I. Modelling with complementarity constraints

Lunteren, January 14-16, 2003

1. Introduction & problem statement
2. Applications:
   2.1. Stackelberg Games
   2.2. Multi Objective Optimization
   2.3. Other Applications
3. A naive solution approach using NLP
4. Summary

Sven Leyffer leyffer@mcs.anl.gov

Mathematics and Computer Science Division Argonne National Laboratory

1. Introduction & problem statement

Mathematical Program with Complementarity Constraints (MPCC)

$$\left\{ \begin{array}{ll} \underset{x}{\text{minimize}} & f(x) \\ \text{subject to} & c(x) \geq 0 \\ & 0 \leq x_1 \ \perp \ x_2 \geq 0 \end{array} \right.$$

where $x = (x_0, x_1, x_2)$ ... partition of variables and

$$0 \leq x_1 \ \perp \ x_2 \geq 0 \quad \Leftrightarrow \quad \text{either } x_{1i} = 0 \text{ or } x_{2i} = 0$$

Applications: [Luo et al., 1996], [Outrata et al., 1998]
& [Ferris and Pang, 1997].

... equality constraints $h(x) = 0$: no problem!

2.1. Stackelberg Games: Nash Equilibrium

Extension of classical Nash Game:

m players choose strategy x_i, $i = 1, \ldots, m$ to

$$(S_i) \begin{cases} \underset{x_i}{\text{minimize}} & f_i(\widehat{x_1}, \ldots, x_i, \ldots, \widehat{x_m}) \\ \text{subject to} & c_i(x_i) \geq 0 \end{cases}$$

given $\widehat{x}_j$, $j \neq i$ strategy of other players.

Nash Equilibrium: x^* such that $x_i^* \in$ solution set of (S_i).

i.e. no player can do better by changing his/her strategy.

2.1. Stackelberg Games: Complementarity

Nash Equilibrium $\Leftrightarrow$ Complementarity Problem

KKT conditions of each (S_i) players optimization

$$\left\{ \begin{array}{ll} \underset{x_i}{\text{minimize}} & f_i(\widehat{x_1}, \ldots, x_i, \ldots, \widehat{x_m}) \\ \text{subject to} & c_i(x_i) \geq 0 \end{array} \right. \quad \Rightarrow \quad \left\{ \begin{array}{r} \nabla_{x_i} f_i^* - \nabla_{x_i} {c_i^*}^T \lambda_i^* = 0 \\ 0 \leq \lambda_i^* \perp c_i(x_i^*) \geq 0 \end{array} \right.$$

all players have same information $\rightarrow$ players are homogeneous

2.1. Stackelberg Games: Leader-Follower

Stackelberg game has one distinct player: leader x_0 controls

Leader, x_0, anticipates/controls response of followers

Followers are Nash players
$\Rightarrow$ equilibrium constraints parameterized in x_0

$$\begin{aligned} \nabla_{x_i} f_i(x_0; \ldots, x_i, \ldots) - \nabla_{x_i} c_i(x_0; x_i)^T \lambda_i &= 0 \\ c_i(x_0; x_i) - s_i &= 0 \\ 0 \leq \lambda_i \perp s_i &\geq 0 \end{aligned}$$

... become constraints in leader's optimization problem ...

2.1. Stackelberg Games: Leader's problem

$$
\left\{
\begin{array}{lll}
\min\limits_{x_0} & f(x_0) & \ldots \text{ leader's objective} \\
\text{s.t.} & c(x_0) \geq 0 & \ldots \text{ leader's constraints} \\
& \boxed{\begin{array}{r}
\nabla_{x_i} f_i(x_0; \ldots, x_i, \ldots) - \nabla_{x_i} c_i(x_0; x_i)^T \lambda_i = 0 \\
c_i(x_0; x_i) - s_i = 0 \\
0 \leq \lambda_i \perp s_i \geq 0
\end{array}} & \ldots \text{ followers}
\end{array}
\right.
$$

Mathematical Program with Complementarity Constraints (MPCCs)

x_0 controls or upper level variables

x_i, λ_i, s_i states or lower level variables $i = 1, \ldots, m$

2.1. Stackelberg Games: Examples

Standard example:

∘ government = leader; sets tax rates x_0

∘ consumers = followers; black market economy ...

Stackelberg games applied in oligopolistic market analysis

∘ analysis/design of electricity markets

2000 California Electricity Market Crash

∘ dominant player exercise market power:

∘ increase prices by withholding NOx permits

Extension:

- EPEC: Equilibrium Problem with Equilibrium Constraints
  - ∘ Nash Game between two (or more) Stackelberg players

    $\Rightarrow$ Equilibrium between several MPCCs (theory?, numerics?)

2.2. Multi Objective Optimization Problems

... optimization problems with conflicting aims ...

Multi-Objective Optimization Problem (MOOP)

$$(P) \begin{cases} \underset{x \in X}{\text{minimize}} & f(x) := (f_1(x), \ldots, f_p(x)) \\ \text{subject to} & c(x) \geq 0 \end{cases}$$

f, c smooth & well behaved.

Applications

◦ Bridge design: minimize total mass & maximize stiffness

◦ Airplane design: maximize fuel efficiency & payload;
minimize the weight & cabin noise.

2.2. MOOP: Finding a single solution ...

Given: weights $w_k \geq 0$, $k = 1, \ldots, p \ : \ \sum w_k = 1$, solve

$$(P_{\mathsf{SUM}}) \left\{ \begin{array}{ll} \underset{x \in X}{\text{minimize}} & \displaystyle\sum_{k=1}^{p} w_k f_k(x) \\ \text{subject to} & c(x) \geq 0 \end{array} \right.$$

Solution x^* is single Pareto point.

$$\boxed{\text{KKT conditions of } (P_{\mathsf{SUM}})}$$

$$\Updownarrow$$

$$\boxed{\text{necessary conditions for Pareto point}}$$

Other approaches:

- Maximum effectiveness method.
- Goal Programming (most popular) ...

2.2. MOOP: Finding multiple solutions ...

Aim: evaluate trade-offs $\Rightarrow$ need more than 1 alternative ...

For different weights $w = w^1, w^2, \ldots, w^q$, solve

$$(P_{\mathsf{SUM}}) \begin{cases} \underset{x \in X}{\text{minimize}} & \displaystyle\sum_{k=1}^{p} w_k^l f_k(x) \\ \text{subject to} & c(x) \geq 0 \end{cases}, \; l = 1, \ldots, q$$

Disadvantages:

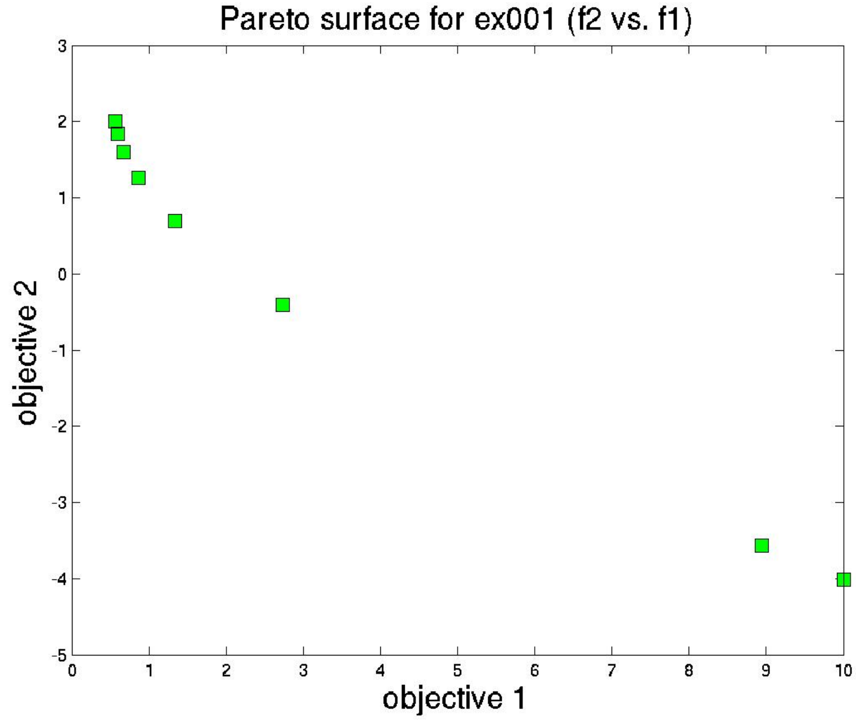
1. Nonconvex Pareto curve

   $\Rightarrow \nexists\, w$, weights to represent nonconvex part of Pareto curve.

2. Uniform spread of weights $\not\Rightarrow$ uniform description of Pareto curve.

2.2. MOOP: Finding multiple solution ...

Example: 2 objectives, 10 uniform weights $w^l = \left(\frac{l-1}{10}, 1 - \frac{l-1}{10}\right)$.



Pareto curve with uniformly distributed weights
$\Rightarrow$ poor representation of Pareto set

2.2. MOOP as bilevel program (new!)

What is "good" representation $\mathcal{R} \subset \mathcal{P}$ of Pareto set? [Sayin, 2000]

1. Coverage error e_c: all elements of $\mathcal{P}$ represented:

$$e_c := \begin{cases} \text{minimize} & e_c \\ \text{subject to} & e_c \geq \|v - u\| \qquad \forall v \in \mathcal{P} \qquad \forall u \in \mathcal{R} \end{cases}$$

   Requires knowledge of $\mathcal{P}$, not available!

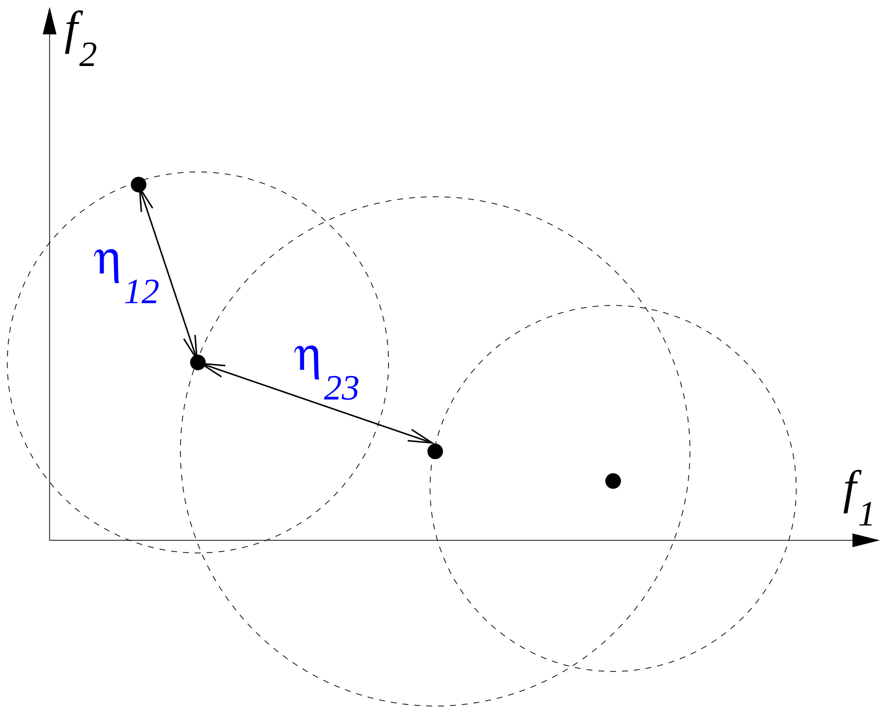
2. Uniformity of representation: no redundancy:

$$\mathcal{R}\ \eta\text{-uniform, iff} \quad \eta \leq \min_{u,v \in \mathcal{R},\ u \neq v} \|u - v\|$$

   Sensible & computable $\Rightarrow$ use here

Uniformity of representation

Let $\eta_{l_1,l_2} := \|f(x^{l_1}) - f(x^{l_2})\|$ for $l_1 \neq l_2$



Representation $\mathcal{R}$

η-uniform, iff

$$\eta = \min_{l_1 \neq l_2} (\eta_{l_1,l_2})$$

2.2. MOOP & bilevel program: Basic idea

Find representation $\mathcal{R}$, maximizing uniformity η

$$\begin{cases} \underset{x,w,\eta}{\text{maximize}} & \eta \\ \text{subject to} & \eta \leq \|f(x^{l_1}) - f(x^{l_2})\|_2^2 \quad l_1 \neq l_2 \\ & x^l \;\; \text{solves} \;\; (P_{\mathsf{SUM}}(w^l)) \qquad \forall l \end{cases}$$

where

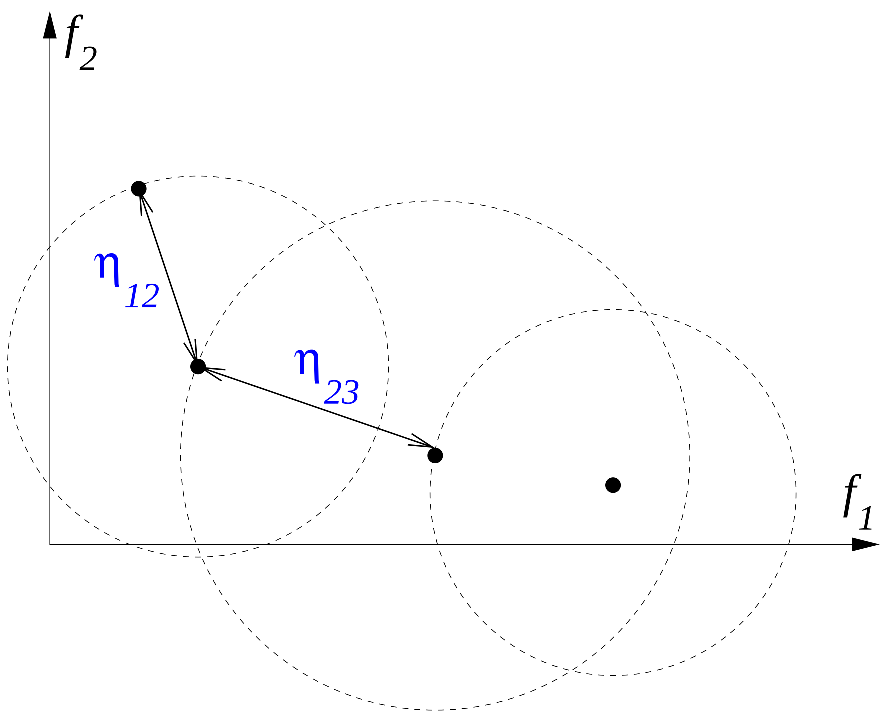
$$(P_{\mathsf{SUM}}(w^l)) \begin{cases} \underset{x \in X}{\text{minimize}} & \displaystyle\sum_{k=1}^{p} w_k^l f_k(x) \\ \text{subject to} & c(x) \geq 0 \end{cases}$$

Bilevel optimization problem (x^l is solution to NLP).

Basic idea (cont.)

Get representation $\mathcal{R}$, maximizing uniformity η by varying weights w^l.

$\Leftrightarrow$ maximize smallest distance between any pair $f(x^{l_1}), f(x^{l_2})$ for $l_1 \neq l_2$ by varying weights w^l.



Interpretation of BLP

2.2. MOOP: Formulation as MPCC

Replace 2nd level NLP $(P_{\mathsf{SUM}}(w^l))$ by KKT cond^s:

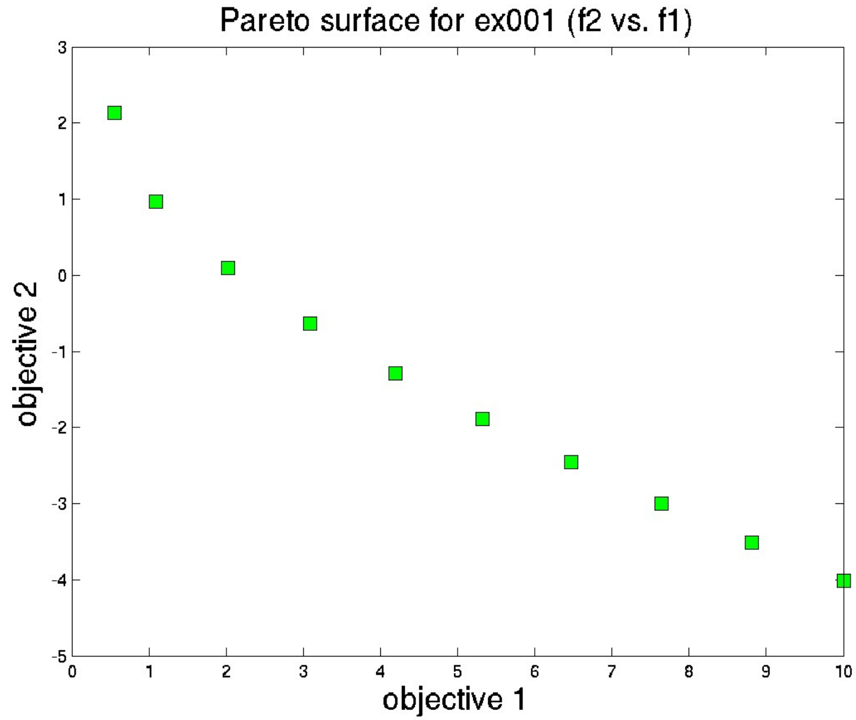
$$(P_{\mathsf{SUM}}(w^l)) \;\Rightarrow\; \begin{cases} 0 = \nabla\left(w^{l^T} f(x^l)\right) - \nabla c(x^l)\lambda^l \\ 0 \le \lambda^l \perp c(x^l) \ge 0 \end{cases}$$

then

$$(\mathsf{BLP}) \;\Leftrightarrow\; \begin{cases} \underset{x^l,\lambda^l,w^l,\eta}{\text{maximize}} & \eta & \\ \text{subject to} & \eta \le \|f(x^{l_1}) - f(x^{l_2})\|_2^2 & l_1 \ne l_2 \\ & 0 = \nabla\left(w^{l^T} f(x^l)\right) - \nabla c(x^l)\lambda^l & \forall l \\ & 0 \le \lambda^l \perp c(x^l) \ge 0 & \forall l \end{cases}$$

2.2. MOOP: Solution of MPCC

Example: 2 objectives, MPCC solution



Optimally chosen weights

$\Rightarrow$ uniform representation of Pareto set

2.3. MPCC: Engineering Applications

- design of structures involving friction [Ferris and Tin-Loi, 1999a]
- brittle fracture identification [Tin-Loi and Que, 2002]
- problems in elastoplasticity [Ferris and Tin-Loi, 1999b]
- process engineering models [Rico-Ramirez and Westerberg, 1999], [Raghunathan and Biegler, 2002]
- floor planning in design of semi-conductor devices [Anjos and Vanelli, 2002]
- obstacle problems (PDE); packaging problems [Outrata et al., 1998]

2.3. MPCC: Engineering Applications (cont.)

Models involving nonsmooth functions; e.g. pipeline network

Q_{ij} = flow through pipe; Δ_{ij} = pressure drop

$$\Delta_{ij} = K \,\text{sign}(Q_{ij})\, Q_{ij}^2$$

... usually model with 0-1 variables & "big-M" $\Rightarrow$ notoriously bad

Split $Q_{ij} = Q_{ij}^+ - Q_{ij}^-$ into positive/negative part ...

$$\begin{cases} \Delta_{ij} = K \left({Q_{ij}^+}^2 - {Q_{ij}^-}^2 \right) \\ 0 \le Q_{ij}^+ \perp Q_{ij}^- \ge 0 \end{cases}$$

$\Rightarrow$ smooth problem ... similar for $\max$ functions

2.3. MPCC: Economic Applications

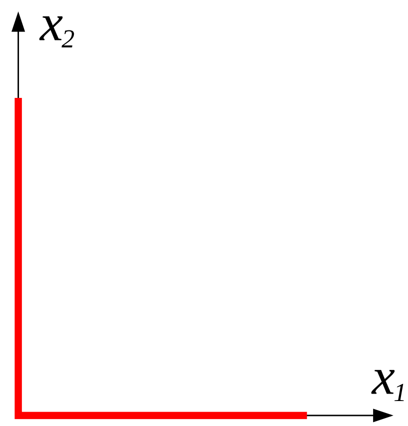
- Stackelberg games [Stackelberg, 1952]
- modeling of competition in deregulated electricity markets [Pieper, 2001], [Hobbs et al., 2000]
- volatility estimation in American option pricing [Wilmott et al., 1993], [Huang and Pang, 1999]
- transportation network design
  - toll road pricing: how to set toll levels = leader
  - consumers move optimally (Wardrop's principle) = followers
  - [Hearn and Ramana, 1997], [Ferris et al., 1999]

3. A naive solution approach

Replace complementarity $0 \le x_1 \perp x_2 \ge 0$ ...
by nonlinear equations $X_1 x_2 = 0$ or $x_1^T x_2 = 0$

Since $x_1, \ x_2 \ge 0 \quad \Rightarrow$ relax "$=$" to "$\le$" ...

$$\begin{array}{ll} \underset{x}{\text{minimize}} & f(x) \\ \text{subject to} & c(x) \ge 0 \\ & x_1, \ x_2 \ge 0 \\ & \boxed{X_1 x_2 \le 0} \end{array}$$



Advantage: standard (?) NLP; use large-scale solvers ...

Snag: NLP violates standard assumptions!

3.1. Mangasarian Fromowitz CQ fails

Mangasarian Fromowitz Constraint Qualification at feasible $\hat{x}$:

$\exists\ s$ such that $\hat{x}_1 + s_1 > 0,\ \hat{x}_2 + s_2 > 0$ and $\hat{X}_2 s_1 + \hat{X}_1 s_2 < 0$

... violated, e.g.

Case 1: $\hat{x}_1 = \hat{x}_2 = 0 \quad \Rightarrow \quad 0 < 0$

Case 2: $\hat{x}_1 > 0,\ \hat{x}_2 = 0 \quad \Rightarrow \quad s_2 > 0$ and $\hat{X}_1 s_2 < 0$

MFCQ is important (minimalistic) stability assumption for NLP

MFCQ $\Leftrightarrow$ Lagrange multiplier set bounded

3.2. Consequences of MFCQ failure

1. constraint gradients are linearly dependent:
   - At $(0, \alpha)$ get $x_1 \geq 0$ and $x_1 \alpha \leq 0$ *both* active

$$\Rightarrow \quad \text{gradients} \begin{bmatrix} 1 & \alpha \\ 0 & 0 \end{bmatrix} \text{ linearly dependent}$$

   $\Rightarrow$ slow convergence

2. central path fails to exist:

   Cannot find $x_1,\ x_2 > 0$ such that $X_1 x_2 < 0$

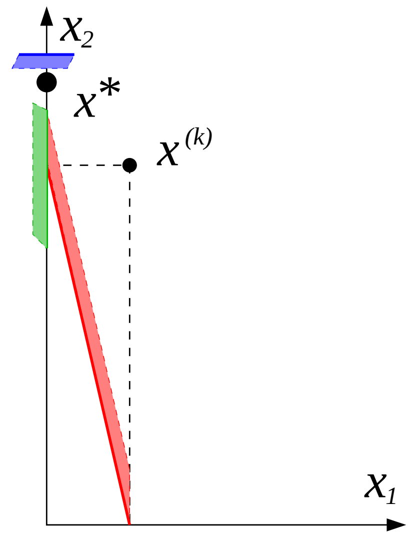
   $\Rightarrow$ multipliers blow up in practice ...

3.2. Consequences of MFCQ failure (cont.)

All numerical methods based on linearization ...

... but linearization inconsistent arbitrarily close to solution

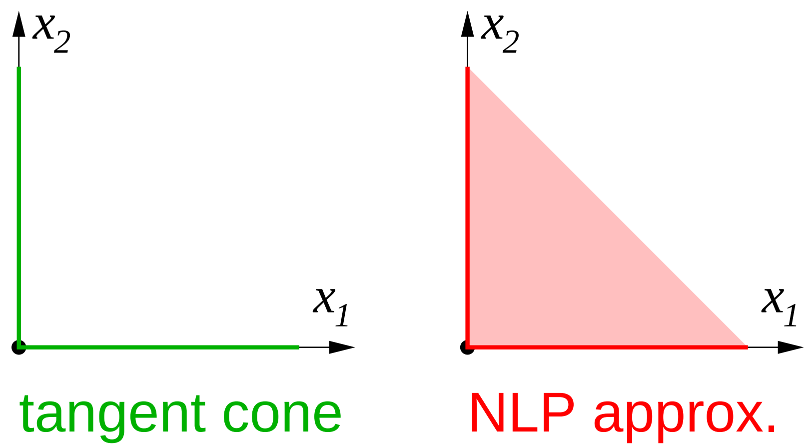
$$\left\{ \begin{array}{ll} \underset{x}{\text{minimize}} & x_1 + x_2 \\ \text{subject to} & x_2^2 \geq 1 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \\ & x_1 \, x_2 \leq 0 \end{array} \right.$$



generic problem ⇒ solvers take arbitrary steps

3.2. Consequences of MFCQ failure (cont.)

"linearization lousy approximation of tangent cone"



... NLP approx. over-estimates tangent cone at $(0, 0)$

3.3. Numerical experience with $X_1 x_2 \leq 0$

Small convex bilevel problems [Bard, 1988]

◦ numerical experience with GRG (reduced gradient method)

$\Rightarrow$ failure on 50 - 70 % of problems

Nonlinear complementarity problem [Ferris and Pang, 1997]

◦ CUTE problem `LUBRIF`: elastohydrodynamic lubrication

◦ `LANCELOT` fails due to complementarity

4. Summary part I

- Mathematical Program with Complementarity Constraints

∘ useful modelling paradigm

∘ many practical applications (engineering & economics)

- Equivalent NLP ($X_1 x_2 \leq 0$) violates Mangasarian Fromowitz CQ

  ⇒ ∘ unbounded multipliers

  ∘ constraint gradients linearly dependent

  ∘ central path fails to exist

  ∘ inconsistent linearizations

⇒ expect all sorts of numerical trouble ???

II. Solving MPCCs

Lunteren, January 14-16, 2003

5. Review of part I: complementarity constraints
6. Special purpose MPCC methods
7. Numerical experience with NLP solvers
8. Convergence of NLP solvers
   8.1. SQP methods lead the way ...
   8.2. Robust Interior Point Methods
9. Alternative NLP formulations
10. Conclusions & Open Questions

Sven Leyffer leyffer@mcs.anl.gov

Mathematics and Computer Science Division Argonne National Laboratory

5. Review of Part I

Mathematical Program with Complementarity Constraints (MPCC)

$$\begin{cases} \underset{x}{\text{minimize}} & f(x) \\ \text{subject to} & c(x) \geq 0 \\ & 0 \leq x_1 \;\perp\; x_2 \geq 0 \end{cases}$$

where $x = (x_0, x_1, x_2)$ ... partition of variables and

$$0 \leq x_1 \;\perp\; x_2 \geq 0 \quad \Leftrightarrow \quad \text{either } x_{1i} = 0 \text{ or } x_{2i} = 0$$

5. Review of Part I (cont.)

Write MPCC as equivalent NLP

$$\begin{cases} \underset{x}{\text{minimize}} & f(x) \\ \text{subject to} & c(x) \geq 0 \\ & x_1, x_2 \geq 0, \quad X_1 x_2 \leq 0 \end{cases}$$

Theoretical & numerical difficulties:

- NLP violates Mangasarian Fromowitz CQ
  - $\Rightarrow$ ∘ unbounded multipliers
  - ∘ constraint gradients linearly dependent
  - ∘ central path fails to exist
  - ∘ inconsistent linearizations

6. Special purpose MPCC methods

Apparent difficulties of NLP motivate development of MPCC solvers:

1. implicit non-smooth techniques [Outrata et al., 1998],
2. smoothing & penalization approaches [Scholtes, 2001],
3. branch-and-bound: branch on $x_{1i} = 0$ or $x_{2i} = 0$ [Bard, 1988],
4. SQPEC and PIPA ... [Luo et al., 1996]

... require significantly more work than NLP approach

6.1. Implicit non-smooth techniques

Key assumptions:

$\forall x_0 \; \exists$ unique $(x_1, x_2) \in \text{sol}(c(x) \leq 0, \; 0 \leq x_1 \perp x_2 \geq 0)$

$\Rightarrow$ obtain non-smooth functions $x_1(x_0)$ and $x_2(x_0)$

MPCC now equivalent to

$$(NS) \left\{ \begin{array}{ll} \underset{x_0}{\text{minimize}} & f(x_0, x_1(x_0), x_2(x_0)) \\ \text{subject to} & x_0 \in X \end{array} \right.$$

... objective nonsmooth.

- apply "bundle method" to solve (NS)
- solve complementarity problem for every (x_0) with `PATH`

6.2. Smoothing & penalization approaches

smoothing

$$\left\{ \begin{array}{ll} \underset{x}{\text{minimize}} & f(x) \\ \text{subject to} & c(x) \geq 0 \\ & x_1, x_2 \geq 0` \\ & X_1 x_2 \leq \tau e \end{array} \right.$$

penalization

$$\left\{ \begin{array}{ll} \underset{x}{\text{minimize}} & f(x) + \rho x_1^T x_2 \\ \text{subject to} & c(x) \geq 0 \\ & x_1, x_2 \geq 0 \end{array} \right.$$

... solve sequence of NLPs: $\tau \to 0$ or $\rho \to \infty$

... NLPs satisfy Mangasarian Fromowitz CQ for $\tau > 0$ or $\rho > 0$

7. Numerical Experience with NLP solvers

Numerical experience with MacMPEC
`www.mcs.anl.gov/~leyffer/MacMPEC/`

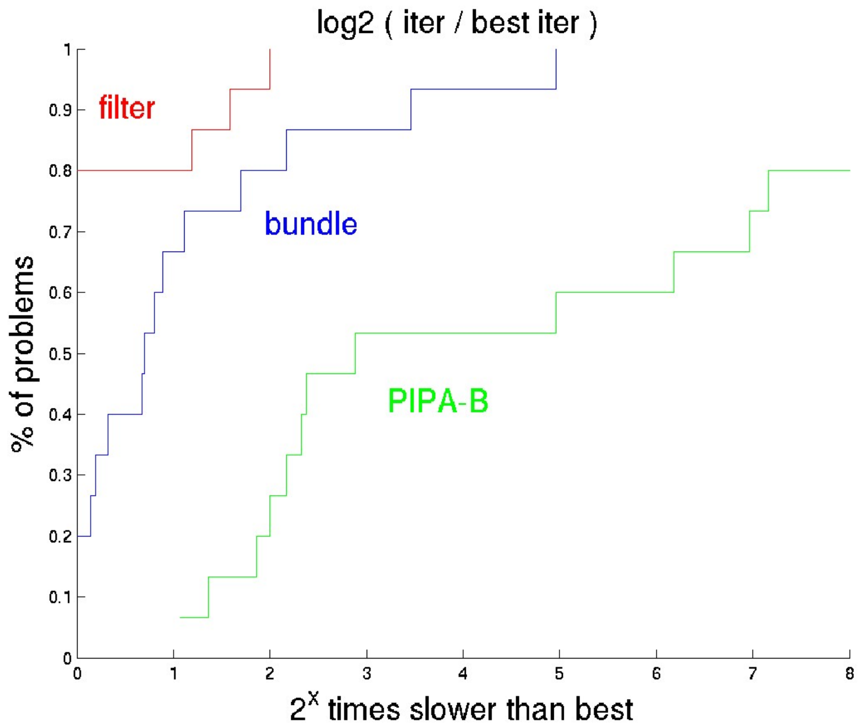
AMPL interface to SQP
use AMPL's `complements`
about 150 problems
up to 7000 variables

MPCC solvers:
- bundle: implicit non-smooth approach
- PIPA-B: penalty interior point algorithm

7.1. SQP vs. MPCC solvers

Performance plots: (subset of 15 problems)

$$\forall \text{ solver } s \quad \log_2\left(\frac{\#\text{ iter}(s,p)}{\text{best\_iter}(p)}\right), \; p \in \text{problem}$$



⇒ SQP faster

& more reliable

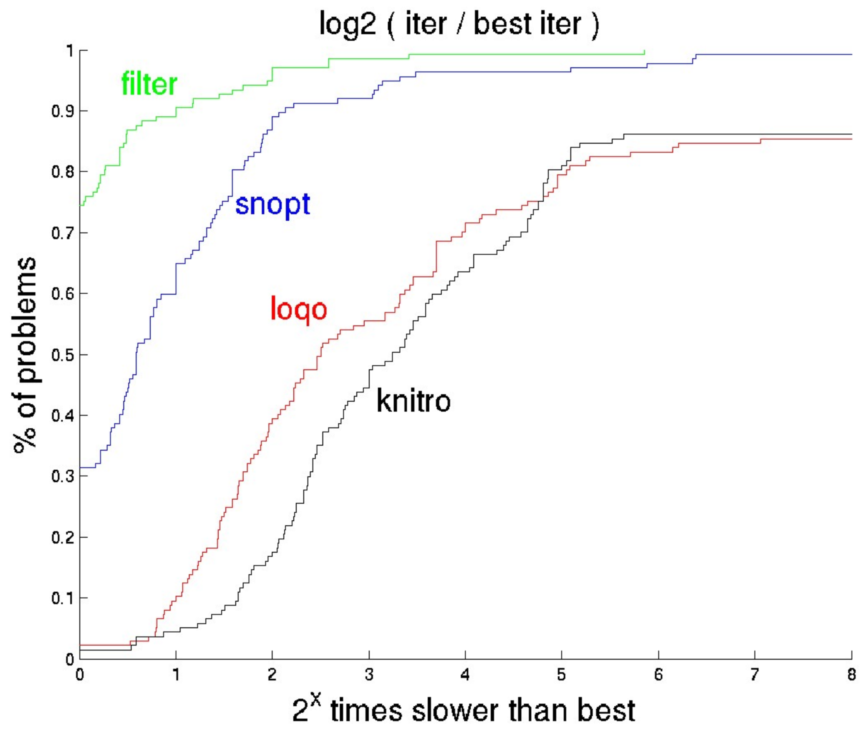
than MPCC solvers

7.2. Comparison of NLP solvers

SQP: filter & snopt

IPM: knitro & loqo

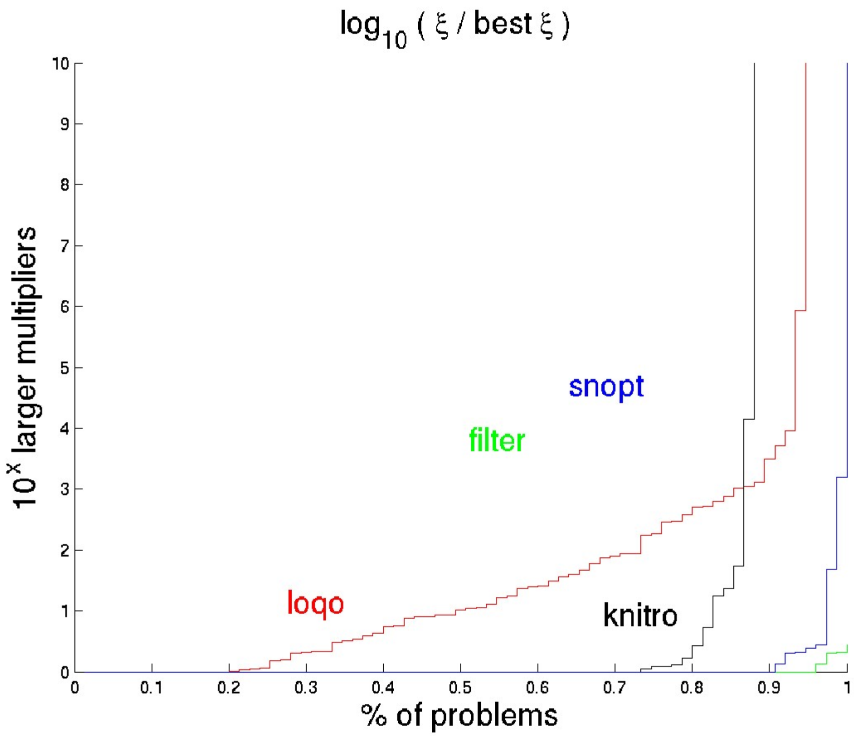
- SQP fast & reliable
- NLP better than MPCC
- IPM solvers less robust



NEOS-server `www-neos.mcs.anl.gov/`

7.2. Comparison of NLP multipliers

IPM multipliers larger
⇒ smaller slacks
⇒ slower convergence



8. Convergence of NLP solvers: Outline

Key idea: strong-stationarity $\Leftrightarrow$ KKT conditions of equivalent NLP

Two techniques:

1. relaxation of complementarity
2. penalization of complementarity $\Rightarrow$ well behaved problem

... apply within Interior Point or SQP method ...

Aim: NLP solver with small modification works for MPCCs

8. Convergence of NLP solvers: Strong stationarity

Let $\mathcal{X}_1 \equiv \{i : x^*_{1i} = 0\}$ and $\mathcal{X}_1^\perp$ complement ... $\mathcal{X}_2$ similar

Relaxed NLP defined as

$$(R) \left\{ \begin{array}{lll} \underset{x}{\text{minimize}} & f(x) & \\ \text{subject to} & c(x) \geq 0 & \\ & x_{1i} = 0 & i \in \mathcal{X}_2^\perp \\ & x_{2i} = 0 & i \in \mathcal{X}_1^\perp \\ & x_{1i} \geq 0 & i \in \mathcal{X}_2 \\ & x_{2i} \geq 0 & i \in \mathcal{X}_1 \end{array} \right.$$

... well behaved NLP ... only use concept in proof

$\hat{x}$ solves (R) and $\widehat{x_1}^T \widehat{x_2} = 0 \Rightarrow$ solved MPCC !!!

8. Convergence of NLP solvers: Strong stationarity

KKT conditions of relaxed NLP:

x^*strongly stationary $\Rightarrow$ $\exists$ multipliers $\lambda^* \geq 0$, $\hat{\nu}_1, \hat{\nu}_2$:

$$\nabla f(x^*) - \nabla c(x^*)^T \lambda^* - \begin{pmatrix} 0 \\ \hat{\nu}_1 \\ \hat{\nu}_2 \end{pmatrix} = 0 \quad 1^{st} \text{ order}$$

$$c(x^*) \geq 0,\ x_1^* \geq 0,\ x_2^* \geq 0 \ \text{ and } \ x_{1i}^* = 0 \text{ or } x_{2i}^* = 0 \quad \text{primal feas.}$$

$$c(x^*)^T \lambda = {x_1^*}^T \hat{\nu}_1 = {x_2^*}^T \hat{\nu}_2 = 0 \quad \text{compl. slack.}$$

$$\hat{\nu}_{1i}, \hat{\nu}_{2i} \geq 0, \ \text{ if } x_{1i}^* = x_{2i}^* = 0$$

$\Rightarrow$ $\exists$ bounded multipliers

8. Convergence of NLP solvers: Strong stationarity (cont.)

KKT conditions of equivalent NLP: $\exists \lambda^*, \nu_1^*, \nu_2^*, \xi^* \geq 0$

$$\nabla f(x^*) - \nabla c(x^*)^T \lambda^* - \begin{pmatrix} 0 \\ \nu_1^* - X_2^* \xi^* \\ \nu_2^* - X_1^* \xi^* \end{pmatrix} = 0$$

$$c(x^*) \geq 0,\ x_1^* \geq 0,\ x_2^* \geq 0 \ \text{ and } \ X_1^* x_2^* \leq 0$$

$$c(x^*)^T \lambda \ = \ {x_1^*}^T \nu_1^* \ = \ {x_2^*}^T \nu_2^* = 0$$

multiplers of relaxed NLP

$$\hat{\nu}_1 := \nu_1 - X_2^* \xi$$
$$\hat{\nu}_2 := \nu_2 - X_1^* \xi$$

remain bounded!!!

Illustrative example ($x^* = (0,1)$)

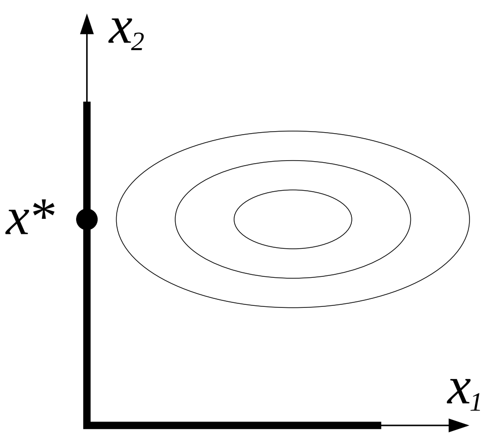
$$\begin{cases} \min_x & \frac{1}{2}(x_1-1)^2 + (x_2-1)^2 \\ \text{s.t.} & 0 \le x_1 \perp x_2 \ge 0 \end{cases}$$

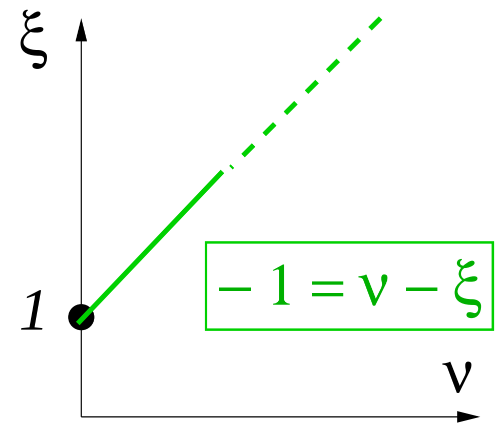
KKT conditions:

$$\begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} \nu_1 \\ 0 \end{pmatrix} - \xi \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

ν multiplier of $x_1 \ge 0$; ξ multiplier of $x_1 x_2 \le 0$.

Equivalent NLP ($x_1 x_2 \le 0$) violates MFCQ $\Rightarrow$ unbounded multipliers





multipliers are a ray !!!

8. Convergence of NLP solvers: Strong stationarity (cont.)

$\Rightarrow$ $\exists$ minimal (basic) multiplier ξ:

$$\xi_i = \begin{cases} 0 & \text{if } x^*_{1i} = x^*_{2i} = 0 \\ \max\left(0, \dfrac{-\hat{\nu}_{1i}}{x^*_{2i}}\right) & \text{if } x^*_{2i} > 0 \\ \\ \max\left(0, \dfrac{-\hat{\nu}_{2i}}{x^*_{1i}}\right) & \text{if } x^*_{1i} > 0. \end{cases}$$

$\Rightarrow$ • multipliers form ray ($\hat{\nu}_1 := \nu_1 - X^*_2 \xi$)

• minimal (basic) multiplier is complementary

$0 \le \xi \perp \nu_1 \ge 0$ and $0 \le \xi \perp \nu_2 \ge 0$

• non-zeros of minimal multiplier $\equiv$ linearly independent gradients

8.1 Convergence of SQP for MPCCs

Sequential Quadratic Programming (SQP)

$$
\left\{
\begin{array}{ll}
\min\limits_{x} & f(x) \\
\text{s.t.} & c(x) \geq 0 \\
 & x_1 \geq 0 \\
 & x_2 \geq 0 \\
 & x_1^T x_2 \leq 0
\end{array}
\right.
\quad \rightarrow \quad
\left\{
\begin{array}{ll}
\min\limits_{d} & \nabla f^{k^T} d + \frac{1}{2} d^T W^k d \\
\text{s.t.} & c^k + \nabla c^{k^T} d \geq 0 \\
 & x_1^k + d_1^k \geq 0 \\
 & x_2^k + d_2^k \geq 0 \\
 & x_1^{k^T} x_2^k + x_2^{k^T} d_1 + x_1^{k^T} d_2 \leq 0
\end{array}
\right.
$$

where $W^k \simeq \nabla^2 f^k - \sum \lambda_i \nabla^2 c^k$ Hessian of the Lagrangian.

Proof in two parts [Fletcher et al., 2002]:

1. $x_1^{k^T} x_2^k = 0$ & close to solution $\Rightarrow$ $x_1^{(k+1)^T} x_2^{(k+1)} = 0$ stay on axis
2. $x_1^{k^T} x_2^k > 0 \ \forall k \Rightarrow$ basis bounded away from singularity

$\Rightarrow$ 2nd order convergence

8.1. Convergence of SQP for MPCCs (Part 1.)

${x_1^{(k)}}^T x_2^{(k)} = 0$, i.e. on axis

wlog $x_1^k > 0 \Rightarrow$ QP includes $d_2 \geq 0$ and $x_1^k d_2 \leq 0$

$\Rightarrow d_2 = 0$ same as QP for relaxed NLP ...

$\Rightarrow$ work like SQP for relaxed NLP (well behaved)

$\Rightarrow {x_1^{(k+1)}}^T x_2^{(k+1)} = 0$; remain on same face.

QP picks non-singular basis $\equiv$ non-zeros of minimal multiplier

$\Rightarrow$ quadratic convergence

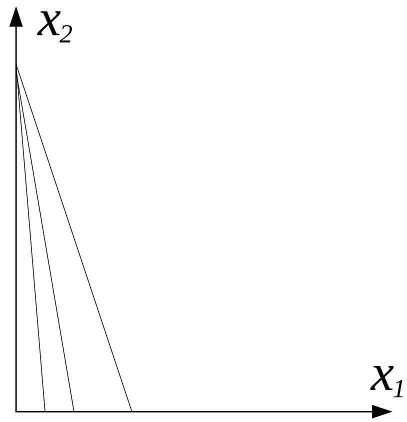
NB: Slacks matter: not true for $0 \leq x_1 \perp F(x) \geq 0$.

8.1. Convergence of SQP for MPCCs (Part 2.)

${x_1^{(k)}}^T x_2^{(k)} > 0$, i.e. off axis (assume wlog $x_1^* = 0$)

QP picks basis, subset of

$$\begin{bmatrix} I & x_2^{(k)} \\ 0 & x_1^{(k)} \end{bmatrix}$$



Assume all QP consistent (yuk!) ... 2 cases:

case 1: true subset $\Rightarrow$ non-singular $\Rightarrow$ quadratic convergence

case 2: full set $\Rightarrow$ $x_1^{(k)} > 0$ (otherwise singular)

$\Rightarrow$ ${x_1^{(k+1)}}^T x_2^{(k+1)} = 0$ now see 1. above.

8.1. Convergence of SQP for MPCCs

No MFCQ $\Rightarrow$ inconsistent linearizations near solution

Remedy: relax linearization of $X_1 x_2 \leq 0$, constants $0 < \delta, \kappa < 1$:

$$X_1^k x_2^k + X_2^k d_1 + X_1^k d_2 \leq \delta \left(x_1^{k^T} x_2^k \right)^{1+\kappa}$$

Many practical MPCCs have consistent linearizations

[Anitescu, 2000] shows convergence for ...
... SQP with elastic mode $\equiv$ penalization (add $\rho x_1^T x_2$ to objective)

8.2. Standard Interior Point Methods

Why/how do IPM solvers fail ?

1. IPMs are lousy for infeasible NLPs; struggle without $X_1 x_2 \leq 0$
2. Central path does not exist for MPCCs.
3. Special problem for knitro
   - step decomposition: normal step + tangential step
     small slacks & large multipliers
     $\Rightarrow$ small step size (0.5), due to tangential step

Surprise:

- IPMs still work OK for 80 % of MPCCs ...
  smallish multipliers (especially loqo)

8.2. Standard Interior Point Methods

Counter example for PIPA:

- small well behaved MPCC; unique minimum
- forces PIPA to converge to a non-stationary point

$\Rightarrow$ contradicts convergence theory of PIPA

Reason for PIPA's failure: direction finding QP problem contains

$$\|d_0\|^2 \leq c \left(\|c^k\| + {x_1^k}^T x_2^k \right)$$

trust-region on step in controls x_0

$\Rightarrow$ when x^k becomes feasible, then rhs $\to 0$

$\Rightarrow$ limits progress to optimality

8.2. Robust Interior Point Methods

Ignore $c(z) \geq 0$ constraints, primal-dual equations ...

$$\left\{ \begin{array}{ll} \min\limits_{x} & f(x) \\ \text{s.t.} & x_1 \geq 0 \\ & x_2 \geq 0 \\ & X_1 x_2 \leq 0 \end{array} \right. \rightarrow \left\{ \begin{array}{l} \nabla f(x) - \begin{pmatrix} 0 \\ \nu_1 - X_2 \xi \\ \nu_2 - X_1 \xi \end{pmatrix} = 0 \\ X_1 x_2 + s = 0 \\ X_1 \nu_1 = \mu e \\ X_2 \nu_2 = \mu e \\ S \xi = \mu e \end{array} \right.$$

... follow central path for $\mu \rightarrow 0$ ... fails to exist ...

8.2. Robust Interior Point Methods

Perturb rhs of complementarity constraint

$$\left\{\begin{array}{ll} \min\limits_{x} & f(x) \\ \text{s.t.} & x_1 \geq 0 \\ & x_2 \geq 0 \\ & X_1 x_2 \leq C\mu e \end{array}\right. \quad \rightarrow \quad \left\{\begin{array}{l} \nabla f(x) - \begin{pmatrix} 0 \\ \nu_1 - X_2 \xi \\ \nu_2 - X_1 \xi \end{pmatrix} = 0 \\ X_1 x_2 + s = C\mu e \\ X_1 \nu_1 = \mu e \\ X_2 \nu_2 = \mu e \\ S\xi = \mu e \end{array}\right.$$

defines central path $(x(\mu), \nu(\mu), \xi(\mu))$ for barrier parameter $\mu > 0$.

[Raghunathan and Biegler, 2002] and [Liu and Sun, 2002]

8.2. Robust Interior Point Methods

Alternative: ℓ_1 penalty of complementarity constraint

$$\left\{ \begin{array}{ll} \min\limits_x & f(x) + \rho x_1^T x_2 \\ \text{s.t.} & x_1 \geq 0 \\ & x_2 \geq 0 \end{array} \right. \quad \rightarrow \quad \left\{ \begin{array}{l} \nabla f(x) - \begin{pmatrix} 0 \\ \nu_1 - \rho x_2 \\ \nu_2 - \rho x_1 \end{pmatrix} = 0 \\ X_1 \nu_1 = \mu e \\ X_2 \nu_2 = \mu e \end{array} \right.$$

- smooth, since IPM keeps $x_1, x_2 > 0$
- exact, for $\rho > \|\xi\|_\infty$ basic multiplier
- if necessary, increase ρ during IPM iteration (while reducing μ)

  ... joint work in progress with Nocedal & Lopez

8.3. 5 Red Herrings

1. Equivalent NLP violates Mangasarian Fromowitz CQ implies ...

   unbounded multipliers ... form a ray $\Rightarrow \exists$ bounded multipliers

   constraint gradients linearly dependent ... QP solver finds basis

   central path fails to exist ... perturb or penalize

   inconsistent linearizations ... perturb or penalize

8.3. 5 Red Herrings

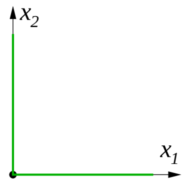
2. Perturbation $x_1^T x_2 \leq 0$ to $x_1^T x_2 \leq -\epsilon \Rightarrow$ inconsistent NLP ...

rhs "0" is structural zero

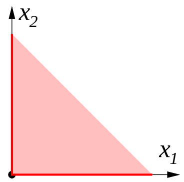
treat like structural zeros in sparse linear algebra ...

$\Rightarrow$ never perturbed to $-\epsilon$

3. Linearization at $(0, 0)$ is lousy approximation of tangent cone



tangent cone



NLP approx.

OK, if multipliers $\nu_1, \nu_2 > 0$ ... or SOSC

8.3. 5 Red Herrings

4. [Bard, 1988] experience with GRG (50-70% failure)
   modern solvers more robust/advanced ...

5. Failure of `LANCELOT` on `LUBRIF`
   mistakes in SIF file ... pressure & thickness mixed up
   $\Rightarrow$ model makes no sense physically!

8.3. 5 Red Herrings

4. [Bard, 1988] experience with GRG (50-70% failure)
   modern solvers more robust/advanced ...

5. Failure of `LANCELOT` on `LUBRIF`
   mistakes in SIF file ... pressure & thickness mixed up
   $\Rightarrow$ model makes no sense physically!
   no complementarity ... minimize $x_1^T x_2 \Rightarrow$ model has MFCQ

9. New NCP functions & formulations

Questions:

- Can we avoid infeasibility of QPs close to solution?
- Are there better formulations of MPCCs as NLPs?

Answer: Look at NCP functions!

$$
\begin{array}{lll}
& 0 \le x_1 \;\perp\; x_2 \ge 0 & \\
\Leftrightarrow & \min(x_1 \,, x_2) = 0 & \text{min-function} \\
\Leftrightarrow & x_1 + x_2 - \sqrt{x_1^2 + x_2^2} = 0 & \text{Fischer-Burmeister}
\end{array}
$$

Snag: NCP functions are non-smooth & nasty at $(0,0)$!

Idea: Keep bounds $x_i \ge 0$ & only use upper bound of NCP functions.

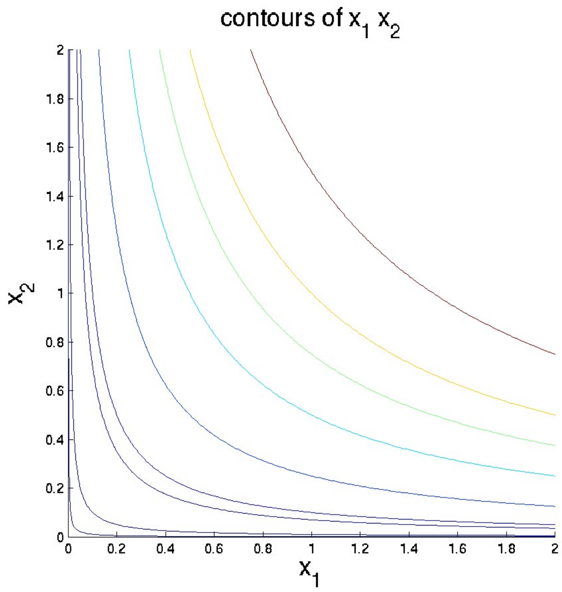
9.1. Scalar product function

$$0 \leq x_1 \perp x_2 \geq 0$$

$$\Leftrightarrow \quad x_1 \geq 0$$

$$x_2 \geq 0$$

$$x_{1i} x_{2i} \leq 0$$



9.2. Fischer-Burmeister function

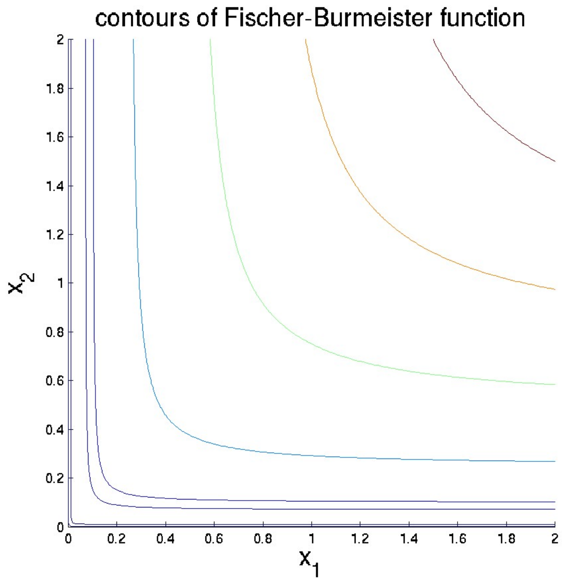
$$0 \le x_1 \;\perp\; x_2 \ge 0$$

$$\Leftrightarrow \quad x_1 \ge 0$$

$$x_2 \ge 0$$

$$x_{1i} + x_{2i} - \sqrt{x_{1i}^2 + x_{2i}^2} \le 0$$

beware at $(0, 0)$



9.3. min-function

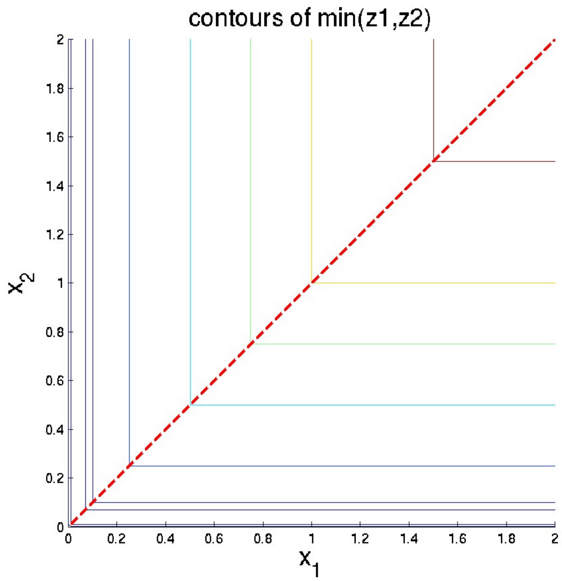
$$0 \le x_1 \;\perp\; x_2 \ge 0$$

$$\Leftrightarrow \quad x_1 \ge 0$$

$$x_2 \ge 0$$

$$\min(x_{1i}, x_{2i}) \le 0$$

large arbitrary jumps



9.4. linearized min-function

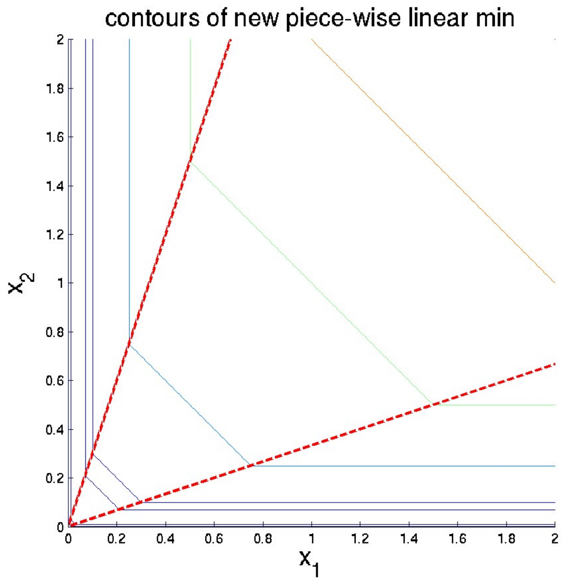
$$0 \leq x_1 \perp x_2 \geq 0$$

$$\Leftrightarrow \quad x_1 \geq 0$$

$$x_2 \geq 0$$

$$\psi_l(x_{1i}, x_{2i}) \leq 0$$

... or even smoother ...



9.5. quadratic min-function

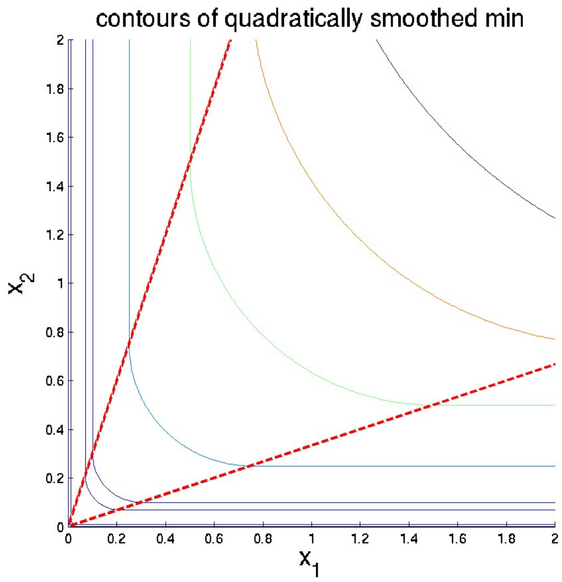
$$0 \leq x_1 \ \perp \ x_2 \geq 0$$

$$\Leftrightarrow \quad x_1 \geq 0$$

$$x_2 \geq 0$$

$$\psi_q(x_{1i}, x_{2i}) \leq 0$$

... readily computed



9.6 Properties of min-functions

$\min$-functions give consistent linearization close to solution!

$\equiv$ constraint qualification $\Rightarrow$ SQP converges without relaxation!

Important how to handle $(0,0)$

$$\min(x_1, x_2) \; = \; \begin{cases} 0 & \boxed{x_1 = x_2 = 0 \textbf{ !!!}} \\ x_2 & x_2 \le x_1 \\ x_1 & x_1 \le x_2 \end{cases}$$

... $\nabla \min(x_1, x_2)$ accordingly

... otherwise get trapped at $(0,0)$.

9.7. Properties of Fischer-Burmeister function

Fischer-Burmeister linearizations can be inconsistent $\Rightarrow$ relax

Gradients of Fischer-Burmeister bounded for $(x_{1i}, x_{2i}) \neq (0,0)$
$\Rightarrow$ special case at $(0,0)$.

Hessian unbounded near $(0,0)$ $\Rightarrow$ do not use Hessian ...
... OK, since no curvature in complementarity!

${x_1^{(k)}}^T x_2^{(k)} = 0$, i.e. on axis

$\Rightarrow$ SQP equivalent to SQP on relaxed NLP

$\Rightarrow$ SQP converges similar to scalar product formulation

... also used in [Facchinei et al., 1996].

9.8. Comparison of different formulations

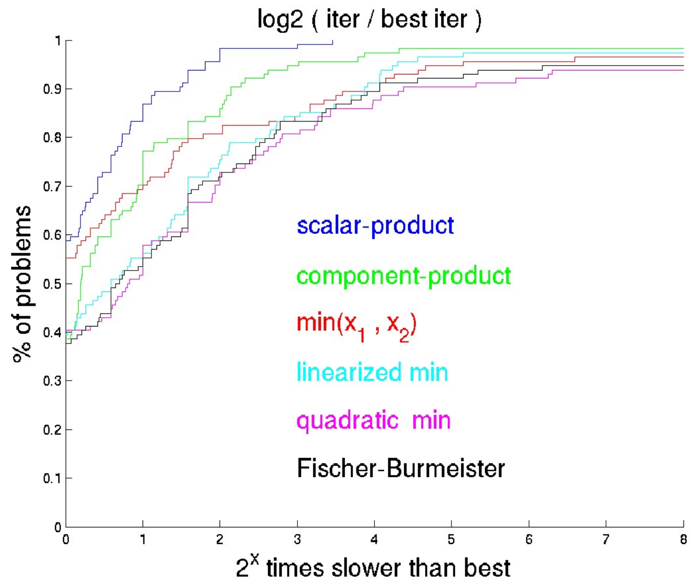
MacMPEC:

www.mcs.anl.gov/~leyffer/MacMPEC/

subset of 114 problems

6 formulations in filterSQP

$\Rightarrow x_1^T x_2 \leq 0$ best



10. Conclusion & Outlook: MPCCs

- MPCCs emerging optimization area: many applications
- NLP solvers work very well; supported by theory!
- Alternative formulations; $\min$ et al. and Fischer-Burmeister.

Open Questions:

- global convergence of NLP solvers?
- avoid convergence to "x-stationary" points
- need global minimizers for certain applications

References

[Anitescu, 2000] Anitescu, M. (2000). On solving mathematical programs with complementarity constraints as nonlinear programs. Preprint ANL/MCS-P864-1200, MCS Division, Argonne National Laboratory, Argonne, IL, USA.

[Anjos and Vanelli, 2002] Anjos, M. and Vanelli, A. (2002). A new mathematical programming framework for facility layout problems. Technical Report UW-E&CE#2002-04,, Department of Electrical & Computer Engineering, University of Waterloo,, Canada,.

[Bard, 1988] Bard, J. (1988). Convex two-level optimization. *Mathematical Programming*, 40(1):15–27.

[Facchinei et al., 1996] Facchinei, F., Jiang, H., and Qi, L. (1996). A smoothing method for mathematical programs with equilibrium constraints. Technical Report 03.96, Universita di Roma, "La Sapienza".

[Ferris et al., 1999] Ferris, M., Meeraus, A., and Rutherford, T. (1999). Computing Wardropian equilibrium in a complementarity framework. *Optimization Methods and Software*, 10:669–685.

[Ferris and Pang, 1997] Ferris, M. and Pang, J. (1997). Engineering and economic applications of complementarity problems. *SIAM Review*, 39(4):669–713.

[Ferris and Tin-Loi, 1999a] Ferris, M. and Tin-Loi, F. (1999a). Limit analysis of frictional block assemblies as a mathematical program with complementarity constraints. Mathematical Programming Technical Report 99-01, University of Wisconsin.

[Ferris and Tin-Loi, 1999b] Ferris, M. and Tin-Loi, F. (1999b). On the solution of a minimum weight elastoplastic problem involving displacement and complementarity constraints. *Computer Methods in Applied Mechanics and Engineering*, 174:107–120.

[Fletcher et al., 2002] Fletcher, R., Leyffer, S., Ralph, D., and Scholtes, S. (2002). Local convergence of SQP methods for mathematical programs with equilibrium constraints. Numerical Analysis Report NA/209, Department of Mathematics, University of Dundee, Dundee, DD1 4HN, UK.

[Hearn and Ramana, 1997] Hearn, D. and Ramana, M. (1997(?)). Solving congestion toll pricing models. Technical report, Department of Industrial and Systems Engineering, University of Florida, `http://www.ise.ufl.edu/hearn/crt.ps`.

[Hobbs et al., 2000] Hobbs, B., Metzler, C., and Pang, J.-S. (2000). Strategic gaming analysis for electric power systems: An mpec approach. *IEEE Transactions on Power Systems*, 15(2):638–645.

[Huang and Pang, 1999] Huang, J. and Pang, J.-S. (1999). A mathematical programming with equilibrium approach to the implied volatility surface of American options. Technical report, Dept. of Mathematical Sciences, The Johns Hopkins University, Baltimore, Maryland 21218-2682, USA.

[Liu and Sun, 2002] Liu, X. and Sun, J. (2002). Generalized stationary points and an interior point method for mathematical programs with equilibrium constraints. Preprint, Singapore, School of Business.

[Luo et al., 1996] Luo, Z.-Q., Pang, J.-S., and Ralph, D. (1996). *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press.

[Outrata et al., 1998] Outrata, J., Kocvara, M., and Zowe, J. (1998). *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*. Kluwer Academic Publishers, Dordrecht.

[Pieper, 2001] Pieper, H. (2001). *Algorithms for Mathematical Programs with Equilibrium Constraints with Applications to deregulated electricity markets*. PhD thesis, Department of Management Science and Engineering, Stanford University.

[Raghunathan and Biegler, 2002] Raghunathan, A. and Biegler, L. T. (2002). Mpec formulations and algorithms in process engineering. Technical report, CMU Chemical Engineering.

[Rico-Ramirez and Westerberg, 1999] Rico-Ramirez, V. and Westerberg, A. (1999). Conditional modeling. 2. solving using complementarity and boundary-crossing formulations. *Industrial Engineering and Chemistry Research*, 38:531–553.

[Sayin, 2000] Sayin, S. (2000). Measuring the quality of discrete representations of efficient sets in multiple objective mathematical programming. *Mathematical Programming*, 87(3):543–560.

[Scholtes, 2001] Scholtes, S. (2001). Convergence properties of regularization schemes for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 11(4):918–936.

[Stackelberg, 1952] Stackelberg, H. V. (1952). *The Theory of Market Economy*. Oxford University Press.

[Tin-Loi and Que, 2002] Tin-Loi, F. and Que, N. (2002). Nonlinear programming approaches for an inverse problem in quasibrittle fracture. *Mechanical Sciences*, 44:843–858.

[Wilmott et al., 1993] Wilmott, P., Dewyne, J., and Howison, S. (1993). *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press.