Formulations and solution algorithms for Complementarity Problems (or seven ways to skin a cat)

Michael C. Ferris

University of Wisconsin, Madison

Funded by AFOSR and NSF

Auckland, April 26, 2018

AVI over polyhedral convex set

An affine function

$$F : \mathbb{R}^n \to \mathbb{R}^n,\ F(z) = Mz + q,\ M \in \mathbb{R}^{n \times n},\ q \in \mathbb{R}^n$$

A polyhedral convex set

$$\mathcal{C} = \{z \in \mathbb{R}^n \mid Az(\geq, =, \leq)a,\ l \leq z \leq u\},\ A \in \mathbb{R}^{m \times n}$$

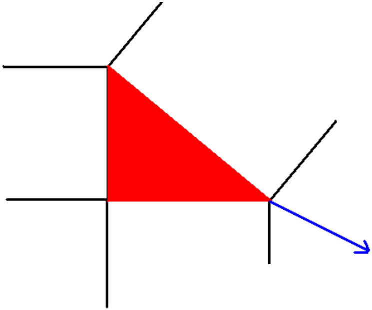
Find a point $z^* \in \mathcal{C}$ satisfying

$$\begin{aligned} &\langle F(z^*), y - z^* \rangle && \geq 0, \quad \forall y \in \mathcal{C} \\ (\Leftrightarrow)\ &\langle -F(z^*), y - z^* \rangle && \leq 0, \quad \forall y \in \mathcal{C} \\ (\Leftrightarrow)\ &-F(z^*) && \in N_{\mathcal{C}}(z^*) \end{aligned}$$

where

$$N_{\mathcal{C}}(z^*) \quad = \{v \mid \langle v, y - z^* \rangle \leq 0, \forall y \in \mathcal{C}\}$$

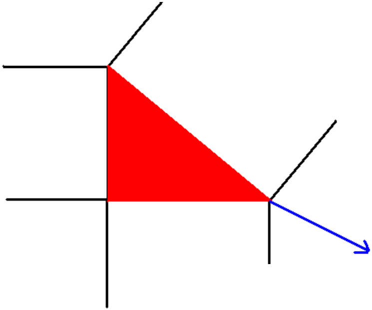
Normal map for polyhedral C



projection: $\pi_C(x)$

$$x - \pi_C(x) \in N_C(\pi_C(x))$$

Normal map for polyhedral C



projection: $\pi_C(x)$

$$x - \pi_C(x) \in N_C(\pi_C(x))$$

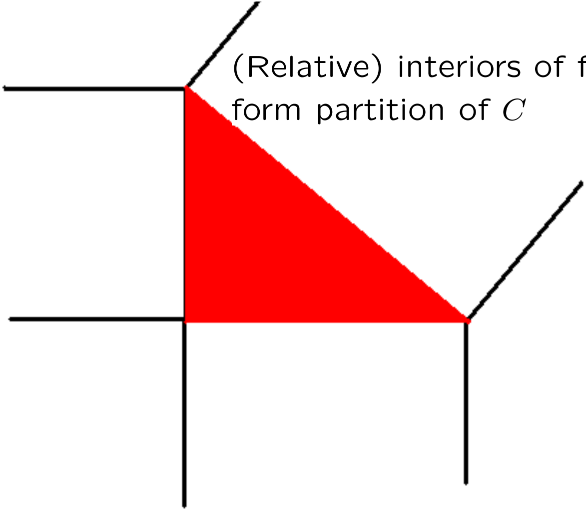
If $-M\pi_C(x) - q = x - \pi_C(x)$ then $z = \pi_C(x)$ solves

$$0 \in Mz + q + N_C(z)$$

if and only if we can find x, a zero of the normal map:

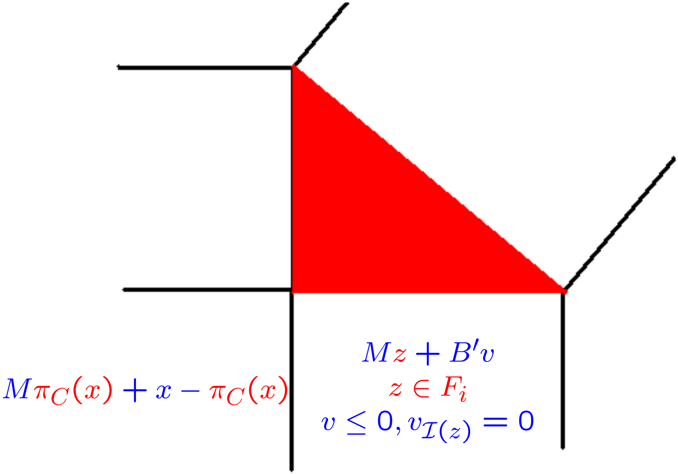
$$0 = M\pi_C(x) + q + x - \pi_C(x)$$

Normal manifold = $\{F_i + N_{F_i}\}$

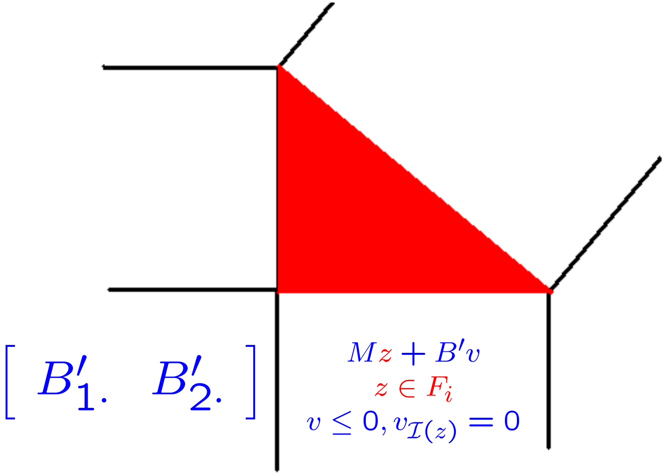


(Relative) interiors of faces F_i form partition of C

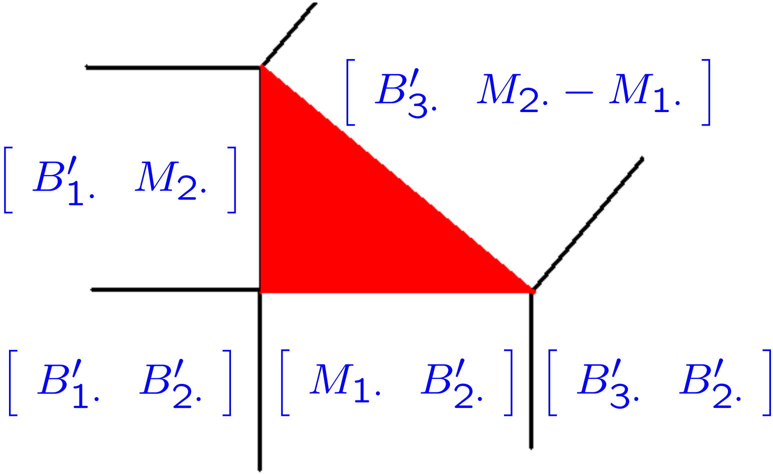
$$C = \{z | Bz \geq b\}, N_C(z) = \{B'v | v \leq 0, v_{\mathcal{I}(z)} = 0\}$$



$C = \{z | Bz \geq b\}, N_C(z) = \{B'v | v \leq 0, v_{\mathcal{I}(z)} = 0\}$

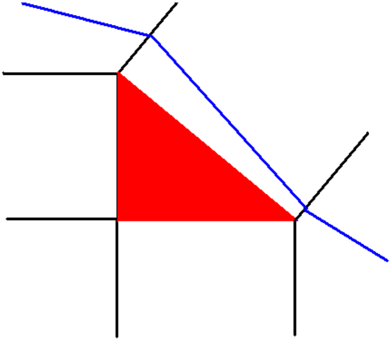


$$C = \{z | Bz \geq b\}, F(z) = Mz + q$$



The PATHAVI algorithm

- Start in cell that has interior (face is an extreme point, so normal cone has interior - primary ray)
- Move towards a zero of affine map in cell
- Update direction when hit boundary (pivot)
- Solves, or determines infeasible if M is copositive-plus on rec(C)
- Solves 2-person bimatrix games, 3-person games too, but these are nonlinear



But algorithm has exponential complexity (von Stengel et al)

Theorem

Suppose $\mathcal{C}$ is a polyhedral convex set and M is an L–matrix with respect to $\mathrm{rec}\mathcal{C}$ which is invertible on the lineality space of $\mathcal{C}$. Then exactly one of the following occurs:

- *PATHAVI solves (AVI)*
- *the following system has no solution*

$$Mz + q \in (\mathrm{rec}\mathcal{C})^D, \qquad z \in \mathcal{C}. \tag{1}$$

Corollary

If M is copositive–plus with respect to $\mathrm{rec}\mathcal{C}$, then exactly one of the following occurs:

- *PATHAVI solves (AVI)*
- *(1) has no solution*

Note also that if $\mathcal{C}$ is compact, then any matrix M is an L–matrix with respect to $\mathrm{rec}\mathcal{C}$. So always solved.

Experimental results: AVI vs MCP

- Run PATHVI over AVI formulation.
- Run PATH over rectangular form (poorer theory as $\mathrm{rec}\mathcal{C}$ larger).
- Structure knowledge leads to improved reliability

Name	(#cons,#vars)	Number of iterations (time/secs)	
		PATHVI	PATH
CVXQP1_M	(500, 1000)	3119 (0.459)	fail
CVXQP2_M	(250, 1000)	33835 (2.927)	fail
CVXQP3_M	(750, 1000)	360 (0.105)	3603 (1.992)
CONT-050	(2401, 2597)	11 (2.753)	382 (272.429)
CONT-100	(9801,10197)	3 (174.267)	fail

Extension to Nonlinear Model

- So now we can solve AVI, what happens when F is nonlinear
- Embed AVI solver in a Newton Method - each Newton step solves an AVI
- Nonlinear equations $F(x) = 0$
- Newton's Method

$$\begin{aligned} F(x^k) + \nabla F(x^k) d^k &= 0 \\ x^{k+1} &= x^k + d^k \end{aligned}$$

- Damp using Armijo linesearch on $\frac{1}{2} \|F(x)\|_2^2$
- Descent direction - gradient of merit function
- Properties
  - Well defined
  - Global and local-fast convergence

Nonsmooth Newton Method

Given x^k

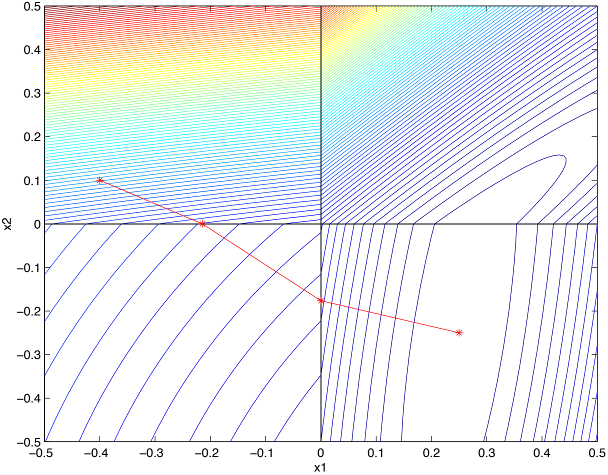
$$\text{solve: } 0 \in F(x^k) + \nabla F(x^k)(x - x^k) + N_C(x)$$
$$d_k = x^* - x^k,\ x^* \text{ from above}$$
$$x^{k+1} = x^k + \alpha d^k$$

- Equivalent piecewise smooth equation $F_+(x) = 0$

$$F_+(x) \equiv F(\pi_C(x)) + x - \pi_C(x)$$

  (when $C = \mathbb{R}^n_+$ then $\pi_C(x) = max(x, 0)$ is easy to compute)
- Nonsmooth Newton Method
  - Iteratively solve piecewise linear system of equations, via pivoting
  - Damp using Armijo search on $\frac{1}{2} \|F_+(x)\|_2^2$
- Properties
  - Global and local-fast convergence
  - Merit function *not* differentiable

Piecewise Linear Example



Fischer-Burmeister Function

$$\phi(a, b) := \sqrt{a^2 + b^2} - a - b$$

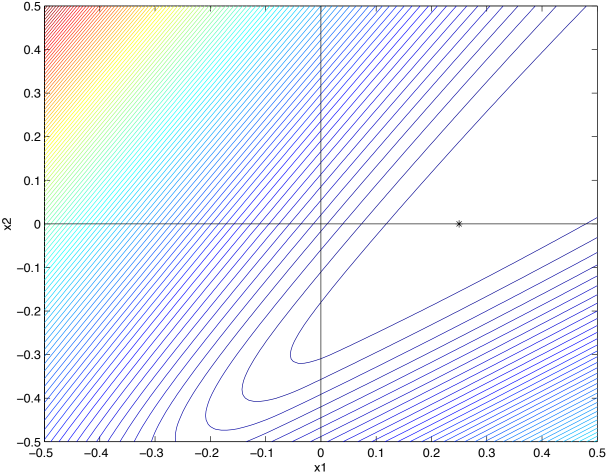
$$\phi(a, b) = 0 \iff 0 \le a \perp b \ge 0$$

- $\Phi(x)$ defined componentwise

$$\Phi_i(x) \equiv \sqrt{(x_i)^2 + (F_i(x))^2} - x_i - F_i(x)$$

- $\Phi(x) = 0$ if and only if x solves NCP(F)
- Not continuously differentiable - semismooth
- Natural merit function ($\frac{1}{2} \|\Phi(x)\|_2^2$) is differentiable

Fischer-Burmeister Example



Review

- Nonlinear Complementarity Problem
- Piecewise smooth system of equations
  - Use nonsmooth Newton Method
  - Solve linear complementarity problem per iteration
  - Merit function not differentiable
- Fischer-Burmeister
  - Differentiable merit function
- Combine to obtain new algorithm
  - Well defined
  - Global and local-fast convergence

Feasible Descent Framework

- Calculate direction using a local method
  - Generates feasible iterates
  - Local fast convergence
  - Used nonsmooth Newton Method
- Accept direction if descent for $\frac{1}{2}\|\Phi(x)\|^2$
- Otherwise use projected gradient step

Theorem

Let $\{x^k\} \subseteq \Re^n$ be a sequence generated by the algorithm that has an accumulation point x^* which is a strongly regular solution of the NCP. Then the entire sequence $\{x^k\}$ converges to this point, and the rate of convergence is Q-superlinear.

- Method is well defined
- Accumulation points are stationary points
- Locally projected gradient steps not used

Computational Details

- Preprocessing to simplify without changing underlying problem
- Crashing method to quickly identify basis
- Nonmonotone search with watchdog
- Perturbation scheme for rank deficiency
- Stable interpolating pathsearch
- Restart strategy
- Projected gradient searches

Nonlinear Complementarity Problems

- Given $F : \Re^n \to \Re^n$
- Find $x \in \Re^n$ such that

$$0 \leq F(x) \qquad x \geq 0$$

$$x^T F(x) = 0$$

- Compactly written

$$0 \leq F(x) \quad \perp \quad x \geq 0$$

- Equivalent to nonsmooth equation (min-map):

$$\min(x, F(x)) = 0$$

Nonsmooth alternatives

The normal map is one nonsmooth equation reformulation of the nonlinear complementarity problem.

We have just seen two alternatives

1. Fischer-Burmeister $\Phi(x) = 0$
2. Min-map $\min(x, F(x)) = 0$

Alternative methods generate generalized derivatives of these nonsmooth functions and use within nonsmooth Newton methods

- Approaches are relatively simple to implement and work well in many (well defined) cases
- Fundamental difference is nonsmoothness is outside F
- PATH tends to perform better (due to the heuristic extensions) on harder/messier problems

Smoothing: The Fischer Function [Burmeister]

- For NCP (with $\mu > 0$):

$$0 = \phi_\mu(x_i, F_i(x)), \quad i = 1, 2, \ldots, n$$

where

$$\phi_\mu(a, b) := \sqrt{a^2 + b^2 + \mu} - a - b$$

- Gives rise to semismooth algorithms
- Need to drive μ to 0, no longer nonsmooth
- Available within NLPEC

MIP formulations for Complementarity

Set $y_i = F_i(x)$, then additionally

$$y_i \geq 0, x_i \geq 0, x_i y_i = 0$$

If we know upper bounds on x_i and y_i we can model as:

$$(x_i, y_i) \in SOS1$$

or introduce binary variable z_i and

$$x_i \leq Mz_i, y_i \leq M(1 - z_i)$$

(or use indicator variables to turn on "fixing" constraints). Works if bounds are good and problem size is not too large. Issues with bounds on multipliers not being evident.

MPEC approaches

- Can use nonlinear programming approaches (e.g. NLPEC)
- Knitro can process MPCC's and uses penalization for complementarity
- Implicit approach: generate $y(x)$ where y solves the parametric (in x) complementarity problem, then solve

  $$\min f(x, y(x))$$

  using a bundle trust region method for example. Difficult to deal with side constraints.

Separable Structure

- Partition variables into (x, y)
- Identify separable structure

$$0 \in \begin{bmatrix} F(x) \\ G(x, y) \end{bmatrix} + \begin{bmatrix} N_{\Re^n_+}(x) \\ N_{\Re^n_+}(y) \end{bmatrix}$$

- Reductions possible if either
  1. $0 \in F(x) + N_{\Re^n_+}(x)$ has a unique solution
  2. $0 \in G(x, y) + N_{\Re^n_+}(y)$ has solution for all x
- Theory provides appropriate conditions
- Solve F and G sequentially

Conclusions

- Many formulations and algorithms for complementarity problems
- PATH algorithm is widely used, available in GAMS, AMPL, AIMMS, JUMP, Matlab, API-format
- Need for more theoretic and algorithmic enhancements in large scale and structured cases
- Need to find all solutions of complementarity problems, or to solve MPEC/MPCC to global optimality