable than φ_{F}),
  - can give better local convergence result than with φ_{F} [32],
  - can give finite termination for LCP [39].

Equation reformulation of NLCP (II): globalization [12, 13]

The quadratic merit function associated with (5) is defined at $x \in \mathbb{R}^n$ by

$$\theta(x) := \frac{1}{2}\|H(x)\|^2 = \frac{1}{2}\|\min(F(x), G(x))\|^2. \tag{6}$$

- θ has directional derivatives and is semismooth.
- Algorithmic goal

Algorithm

- *Compute $d \in \mathbb{R}^n$ such that*
  - *it is a descent direction of θ, i.e., $\theta'(x; d) < 0$,*
  - *it is efficient locally (quadratic or finite convergence).*
- *Do a standard Armijo line-search on θ: find a not too small $\alpha > 0$ such that ($\omega \in (0, 1)$)*

$$\theta(x + \alpha d) \leqslant \theta(x) + \omega\alpha\theta'(x; d).$$

- *Update $x_+ = x + \alpha d$.*

- Certify the algorithm by some kind of global convergence.

Josephy-Newton (JN) method

For a function Φ and a multifunction N, the JN algorithm [46] aims at solving

$$\Phi(x) + N(x) \ni 0,$$

by linearizing Φ, while keeping N unchanged. Hence $x_+ = x + d$, where d solves

$$\Phi(x) + \Phi'(x)d + N(x+d) \ni 0.$$

Applied to the NLCP "$0 \leqslant F(x) \perp G(x) \geqslant 0$" $\iff$ "$F(x) + (\mathrm{N}_{\mathbb{R}^n_+} \circ G)(x) \ni 0$", it computes $x_+ = x + d$ where d solves

$$\text{(JN)} \qquad 0 \leqslant \Big(F(x) + F'(x)d\Big) \perp \Big(G(x) + G'(x)d\Big) \geqslant 0.$$

Properties (similar to those of the SQP algorithm in constrained optimization):

- ⊕ fast local convergence (quadratic) with realistic assumptions,
- ⊕ yields descent directions of the quadratic merit function θ,
- ⊕ global convergence,
- ⊖ expensive iteration (one LCP to solve),
- ⊖ makes no sense for solving the LCP, since (JN) $\equiv$ (LCP).

Josephy-Newton (JN) method

For a function Φ and a multifunction N, the JN algorithm [46] aims at solving

$$\Phi(x) + N(x) \ni 0,$$

by linearizing Φ, while keeping N unchanged. Hence $x_+ = x + d$, where d solves

$$\Phi(x) + \Phi'(x)d + N(x+d) \ni 0.$$

Applied to the NLCP "$0 \leqslant F(x) \perp G(x) \geqslant 0$" $\iff$ "$F(x) + (\mathrm{N}_{\mathbb{R}^n_+} \circ G)(x) \ni 0$", it computes $x_+ = x + d$ where d solves

$$\text{(JN)} \qquad 0 \leqslant \Big(F(x) + F'(x)d\Big) \perp \Big(G(x) + G'(x)d\Big) \geqslant 0.$$

Properties (similar to those of the SQP algorithm in constrained optimization):

- $\oplus$ fast local convergence (quadratic) with realistic assumptions,
- $\oplus$ yields descent directions of the quadratic merit function θ,
- $\oplus$ global convergence,
- $\ominus$ expensive iteration (one LCP to solve),
- $\ominus$ makes no sense for solving the LCP, since (JN) $\equiv$ (LCP).

B-Newton method

For a locally Lipschitz function H, the B-Newton algorithm [55] aims at solving $H(x) = 0$ by taking $x_+ = x + d$, where d solves

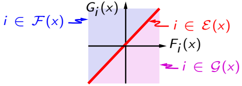
$$H(x) + H'(x; d) = 0.$$

Applied to the NLCP [55, 56] and $H = \min(F, G)$, it computes $x_+ = x + d$ where d solves

$$\text{(BN)} \qquad \begin{cases} (F(x) + F'(x)d)_{\mathcal{F}(x)} = 0, \\ (G(x) + G'(x)d)_{\mathcal{G}(x)} = 0, \\ 0 \leqslant (F(x) + F'(x)d)_{\mathcal{E}(x)} \perp (G(x) + G'(x)d)_{\mathcal{E}(x)} \geqslant 0, \end{cases}$$

where

$$\begin{aligned} \mathcal{E}(x) &:= \{i \in [1:n] : F_i(x) = G_i(x)\}, \\ \mathcal{F}(x) &:= \{i \in [1:n] : F_i(x) < G_i(x)\}, \\ \mathcal{G}(x) &:= \{i \in [1:n] : F_i(x) > G_i(x)\}. \end{aligned}$$



Properties:

- ⊕ yields descent directions of the quadratic merit function θ,
- ⊕ global convergence,
- ⊖ a limit point $\bar{x}$ is a solution if it is "regular" and satisfies $F_i(\bar{x}) = G_i(\bar{x}) = 0$ for $i \in \mathcal{E}(\bar{x})$,
- ⊖ much less expensive iteration than JN ($|\mathcal{E}(x)|$ small), but still one LCP to solve,
- ⊖ makes no sense for solving the LCP, since (BN) ≡ (JN) when $\mathcal{E}(x) = [1:n]$.

B-Newton method

For a locally Lipschitz function H, the B-Newton algorithm [55] aims at solving $H(x) = 0$ by taking $x_+ = x + d$, where d solves

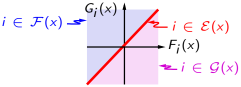
$$H(x) + H'(x; d) = 0.$$

Applied to the NLCP [55, 56] and $H = \min(F, G)$, it computes $x_+ = x + d$ where d solves

$$(\text{BN}) \qquad \begin{cases} (F(x) + F'(x)d)_{\mathcal{F}(x)} = 0, \\ (G(x) + G'(x)d)_{\mathcal{G}(x)} = 0, \\ 0 \leqslant (F(x) + F'(x)d)_{\mathcal{E}(x)} \perp (G(x) + G'(x)d)_{\mathcal{E}(x)} \geqslant 0, \end{cases}$$

where

$$\begin{aligned} \mathcal{E}(x) &:= \{i \in [1:n] : F_i(x) = G_i(x)\}, \\ \mathcal{F}(x) &:= \{i \in [1:n] : F_i(x) < G_i(x)\}, \\ \mathcal{G}(x) &:= \{i \in [1:n] : F_i(x) > G_i(x)\}. \end{aligned}$$



Properties:

- ⊕ yields descent directions of the quadratic merit function θ,
- ⊕ global convergence,
- ⊖ a limit point $\bar{x}$ is a solution if it is "regular" and satisfies $F_i(\bar{x}) = G_i(\bar{x}) = 0$ for $i \in \mathcal{E}(\bar{x})$,
- ⊖ much less expensive iteration than JN ($|\mathcal{E}(x)|$ small), but still one LCP to solve,
- ⊖ makes no sense for solving the LCP, since (BN) $\equiv$ (JN) when $\mathcal{E}(x) = [1:n]$.

Semismooth Newton method

- Algorithm for solving $H(x) := \min(F(x), G(x)) = 0$
  - Choose a nonsingular Jacobian

$$J \in \partial_B H(x) \subseteq \partial_B H_1(x) \times \cdots \times \partial_B H_n(x) =: \partial_B^{\times} H(x) \quad \text{or}$$
$$J \in \partial_C H(x) \subseteq \partial_C H_1(x) \times \cdots \times \partial_C H_n(x) =: \partial_C^{\times} H(x).$$

  - Determine d by $H(x) + Jd = 0$.
  - If d is descent direction of θ, do a LS along d to get $x_+ := x + \alpha d$.
- Discussion
  - Define the piecewise affine model $\mathcal{L}_x H$ of H at $x \in \mathbb{R}^n$ by

$$y \in \mathbb{R}^n \mapsto (\mathcal{L}_x H)(y) := \min\big(F(x) + F'(x)(y-x), G(x) + G'(x)(y-x)\big).$$

    Then,

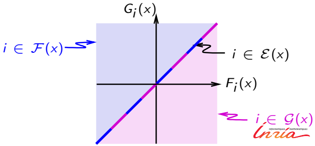
$$\partial_B(\mathcal{L}_x H)(x) \subseteq \partial_B H(x) \qquad \text{and} \qquad \partial_C(\mathcal{L}_x H)(x) \subseteq \partial_C H(x).$$

  - Computing a single Jacobian J of $\partial_B(\mathcal{L}_x H)(x)$, hence of $\partial_B H(x)$, is easy (all the Jacobians is difficult) [29]. Same observation for ∂_C.
  - Having J nonsingular is a matter of assumption (not guaranteed in general).
  - But d is not necessarily a descent direction of θ (a counter-example in a while).

Plain Newton-min method

- Algorithm for solving $H(x) := \min(F(x), G(x)) = 0$
  - Choose a nonsingular Jacobian
$$J \in \partial_B H_1(x) \times \cdots \times \partial_B H_n(x) =: \partial_B^{\times} H(x) \quad \text{or}$$
$$J \in \partial_C H_1(x) \times \cdots \times \partial_C H_n(x) =: \partial_C^{\times} H(x).$$
  - Determine d by $H(x) + Jd = 0$.
  - If d is descent direction of θ, do a LS along d to get $x_+ := x + \alpha d$.

- Discussion
  - For $i \in [1:n]$, one has
$$\partial_B H_i(x) = \begin{cases} \{F_i'(x)\} & \text{if } F_i(x) < G_i(x) \;\Leftrightarrow\; i \in \mathcal{F}(x), \\ \{F_i'(x), G_i'(x)\} & \text{if } F_i(x) = G_i(x) \;\Leftrightarrow\; i \in \mathcal{E}(x), \\ \{G_i'(x)\} & \text{if } F_i(x) > G_i(x) \;\Leftrightarrow\; i \in \mathcal{G}(x). \end{cases}$$
  - Hence d with $J \in \partial_B^{\times} H(x)$ is defined by
$$\begin{cases} F_i(x) + F_i'(x)d = 0 & \text{if } i \in \tilde{\mathcal{F}}(x) \\ G_i(x) + G_i'(x)d = 0 & \text{if } i \in \tilde{\mathcal{G}}(x), \end{cases} \qquad (7)$$
    where $(\tilde{\mathcal{F}}(x), \tilde{\mathcal{G}}(x))$ forms a partition of $[1:n]$ with $\tilde{\mathcal{F}}(x) \supseteq \mathcal{F}(x)$ and $\tilde{\mathcal{G}}(x) \supseteq \mathcal{G}(x)$.



The (semismooth Newton/Newton-min) direction can be an ascent direction for θ

Consider the LCP (3), which is $0 \leqslant x \perp (Mx + q) \geqslant 0$, with

$$M = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}, \qquad q = \begin{pmatrix} -4 \\ -2 \end{pmatrix}, \quad x = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad \text{so that} \quad Mx + q = \begin{pmatrix} -2 \\ -1 \end{pmatrix}. \qquad (8)$$

One has $\mathcal{E}(x) = \{1\}$, $\mathcal{F}(x) = \{2\}$, $\mathcal{G}(x) = \varnothing$.

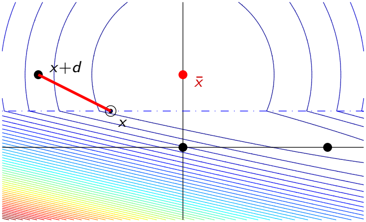
The (semismooth Newton/Newton-min) direction can be an ascent direction for θ

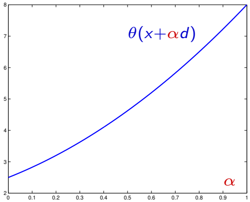
Consider the LCP (3), which is $0 \leqslant x \perp (Mx+q) \geqslant 0$, with

$$M = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}, \qquad q = \begin{pmatrix} -4 \\ -2 \end{pmatrix}, \quad x = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad \text{so that} \quad Mx+q = \begin{pmatrix} -2 \\ -1 \end{pmatrix}. \qquad (8)$$

One has $\mathcal{E}(x) = \{1\}$, $\mathcal{F}(x) = \{2\}$, $\mathcal{G}(x) = \varnothing$.

Take $\tilde{\mathcal{F}}(x) = \{1, 2\}$ and $\tilde{\mathcal{G}}(x) = \varnothing$ in (7), then d is an ascent direction of θ at x:





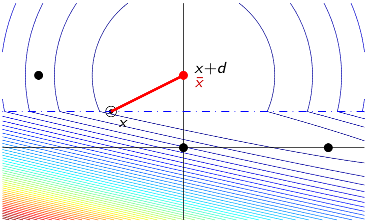
The (semismooth Newton/Newton-min) direction can be an ascent direction for θ

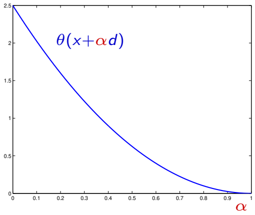
Consider the LCP (3), which is $0 \leqslant x \perp (Mx + q) \geqslant 0$, with

$$M = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}, \qquad q = \begin{pmatrix} -4 \\ -2 \end{pmatrix}, \quad x = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad \text{so that} \quad Mx + q = \begin{pmatrix} -2 \\ -1 \end{pmatrix}. \qquad (8)$$

One has $\mathcal{E}(x) = \{1\}$, $\mathcal{F}(x) = \{2\}$, $\mathcal{G}(x) = \varnothing$.

Take $\tilde{\mathcal{F}}(x) = \{2\}$ and $\tilde{\mathcal{G}}(x) = \{1\}$ in (7), then d is a descent direction of θ at x:





Outline

Orientation

Slightly modify the plain Newton-min direction such that:

- ⊕⊖ it computes a point in a convex polyhedron (harder than a LS, easier than an LCP):
  - ⊕ very few inequalities define the convex polyhedron,
  - ⊖ the computation of d is more expensive, but polynomial,
  - ⊕ there is a bypass that accepts the plain NM direction most of the iterations,
- ⊕ it becomes a descent direction of θ,
- ⊕ it yields some global convergence.

Ensuring descent

For the quadratic merit function $\theta(x) = \frac{1}{2}\|H(x)\|^2 = \frac{1}{2}\|\min(F(x), G(x))\|^2$, one has

$$\begin{aligned}\theta'(x;d) &= H(x)^\mathsf{T} H'(x;d)\\ &= F_{\mathcal{F}(x)}(x)^\mathsf{T} F'_{\mathcal{F}(x)}(x)d + G_{\mathcal{G}(x)}(x)^\mathsf{T} G'_{\mathcal{G}(x)}(x)d + F_{\mathcal{E}(x)}(x)^\mathsf{T} \min(F'_{\mathcal{E}(x)}(x)d, G'_{\mathcal{E}(x)}(x)d).\end{aligned}$$

If $(F(x) + F'(x)d)_{\mathcal{F}(x)} = 0$ and $(G(x) + G'(x)d)_{\mathcal{G}(x)} = 0$, it follows

$$\begin{aligned}\theta'(x;d) =& -\|F_{\mathcal{F}(x)}(x)\|^2 - \|G_{\mathcal{G}(x)}(x)\|^2 - \|F_{\mathcal{E}(x)}(x)\|^2\\ &+ F_{\mathcal{E}(x)}(x)^\mathsf{T} \min(F_{\mathcal{E}(x)}(x) + F'_{\mathcal{E}(x)}(x)d, G_{\mathcal{E}(x)}(x) + G'_{\mathcal{E}(x)}(x)d)\\ =& -2\,\theta(x) + F_{\mathcal{E}(x)}(x)^\mathsf{T} \min(F_{\mathcal{E}(x)}(x) + F'_{\mathcal{E}(x)}(x)d, G_{\mathcal{E}(x)}(x) + G'_{\mathcal{E}(x)}(x)d).\end{aligned}$$

How can we get $\theta'(x;d) < 0$ when $\theta(x) \neq 0$?

- If $F_i(x) = G_i(x) \geqslant 0$, the last term is $\leqslant 0$ when
  $$F_i(x) + F_i'(x)d = 0 \quad \text{or} \quad G_i(x) + G_i'(x)d = 0.$$
- If $F_i(x) = G_i(x) < 0$, the last term is $\leqslant 0$ when
  $$F_i(x) + F_i'(x)d \geqslant 0 \quad \text{and} \quad G_i(x) + G_i'(x)d \geqslant 0.$$

This leads to the following direction definition.

Ensuring descent

For the quadratic merit function $\theta(x) = \frac{1}{2}\|H(x)\|^2 = \frac{1}{2}\|\min(F(x), G(x))\|^2$, one has

$$\begin{aligned}\theta'(x;d) &= H(x)^{\mathsf{T}} H'(x;d)\\ &= F_{\mathcal{F}(x)}(x)^{\mathsf{T}} F'_{\mathcal{F}(x)}(x)d + G_{\mathcal{G}(x)}(x)^{\mathsf{T}} G'_{\mathcal{G}(x)}(x)d + F_{\mathcal{E}(x)}(x)^{\mathsf{T}} \min(F'_{\mathcal{E}(x)}(x)d, G'_{\mathcal{E}(x)}(x)d).\end{aligned}$$

If $(F(x) + F'(x)d)_{\mathcal{F}(x)} = 0$ and $(G(x) + G'(x)d)_{\mathcal{G}(x)} = 0$, it follows

$$\begin{aligned}\theta'(x;d) &= -\|F_{\mathcal{F}(x)}(x)\|^2 - \|G_{\mathcal{G}(x)}(x)\|^2 - \|F_{\mathcal{E}(x)}(x)\|^2\\ &\quad + F_{\mathcal{E}(x)}(x)^{\mathsf{T}} \min(F_{\mathcal{E}(x)}(x) + F'_{\mathcal{E}(x)}(x)d, G_{\mathcal{E}(x)}(x) + G'_{\mathcal{E}(x)}(x)d)\\ &= -2\,\theta(x) + F_{\mathcal{E}(x)}(x)^{\mathsf{T}} \min(F_{\mathcal{E}(x)}(x) + F'_{\mathcal{E}(x)}(x)d, G_{\mathcal{E}(x)}(x) + G'_{\mathcal{E}(x)}(x)d).\end{aligned}$$

How can we get $\theta'(x;d) < 0$ when $\theta(x) \neq 0$?

- If $F_i(x) = G_i(x) \geqslant 0$, the last term is $\leqslant 0$ when

$$F_i(x) + F_i'(x)d = 0 \qquad \text{or} \qquad G_i(x) + G_i'(x)d = 0.$$

- If $F_i(x) = G_i(x) < 0$, the last term is $\leqslant 0$ when

$$F_i(x) + F_i'(x)d \geqslant 0 \qquad \text{and} \qquad G_i(x) + G_i'(x)d \geqslant 0.$$

This leads to the following direction definition.

Ensuring descent

For the quadratic merit function $\theta(x) = \frac{1}{2}\|H(x)\|^2 = \frac{1}{2}\|\min(F(x), G(x))\|^2$, one has

$$\begin{aligned}\theta'(x; d) &= H(x)^{\mathsf{T}} H'(x; d)\\ &= F_{\mathcal{F}(x)}(x)^{\mathsf{T}} F'_{\mathcal{F}(x)}(x)d + G_{\mathcal{G}(x)}(x)^{\mathsf{T}} G'_{\mathcal{G}(x)}(x)d + F_{\mathcal{E}(x)}(x)^{\mathsf{T}} \min(F'_{\mathcal{E}(x)}(x)d, G'_{\mathcal{E}(x)}(x)d).\end{aligned}$$

If $(F(x) + F'(x)d)_{\mathcal{F}(x)} = 0$ and $(G(x) + G'(x)d)_{\mathcal{G}(x)} = 0$, it follows

$$\begin{aligned}\theta'(x; d) &= -\|F_{\mathcal{F}(x)}(x)\|^2 - \|G_{\mathcal{G}(x)}(x)\|^2 - \|F_{\mathcal{E}(x)}(x)\|^2\\ &\quad + F_{\mathcal{E}(x)}(x)^{\mathsf{T}} \min(F_{\mathcal{E}(x)}(x) + F'_{\mathcal{E}(x)}(x)d, G_{\mathcal{E}(x)}(x) + G'_{\mathcal{E}(x)}(x)d)\\ &= -2\,\theta(x) + F_{\mathcal{E}(x)}(x)^{\mathsf{T}} \min(F_{\mathcal{E}(x)}(x) + F'_{\mathcal{E}(x)}(x)d, G_{\mathcal{E}(x)}(x) + G'_{\mathcal{E}(x)}(x)d).\end{aligned}$$

How can we get $\theta'(x; d) < 0$ when $\theta(x) \neq 0$?

- If $F_i(x) = G_i(x) \geqslant 0$, the last term is $\leqslant 0$ when

$$F_i(x) + F_i'(x)d = 0 \qquad \text{or} \qquad G_i(x) + G_i'(x)d = 0.$$

- If $F_i(x) = G_i(x) < 0$, the last term is $\leqslant 0$ when

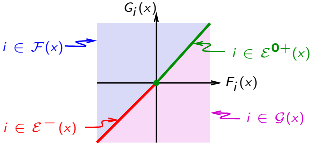
$$F_i(x) + F_i'(x)d \geqslant 0 \qquad \text{and} \qquad G_i(x) + G_i'(x)d \geqslant 0.$$

This leads to the following direction definition.

Plain polyhedral Newton-min direction

A *plain polyhedral Newton-min* (plain PNM) *direction* is a direction d that satisfies

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



where $(\mathcal{E}_{\mathcal{F}}^{0+}(x), \mathcal{E}_{\mathcal{G}}^{0+}(x))$ is a partition of

$$\mathcal{E}^{0+}(x) := \{i \in [1:n] : F_i(x) = G_i(x) \geqslant 0\}$$

and

$$\mathcal{E}^{-}(x) := \{i \in [1:n] : F_i(x) = G_i(x) < 0\}.$$

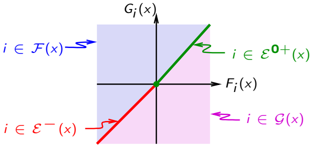
Features of the algorithm:

- ⊖ d must be found in a convex polyhedron (instead of the solution to a LS),
- ⊕ the number of inequalities $2|\mathcal{E}^{-}(x)|$ should be very small (in exact arithmetic!),
- ⊕ can be computed in polynomial time (by LO or QO),
- ⊕ there is a bypass to avoid this computation most of the time (see below),
- ⊕ d is a descent direction of θ,
- ⊖ we were not able to prove global convergence with that d.

Plain polyhedral Newton-min direction

A *plain polyhedral Newton-min* (plain PNM) *direction* is a direction d that satisfies

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



where $(\mathcal{E}^{0+}_{\mathcal{F}}(x), \mathcal{E}^{0+}_{\mathcal{G}}(x))$ is a partition of

$$\mathcal{E}^{0+}(x) := \{i \in [1:n] : F_i(x) = G_i(x) \geqslant 0\}$$

and

$$\mathcal{E}^{-}(x) := \{i \in [1:n] : F_i(x) = G_i(x) < 0\}.$$

Features of the algorithm:

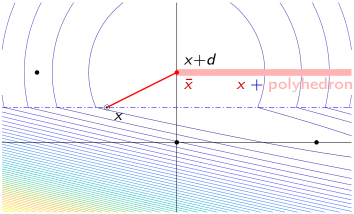
- ⊖ d must be found in a convex polyhedron (instead of the solution to a LS),
- ⊕ the number of inequalities $2|\mathcal{E}^{-}(x)|$ should be very small (in exact arithmetic!),
- ⊕ can be computed in polynomial time (by LO or QO),
- ⊕ there is a bypass to avoid this computation most of the time (see below),
- ⊕ d is a descent direction of θ,
- ⊖ we were not able to prove global convergence with that d.

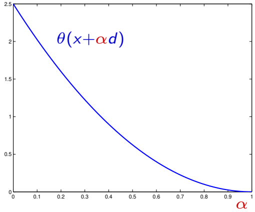
Behavior on the baby problem (8)

Since $\mathcal{E}(x) = \{1\}$, $\mathcal{F}(x) = \{2\}$, $\mathcal{G}(x) = \varnothing$, the algorithm computes the solution to

$$\begin{cases} \min \ \frac{1}{2}\,\|d\|_2^2 \\ M_{2:}\,d + y_2 = 0 \\ M_{1:}\,d + y_1 \geqslant 0 \\ d_1 + x_1 \geqslant 0 \end{cases} \qquad \text{or} \qquad \begin{cases} \min \ \frac{1}{2}\,(d_1^2 + 1) \\ d_1 \geqslant 2, \end{cases} \quad d_2 = 1.$$

A little by chance, it is the right direction $d = (2, 1)$.





Difficulty with global convergence

Let $\bar{x}$ be an accumulation point of the sequence $\{x_k\}_{k\geqslant 1}$ (it may not exist) generated by

$$x_{k+1} := x_k + \alpha_k d_k$$

where $\alpha_k > 0$ is the largest stepsize of the form 2^{-i} for $i \in \mathbb{N}$ such that

$$\theta(x_k + \alpha_k d_k) \leqslant \theta(x_k) + 10^{-4}\alpha_k(\text{"sth negative"}). \tag{9a}$$

We want to show that $\bar{x}$ is a solution of the NLCP (with a regularity assumption).

- If $\limsup_k \alpha_k > 0$, it is easy to show that $\theta(x_k) \downarrow 0$ and that $\bar{x}$ is a solution.
- If $\limsup_k \alpha_k = 0$, it is more difficult.

  Necessarily (9a) is not satisfied for $\check{\alpha}_k = 2\alpha_k$:

$$\theta(x_k + \check{\alpha}_k d_k) > \theta(x_k) + 10^{-4}\check{\alpha}_k(\text{"sth negative"}). \tag{9b}$$

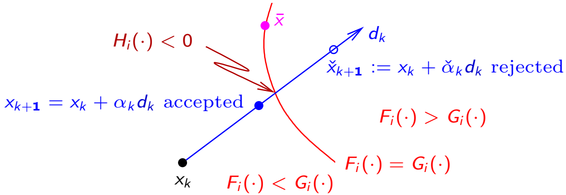
  To get convergence, it is necessary to get information from both (9a) and (9b).

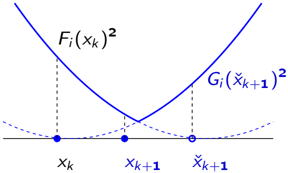
Difficulty with global convergence (negative kink)

- Near a negative kink, one can have with $\check{x}_{k+1} := x_k + \check{\alpha}_k d_k$:

$$F_i(x_{k+1}) < G_i(x_{k+1}) < 0, \qquad 0 > F_i(\check{x}_{k+1}) > G_i(\check{x}_{k+1}),$$

$$0 < H_i(x_{k+1})^2 = F_i(x_{k+1})^2, \qquad 0 < H_i(\check{x}_{k+1})^2 = G_i(\check{x}_{k+1})^2 > F_i(\check{x}_{k+1})^2.$$





- Hence $\check{x}_{k+1}$ is rejected because of $G_i(\check{x}_{k+1})^2$, but one has no information on $G_i(x_k) + G_i'(x_k)d_k$.

- Remedy: for x_k near a <u>negative</u> kink of H,

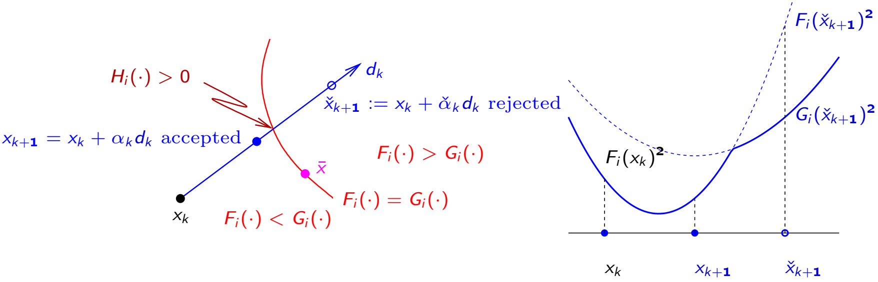
$$F_i(x_k) + F_i'(x_k)d_k = 0 \quad \curvearrowright \quad \begin{cases} F_i(x_k) + F_i'(x_k)d_k \geqslant 0 \\ G_i(x_k) + G_i'(x_k)d_k \geqslant 0. \end{cases}$$

Difficulty with global convergence (positive kink)

- Near a positive kink, one can have with $\check{x}_{k+1} := x_k + \check{\alpha}_k d_k$:

$$0 < F_i(x_{k+1}) < G_i(x_{k+1}), \qquad F_i(\check{x}_{k+1}) > G_i(\check{x}_{k+1}) > 0,$$

$$0 < H_i(x_{k+1})^2 = F_i(x_{k+1})^2, \qquad 0 < H_i(\check{x}_{k+1})^2 = G_i(\check{x}_{k+1})^2 < F_i(\check{x}_{k+1})^2.$$



- Hence $\check{x}_{k+1}$ is rejected because of $G_i(\check{x}_{k+1})^2$ and would also be rejected because of $F_i(\check{x}_{k+1})^2$.

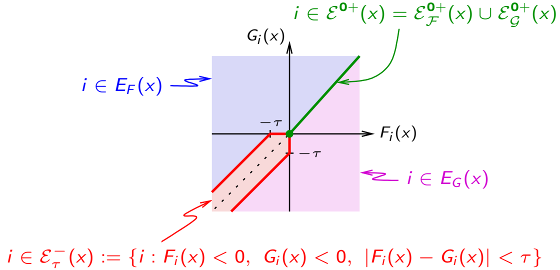
- Since we have information on $F_i(x_k) + F_i'(x_k)d_k = 0$, there's no need for a remedy.

Secure polyhedral Newton-min algorithm

A *secure polyhedral Newton-min* (PNM) *direction* is a direction d satisfying

$$\left\{\begin{array}{ll} F_i(x)+F_i'(x)d=0 & \text{if } i\in E_F(x):=\left[\mathcal{F}(x)\setminus\mathcal{E}_\tau^-(x)\right]\cup\mathcal{E}_{\mathcal{F}}^{\mathbf{0}+}(x)\\ G_i(x)+G_i'(x)d=0 & \text{if } i\in E_G(x):=\left[\mathcal{G}(x)\setminus\mathcal{E}_\tau^-(x)\right]\cup\mathcal{E}_{\mathcal{G}}^{\mathbf{0}+}(x)\\ F_i(x)+F_i'(x)d\geqslant 0 & \text{if } i\in I(x):=\mathcal{E}_\tau^-(x)\\ G_i(x)+G_i'(x)d\geqslant 0 & \text{if } i\in I(x):=\mathcal{E}_\tau^-(x), \end{array}\right. \tag{10}$$

where, for some kink tolerance parameter $\tau\in(0,\infty)$,



PNM regularity condition

- The usual regularity at a limit point $\bar{x}$ assumes that the system to solve has a solution, whatever the vectors defining it are.
- Here, there must be a d satisfying the system below, whatever $F_i(\bar{x})$, $G_i(\bar{x})$, $F_i(\bar{x})$, $G_i(\bar{x})$ are:

$$\begin{cases} F_i(\bar{x}) + F_i'(\bar{x})d = 0 & \text{if } i \in E_F(\bar{x}) \\ G_i(\bar{x}) + G_i'(\bar{x})d = 0 & \text{if } i \in E_G(\bar{x}) \\ F_i(\bar{x}) + F_i'(\bar{x})d \geqslant 0 & \text{if } i \in I(\bar{x}) \\ G_i(\bar{x}) + G_i'(\bar{x})d \geqslant 0 & \text{if } i \in I(\bar{x}). \end{cases}$$

- This is guaranteed by the Mangasarian-Fromovitz "constraint qualification" (MFCQ):

$$\textstyle\sum_{i\in E_F(\bar{x})} \alpha_i \nabla F_i(\bar{x}) + \sum_{i\in E_G(\bar{x})} \beta_i \nabla G_i(\bar{x}) + \sum_{i\in I(\bar{x})} \big[\alpha_i \nabla F_i(\bar{x}) + \beta_i \nabla G_i(\bar{x})\big] = 0$$

and $(\alpha_{I(\bar{x})}, \beta_{I(\bar{x})}) \geqslant 0$ imply that $(\alpha, \beta) = 0$.

- Must be reinforced to have a "diffusion property" near $\bar{x}$ (difficulty with the index sets that change with $\bar{x}$). This yields the PNM regularity. Ensures
  - existence of a d satisfying (10) for x near $\bar{x}$,
  - boundedness of the d's.

Features of the PNM algorithm:

- $\ominus$ d must be found in a convex polyhedron (instead of the solution to a LS),
- $\oplus$ the number of inequalities $2|\mathcal{E}_\tau^-(x)|$ should be very small ($\tau > 0$ can be very small),
- $\oplus$ can be computed in polynomial time (by LO or QO),
- $\oplus$ there is a bypass to avoid this computation most of the time (see below),
- $\oplus$ d is a descent direction of θ,
- $\oplus$ global convergence.

Theorem (global convergence of the PNM algorithm)

If
- *F and $G : \Omega \to \mathbb{R}^n$ are differentiable,*
- *the PNM algorithm generates a sequence $\{x_k\} \subseteq \Omega$,*
- *$\bar{x} \in \Omega$ is an accumulation point of $\{x_k\}$ that is PNM regular,*
- *F' and G' are continuous at $\bar{x}$,*

then, $\{\theta(x_k)\}_{k \geqslant 1} \downarrow 0$ and $\bar{x}$ is a solution to the NLCP (2).

Acceptation criterion (sufficient decrease condition)

One Looks for a criterion for accepting the cheap plain Newton-min direction (7).

- Newton direction for smooth H satisfies $\theta'(x;d) = -2\theta(x)$, hence requiring for some $\eta \in (0,1)$:

  $\theta'(x;d) \leqslant -2(1-\eta)\theta(x) \qquad \longrightarrow \qquad$ not strong enough to get global convergence.

- One requires instead, for some $\eta \in (0,1)$, close to 1:

$$\underbrace{-\sum_{i\in[1:n]} (1-\rho_i(x,d))\, H_i(x)^2}_{\text{upper bound on } \theta'(x;d)} \leqslant -2(1-\eta)\,\theta(x), \tag{11}$$

where

$$\rho_i(x,d) := \begin{cases} \frac{F_i(x)+F_i'(x)d}{F_i(x)} & \text{if } i \in E_F(x) \text{ and } F_i(x) \neq 0 \\ 0 & \text{if } i \in E_F(x) \text{ and } F_i(x) = 0 \\ \frac{G_i(x)+G_i'(x)d}{G_i(x)} & \text{if } i \in E_G(x) \text{ and } G_i(x) \neq 0 \\ 0 & \text{if } i \in E_G(x) \text{ and } G_i(x) = 0 \\ \max\left(\frac{F_i(x)+F_i'(x)d}{F_i(x)}, \frac{G_i(x)+G_i'(x)d}{G_i(x)}\right) & \text{if } i \in I(x), \end{cases}$$

Acceptation criterion (sufficient decrease condition)

One Looks for a criterion for accepting the cheap plain Newton-min direction (7).

- Newton direction for smooth H satisfies $\theta'(x;d) = -2\theta(x)$, hence requiring for some $\eta \in (0,1)$:

  $\theta'(x;d) \leqslant -2(1-\eta)\theta(x) \qquad \longrightarrow \qquad$ not strong enough to get global convergence.

- One requires instead, for some $\eta \in (0,1)$, close to 1:

$$\boxed{\underbrace{-\sum_{i\in[1:n]} (1-\rho_i(x,d))\, H_i(x)^2}_{\text{upper bound on } \theta'(x;d)} \leqslant -2(1-\eta)\,\theta(x),} \tag{11}$$

where

$$\rho_i(x,d) := \begin{cases} \frac{F_i(x)+F_i'(x)d}{F_i(x)} & \text{if } i \in E_F(x) \text{ and } F_i(x) \neq 0 \\ 0 & \text{if } i \in E_F(x) \text{ and } F_i(x) = 0 \\ \frac{G_i(x)+G_i'(x)d}{G_i(x)} & \text{if } i \in E_G(x) \text{ and } G_i(x) \neq 0 \\ 0 & \text{if } i \in E_G(x) \text{ and } G_i(x) = 0 \\ \max\left(\frac{F_i(x)+F_i'(x)d}{F_i(x)}, \frac{G_i(x)+G_i'(x)d}{G_i(x)}\right) & \text{if } i \in I(x), \end{cases}$$

Hybrid polyhedral Newton-min algorithm

Hybrid Polyhedral NM algorithm (HPNM)

- If the plain Newton-min direction d in (7) satisfies (11), take it (very cheap),
- Else take the secure polyhedral Newton-min direction d (more expensive).

Features of the HPNM algorithm:

- ⊕ in most iterations, a plain NM direction (7) is computed (a single LS to solve),
- ⊕ the number of inequalities $2|\mathcal{E}_\tau^-(x)|$ should be very small ($\tau > 0$ can be very small),
- ⊕ can be computed in polynomial time (by LO or QO),
- ⊕ d is a decrease direction of θ,
- ⊕ global convergence.

Theorem (global convergence of the HPNM algorithm)

If
- *F and $G : \Omega \to \mathbb{R}^n$ are differentiable,*
- *the HPNM algorithm generates a sequence $\{x_k\} \subseteq \Omega$,*
- *$\bar{x} \in \Omega$ is an accumulation point of $\{x_k\}$ that is NM and PNM regular,*
- *F' and G' are continuous at $\bar{x}$,*

then, $\{\theta(x_k)\}_{k \geqslant 1} \downarrow 0$ and $\bar{x}$ is a solution to the NLCP (2).

Outline

Comparison of 3 solvers [40]

- PNM (Polyhedral Newton-Min algorithm [26, 17])
  - Direction determined by solving the quadratic optimization problem (QP)

$$\min \frac{1}{2}\|d\|_2^2 \text{ s.t. } \begin{cases} F_i(x) + F_i'(x)d = 0 & \text{if } i \in E_F(x) \\ G_i(x) + G_i'(x)d = 0 & \text{if } i \in E_G(x) \\ F_i(x) + F_i'(x)d \geqslant 0 & \text{if } i \in I(x) \\ G_i(x) + G_i'(x)d \geqslant 0 & \text{if } i \in I(x). \end{cases} \tag{12}$$

  - Kink tolerance τ determined to try to have $|\mathrm{qp}| \leqslant 10$.

- HPNM (Hybrid Polyhedral Newton-Min algorithm [26, 17])
  - Take the plain Newton-min direction if it satisfies the sufficient decrease criterion (11).
  - Otherwise, take the minimum-norm PNM direction (12).
  - Kink tolerance τ determined to try to have $|\mathrm{qp}| \leqslant 10$.

- PATH (pathlcp)
  - The reference CP solver by Dirkse, Ferris, Li, Munson [27, 35, 36, 50].
  - Uses the normal map reformulation [62]: x solves (2) if and only if (x, z) solves

$$F(x) = z^+ \quad \text{and} \quad G(x) = z^-.$$

Dense random problems

Dense random problems of Harker and Pang [43]

- $M = A^{\mathsf{T}}A + \mathrm{Diag}(d) + Z \in \mathbf{P}$, with random $A \in \mathbb{R}^{n\times n}$, $d \in \mathbb{R}^n_{++}$, and $Z \in \mathcal{Z}^n$.
- q such that $0 = x_A < y_A$, $x_I > y_I = 0$, $x_E = y_E = 0$ where na $:= |A|$, ni $:= |I|$, ne $:= |E|$ are given.

			PNM					HPNM					PATH
n	na	ni	iter	#qp	|qp|	α	sec	iter	#qp	|qp|	α	sec	sec
512	128	256	29	27	7.8	$3\,10^{-1}$	0.81	6	4	8.5	$1\,10^{-0}$	0.61	0.21
1024	256	512	47	45	7.9	$2\,10^{-1}$	1.46	7	5	9.0	$1\,10^{-0}$	0.61	1.55
2048	512	1024	62	60	9.6	$1\,10^{-1}$	5.17	7	4	10.0	$1\,10^{-0}$	1.04	7.26
4096	1024	2048	134	132	8.8	$4\,10^{-2}$	57.30	8	1	10.0	$1\,10^{-0}$	3.14	45.10
8192	2048	4096	223	221	9.4	$3\,10^{-2}$	700.14	7	0	–	$1\,10^{-0}$	14.96	233.10
16384	4096	8192	425	423	9.9	$1\,10^{-2}$	9516.20	7	0	–	$1\,10^{-0}$	100.08	stuck!
$O(n^p)$ with $p =$			0.78				2.79	0.04				1.49	2.51

#qp = number of QP's, |qp| = mean size of the QP's, α = $\log_{10}$-mean stepsize, sec = tic-toc time

Academic difficult problems (Murty [54])

Problem yielding exponential complexity of the Lemke algorithms for an LCP with a **P**-matrix:

$$M = L_M := \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 2 & 1 & 0 & \ddots \\ 2 & 2 & 1 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix} \in \mathbf{P}, \qquad q = -e, \qquad \text{and} \qquad x_1 = 0. \tag{13}$$

Murty problem (S2)

		PNM					HPNM					PATH
n	sec	iter	#qp	|qp|	α	sec	iter	#qp	|qp|	α	sec	sec
512	0.00	396	394	9.8	$1\,10^{-2}$	2.65	480	49	9.7	$1\,10^{-2}$	1.66	0.03
1024	0.02	1094	1092	9.9	$3\,10^{-3}$	8.07	1061	142	10.0	$4\,10^{-3}$	5.03	0.13
2048	0.08	1850	1848	9.9	$2\,10^{-3}$	27.88	2421	412	10.0	$1\,10^{-3}$	32.98	0.63
4096	0.55	3951	3949	10.0	$1\,10^{-3}$	224.11	5821	1494	10.0	$4\,10^{-4}$	340.30	2.44
8192	2.67	7756	7754	10.0	$5\,10^{-4}$	2864.29	12880	4032	10.0	$1\,10^{-4}$	5905.34	13.10
$O(n^p)$, $p =$		1.04				2.50	1.19				2.97	2.18

#qp = number of QP's, |qp| = mean size of the QP's, α = $\log_{10}$-mean stepsize, sec = tic-toc time

Academic difficult problems (Fathi [33, 30])

Problem yielding exponential complexity of the Lemke algorithms for an LCP with a **PD**-matrix:

$$M = L_M L_M^{\mathsf{T}} \in \mathbf{PD}, \qquad q = -e, \qquad \text{and} \qquad x_1 = 0, \tag{14}$$

Fathi problem (S2)

		PNM					HPNM					PATH
n	sec	iter	#qp	|qp|	α	sec	iter	#qp	|qp|	α	sec	sec
512	0.00	255	214	5.9	$2\,10^{-2}$	2.07	248	18	10.0	$2\,10^{-2}$	1.57	2.08
1024	0.02	468	318	5.9	$1\,10^{-2}$	4.98	430	12	10.0	$2\,10^{-2}$	5.08	24.86
2048	0.09	1005	686	5.7	$4\,10^{-3}$	35.67	883	20	10.0	$4\,10^{-3}$	50.71	370.13
4096	0.55	2220	1563	5.5	$1\,10^{-3}$	525.28	1488	42	10.0	$6\,10^{-3}$	340.88	2726.22
8192	2.98	5145	3369	4.4	$7\,10^{-4}$	4574.70	2844	36	10.0	$2\,10^{-3}$	4350.27	
$O(n^p)$, $p =$		1.09				2.89	0.88				2.89	3.50

#qp = number of QP's, |qp| = mean size of the QP's, α = $\log_{10}$-mean stepsize, sec = `tic-toc` time

Diphasic flow in a porous media [8]

	PNM					HPNM					PATH
n	iter	#qp	|qp|	α	sec	iter	#qp	|qp|	α	sec	sec
201	4	0	–	$1\,10^{-0}$	0.25	4	0	–	$1\,10^{-0}$	0.27	0.04
501	4	0	–	$1\,10^{-0}$	0.26	4	0	–	$1\,10^{-0}$	0.26	0.22

#qp = number of QP's, |qp| = mean size of the QP's, α = $\log_{10}$-mean stepsize, sec = tic-toc time

Outline

Conclusion

- We have proposed a means to globalize the NM/SSN algorithm for complementarity problems.
- Sometimes spectacularly efficient (random, diphasic flow, many practical applications), but not on particular problems (Murty).
- There is still much to understand and to do, but it seems worth the effort.
  - Baptiste Plaquevent-Jourdain (PhD) works on the Levenberg-Marquardt globalization (to avoid convergence to meaningless points and weaken the regularity condition).
  - A thorough experiment campaign on LCP is programmed (with Mathieu Frappier).
  - To do: asymptotic analysis of the algorithm (admissibility of the unit stepsize, quadratic convergence, finite termination on LCP(**P**)).
  - To do: robustness of the algorithm away from a regular solution (i.e., deal with the possible infeasibility of the linearized system (10)).
  - To do: application of the same solution principle to optimization.
  - To do: application of the same solution principle to other nonsmooth systems, if any.

V. Acary, B. Brogliato (2008).
***Numerical Methods for Nonsmooth Dynamical Systems - Applications in Mechanics and Electronics*.**
Lecture Notes in Applied and Computational Mechanics 35. Springer.
[doi].

Muhamed Aganagić (1984).
Newton's method for linear complementarity problems.
Mathematical Programming, 28, 349–362.
[doi].

Zakia Bazari (2018).
***Modélisation du contact pneumatique/chaussée pour l'évaluation du bruit de roulement*.**
Thèse de doctorat - spécialité génie civil, Université de Lyon, France.
[hal].

L. Beaude, K. Brenner, S. Lopez, R. Masson, F. Smai (2019).
Non-isothermal compositional liquid gas Darcy flow: formulation, soil-atmosphere boundary condition and application to high-energy geothermal simulations.
Computational Geosciences, 23(3), 443–470.
[doi].

I. Ben Gharbia, J. Dabaghi, V. Martin, M. Vohralík (2020).
A posteriori error estimates for a compositional two-phase flow with nonlinear complementarity constraints.
Computational Geosciences, 24(3), 1031–1055.
[doi].

I. Ben Gharbia, E. Flauraud (2019).
Study of compositional multiphase flow formulation using complementarity conditions.
Oil & Gas Sciences and Technology, 74, 1–15.
[doi].

I. Ben Gharbia, J. Jaffré (2014).
Gas phase appearance and disappearance as a problem with complementarity constraints.
Mathematics and Computers in Simulation, 99, 28–36.
[doi].

Ibtihel Ben Gharbia (2012).
***Résolution de Problèmes de Complémentarité – Application à un Écoulement Diphasique Dans un Milieu Poreux*.**

PhD Thesis, Université Paris-Dauphine.

Ibtihel Ben Gharbia, J.Ch. Gilbert (2012).
Nonconvergence of the plain Newton-min algorithm for linear complementarity problems with a P-matrix.
Mathematical Programming, 134, 349–364.
[doi].

Ibtihel Ben Gharbia, J.Ch. Gilbert (2013).
An algorithmic characterization of P-matricity.
SIAM Journal on Matrix Analysis and Applications, 34(3), 904–916.
[doi].

Ibtihel Ben Gharbia, J.Ch. Gilbert (2019).
An algorithmic characterization of P-matricity II: adjustments, refinements, and validation.
SIAM Journal on Matrix Analysis and Applications, 40(2), 800–813.
[doi].

J.F. Bonnans, J.Ch. Gilbert, C. Lemaréchal, C. Sagastizábal (1997).
Optimisation Numérique – Aspects théoriques et pratiques.
Mathématiques et Applications 27. Springer Verlag, Berlin.
[editor].

J.F. Bonnans, J.Ch. Gilbert, C. Lemaréchal, C. Sagastizábal (2006).
Numerical Optimization – Theoretical and Practical Aspects (second edition).
Universitext. Springer Verlag, Berlin.
[authors] [editor] [doi].

B. Brogliato (2016).
Nonsmooth Mechanics - Models, Dynamics and Control (third edition).
Springer.
[doi].

H. Buchholzer, Ch. Kanzow, P. Knabner, S. Kräutle (2011).
The semismooth Newton method for the solution of reactive transport problems including mineral precipitation-dissolution reactions.
Computational Optimization and Applications, 50(2), 193–221.
[doi].

Q.M. Bui, H.C. Elman (2020).

Semi-smooth Newton methods for nonlinear complementarity formulation of compositional two-phase flow in porous media.
Journal of Computational Physics, 407, 109163.
[doi].

A. Chiche, J.Ch. Gilbert (2016).

How the augmented Lagrangian algorithm can deal with an infeasible convex quadratic optimization problem.
Journal of Convex Analysis, 23(2), 425–459.
[pdf] [editor].

S.J. Chung (1989).

NP-completeness of the linear complementarity problem.
Journal of Optimization Theory and Applications, 60, 393–399.
[doi].

F.H. Clarke (1983).

Optimization and Nonsmooth Analysis.
John Wiley & Sons, New York.
Reprinted in 1990 by SIAM, Classics in Applied Mathematics 5 [doi].

R.W. Cottle (1964).

Nonlinear Programs with Positively Bounded Jacobians.
PhD Thesis, University of California, Berkeley, USA.

R.W. Cottle (1966).

Nonlinear programs with positively bounded jacobians.
SIAM Journal on Applied Mathematics, 14, 147–158.
[doi].

R.W. Cottle, J.-S. Pang, R.E. Stone (1992).

The Linear Complementarity Problem.
Academic Press, Boston.

J. Dabaghi, V. Martin, M. Vohralík (2020).

Adaptive inexact semismooth Newton methods for the contact problem between two membranes.
Journal of Scientific Computing, 84(2).
[doi].

E. de Klerk, M. E.-Nagy (2011).

On the complexity of computing the handicap of a sufficient matrix.
Mathematical Programming, 129(2), 383–402.
[doi].

T. De Luca, F. Facchinei, C. Kanzow (2000).
A theoretical and numerical comparison of some semismooth algorithms for complementarity problems.
Computational Optimization and Applications, 16, 173–205.
[doi].

F. Delbos, J.Ch. Gilbert (2005).
Global linear convergence of an augmented Lagrangian algorithm for solving convex quadratic optimization problems.
Journal of Convex Analysis, 12(1), 45–69.
[preprint] [editor].

S.P. Dirkse, M.C. Ferris (1995).
The PATH solver: a non-monotone stabilization scheme for mixed complementarity problems.
Optimization Methods and Software, 5(2), 123–156.

J.-P. Dussault, M. Frappier, J.Ch. Gilbert (2023).
Polyhedral Newton-min algorithms for complementarity problems.
Mathematical Programming (submitted).
[hal-02306526].

J.-P. Dussault, J.Ch. Gilbert, B. Plaquevent-Jourdain (2023).
On the B-differential of the componentwise minimum of two affine vector functions.
Mathematical Programming Computation (submitted).
[hal-04048393].

Jean-Pierre Dussault, M. Frappier, J.Ch. Gilbert (2019).
A lower bound on the iterative complexity of the Harker and Pang globalization technique of the Newton-min algorithm for solving the linear complementarity problem.
EURO Journal on Computational Optimization, 7(4), 359–380.
[doi].

K. Erleben (2013).
Numerical methods for linear complementarity problems in physics-based animation.
In *ACM SIGGRAPH 2013 Courses*, SIGGRAPH '13, pages 8:1–8:42. ACM, New York, NY, USA.
[doi].

F. Facchinei, J.-S. Pang (2003).

Finite-Dimensional Variational Inequalities and Complementarity Problems (two volumes).
Springer Series in Operations Research. Springer.

Y. Fathi (1979).

Computational complexity of LCPs associated with positive definite symmetric matrices.
Mathematical Programming, 17, 335–344.

M.C. Ferris, J.-S. Kanzow, T.S. Munson (1999).

Feasible descent algorithms for mixed complementarity problems.
Mathematical Programming, 86, 475–497.
[doi].

M.C. Ferris, T.S. Munson (1999).

Interfaces to PATH 3.0: design, implementation and usage.
Computational Optimization and Applications, 12, 207–227.

M.C. Ferris, T.S. Munson (2000).

Complementarity problems in GAMS and the PATH solver.
Journal of Economic Dynamics and Control, 24(2), 165–188.
[doi].

M.C. Ferris, J.-S. Pang (1997).

Engineering and economic applications of complementarity problems.
SIAM Review, 39, 669–713.
[doi].

Andreas Fischer (1992).

A special Newton-type optimization method.
Optimization, 24, 269–284.
[doi].

Andreas Fischer, C. Kanzow (1996).

On finite termination of an iterative method for linear complementarity problems.
Mathematical Programming, 74, 279–292.
[doi].

Mathieu Frappier (2019).

Reformulation semi-lisse appliquée au problème de complémentarité.
Master's thesis, Département de Mathématiques, Faculté des Sciences, Université de Sherbrooke, Canada.
[internet].

Y. Gao, H. Song, X. Wang, K. Zhang (2020).
Primal-dual active set method for pricing American better-of option on two assets.
Communications in Nonlinear Science and Numerical Simulation, 80.
[doi].

P.T. Harker, J.-S. Pang (1990).
Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications.
Mathematical Programming, 48, 161–220.
[doi].

P.T. Harker, J.-S. Pang (1990).
A damped-Newton method for the linear complementarity problem.
In E.L. Allgower, K. Georg (editors), *Computational Solution of Nonlinear Systems of Equations*, Lecture in Applied Mathematics 26. AMS, Providence, RI.

P. Hartman, G. Stampacchia (1966).
On some non-linear elliptic differential-functional equations.
Acta Mathematica, 115, 271–310.
[doi].

G. Isac (1992).
Complementarity Problems.
Lecture Notes in Mathematics 1528. Springer, Berlin.
[doi].

N.H. Josephy (1979).
Newton's method for generalized equations.
Technical Summary Report 1965, Mathematics Research Center, University of Wisconsin, Madison, WI, USA.

M. Kojima, Nimrod Megiddo, T. Noma, A. Yoshise (1991).
A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems.
Lecture Notes in Computer Science 538. Springer, Berlin.
[doi].

S. Kräutle (2011).

The semismooth Newton method for multicomponent reactive transport with minerals.
Advances in Water Resources, 34(1), 137–151.
[doi].

B. Kummer (1988).

Newton's method for nondifferentiable functions.
In J. Guddat, B. Bank, H. Hollatz, P. Kall, D. Klatte, B. Kummer, K. Lommatzsch, L. Tammer, M. Vlach, K. Zimmerman (editors), *Advances in Mathematical Optimization*, pages 114–125. Akademie-Verlag, Berlin.

Q. Li, M.C. Ferris, T. Munson (2009).

Linear algebra enhancements to the PATH solver.
Report ANL/MCS-P1565-1208, Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, USA.

V.L.R. Lopes, J.M. Martínez, R. Pérez (1999).

On the local convergence of quasi-Newton methods for nonlinear complementarity problems.
Applied Numerical Mathematics, 30, 3–22.
[doi].

E. Marchand, T. Müller, P. Knabner (2012).

Fully coupled generalised hybrid-mixed finite element approximation of two-phase two-component flow in porous media. Part II: numerical scheme and numerical results.
Computational Geosciences, 16(3), 691–708.
[doi].

E. Marchand, T. Müller, P. Knabner (2013).

Fully coupled generalised hybrid-mixed finite element approximation of two-phase two-component flow in porous media. Part I: formulation and properties of the mathematical model.
Computational Geosciences, 17(2), 431–442.
[doi].

K.G. Murty (1978).

Computational complexity of complementarity pivot methods.
Mathematical Programming Study, 7, 61–73.

J.-S. Pang (1990).

Newton's method for B-differentiable equations.

Mathematics of Operations Research, 15, 311–341.
[doi].

J.-S. Pang (1991).
A B-differentiable equation-based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems.
Mathematical Programming, 51(1-3), 101–131.
[doi].

J.-S. Pang (1995).
Complementarity problems.
In R. Horst, P.M. Pardalos (editors), *Handbook of Global Optimization*, volume 2 of *Nonconvex Optimization and Its Applications*, pages 271–338. Kluwer, Dordrecht.
[doi].

S. Pieraccini, M.G. Gasparo, A. Pasquali (2003).
Global Newton-type methods and semismooth reformulations for NCP.
Applied Numerical Mathematics, 44, 367–384.
[doi].

F.A. Potra, X. Liu (2005).
Predictor-corrector methods for sufficient linear complementarity problems in a wide neighborhood of the central path.
Optimization Methods and Software, 20(1), 145–168.
[doi].

Liqun Qi (1993).
Convergence analysis of some algorithms for solving nonsmooth equations.
Mathematics of Operations Research, 18, 227–244.
[doi].

Liqun Qi, Jie Sun (1993).
A nonsmooth version of Newton's method.
Mathematical Programming, 58, 353–367.
[doi].

S.M. Robinson (1992).
Normal maps induced by linear transformations.
Mathematics of Operations Research, 17, 691–714.