An Extended Mathematical Programming Framework

Michael C. Ferris, Steven Dirkse, Jan Jagla, Alex Meeraus

University of Wisconsin, Madison

IE&OR, Berkeley: April 27, 2009

Optimal Power Flow

OPF(α):	$\min_y$	energy dispatch cost (y, α)
	s.t.	conservation of power flow at nodes
		Kirchoff's voltage law, and simple bound constraints

α are (given) price bids, parametric optimization

Optimal Power Flow

$$\begin{array}{lll} \text{OPF}(\alpha): & \min_y & \text{energy dispatch cost } (y, \alpha) \\ & \text{s.t.} & \text{conservation of power flow at nodes} \\ & & \text{Kirchoff's voltage law, and simple bound constraints} \end{array}$$

α are (given) price bids, parametric optimization

$$\begin{array}{lll} \text{Leader}(\bar{\alpha}_{-i}): & \max_{\alpha_i, y, \lambda} & \text{firm } i\text{'s profit } (\alpha_i, y, \lambda) \\ & \text{s.t.} & 0 \leq \alpha_i \leq \hat{\alpha}_i \\ & & y \text{ solves OPF}(\alpha_i, \bar{\alpha}_{-i}) \end{array}$$

Note that objective involves multiplier from OPF problem

Optimal Power Flow

$$\text{OPF}(\alpha): \quad \min_y \quad \text{energy dispatch cost } (y, \alpha)$$
$$\text{s.t.} \quad \text{conservation of power flow at nodes}$$
$$\text{Kirchoff's voltage law, and simple bound constraints}$$

α are (given) price bids, parametric optimization

$$\text{Leader}(\bar{\alpha}_{-i}): \quad \max_{\alpha_i, y, \lambda} \quad \text{firm } i\text{'s profit } (\alpha_i, y, \lambda)$$
$$\text{s.t.} \quad 0 \leq \alpha_i \leq \hat{\alpha}_i$$
$$y \text{ solves OPF}(\alpha_i, \bar{\alpha}_{-i})$$

Note that objective involves multiplier from OPF problem

$$\text{Leader}(\bar{\alpha}_{-i}): \quad \max_{\alpha_i, y, \lambda} \quad \text{firm } i\text{'s profit } (y, \lambda, \alpha)$$
$$\text{s.t.} \quad 0 \leq \alpha_i \leq \bar{\alpha}_i$$
$$y, \lambda \text{ solves KKT(OPF}(\alpha_i, \bar{\alpha}_{-i}))$$

This is an example of an MPCC since KKT form complementarity constraints

Multi-player and security constraints

- $(\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_m)$ is an equilibrium if

$$\bar{\alpha}_i \text{ solves Leader}(\bar{\alpha}_{-i}), \quad \forall i$$

- (Nonlinear) Nash Equilibrium where each player solves an MPCC

Multi-player and security constraints

- $(\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_m)$ is an equilibrium if

$$\bar{\alpha}_i \text{ solves Leader}(\bar{\alpha}_{-i}), \quad \forall i$$

- (Nonlinear) Nash Equilibrium where each player solves an MPCC
- MPCC is hard (lacks a constraint qualification)
- Nash Equilibrium is PPAD-complete (Chen et al, Papadimitriou et al)
- In practice, also require contingency (scenario) constraints imposed in the OPF problem

- Leader/follower game: Stackleberg
- Supply chains with "market leader"

Biological Pathway Models

Opt knock (a bilevel program)

max bioengineering objective (through gene knockouts)

s.t. max cellular objective (over fluxes)

s.t. fixed substrate uptake

network stoichiometry

blocked reactions (from outer problem)

number of knockouts $\leq$ limit

Biological Pathway Models

Opt knock (a bilevel program)

max bioengineering objective (through gene knockouts)
s.t. max cellular objective (over fluxes)
s.t. fixed substrate uptake
network stoichiometry
blocked reactions (from outer problem)
number of knockouts $\leq$ limit

Also prediction models of the form:

$$\begin{aligned} \min \sum_{i,j} \|w_i - v_j\| \\ \text{s.t. } Sv = w \\ -\bar{v}_L \leq v \leq \bar{v}_U, \; w_j = \bar{w}_j \end{aligned}$$

Can be modeled as an SOCP.

Mathematical programming: modeling

- Optimization models improve understanding of underlying systems and facilitate operational/strategic improvements
- Modeling systems enable application interfacing, prototyping of optimization capability
- Problem format is old/traditional

$$\min_x f(x) \text{ s.t. } g(x) \leq 0, h(x) = 0$$

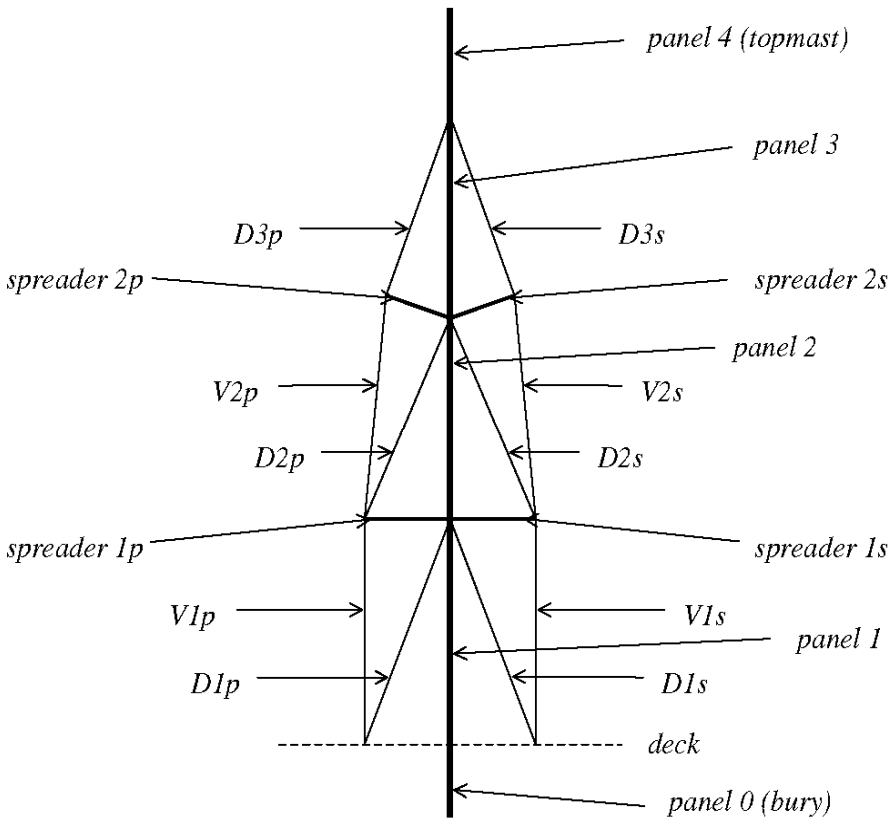
  - Support for integer, sos, semicontinuous variables
  - Limited support for logical constructs
  - Support for complementarity constraints

Optimal Yacht Rig Design

- Current mast design trends use a large primary spar that is supported laterally by a set of tension and compression members, generally termed the rig
- Reduction in either the weight of the rig or the height of the VCG will improve performance
- Design must work well under a variety of weather conditions

Complementarity feature

- Stays are tension only members (in practice) - Hookes Law
- When tensile load becomes zero, the stay goes slack (low material stiffness)
- $0 \geq s \perp s - k\delta \leq 0$
  - s axial load
  - k member spring constant
  - δ member extension
- Either $s_i = 0$ or $s_i = k\delta_i$



MPCC: complementarity constraints

$$\begin{array}{ll} \min\limits_{x,s} & f(x,s) \\ \text{s.t.} & g(x,s) \leq 0, \\ & 0 \leq s \perp h(x,s) \geq 0 \end{array}$$

- g, h model "engineering" expertise: finite elements, etc
- $\perp$ models complementarity, disjunctions
- Complementarity "$\perp$" constraints available in AIMMS, AMPL and GAMS

MPCC: complementarity constraints

$$\begin{array}{ll} \min\limits_{x,s} & f(x,s) \\ \text{s.t.} & g(x,s) \leq 0, \\ & 0 \leq s \perp h(x,s) \geq 0 \end{array}$$

- g, h model "engineering" expertise: finite elements, etc
- $\perp$ models complementarity, disjunctions
- Complementarity "$\perp$" constraints available in AIMMS, AMPL and GAMS
- NLPEC: use the convert tool to automatically reformulate as a parameteric sequence of NLP's
- Solution by repeated use of standard NLP software
  - Problems solvable, local solutions, hard
  - Southern Spars Company (NZ): improved from 5-0 to 5-2 in America's Cup!

Use of complementarity

- Pricing electricity markets and options
- Video games: model contact problems
  - Friction only occurs if bodies are in contact
- Structure design
  - how springy is concrete
  - optimal sailboat rig design
- Computer/traffic networks
  - The price of anarchy measures difference between "system optimal" (MPCC) and "individual optimization" (MCP)

- Complementarity facilitates modeling of competition, nonsmoothness and "switching"
- Large scale models involving complementarity now solvable
- Do you (or should you) care?

EMP(i): Variational inequalities

- Find $z \in C$ such that

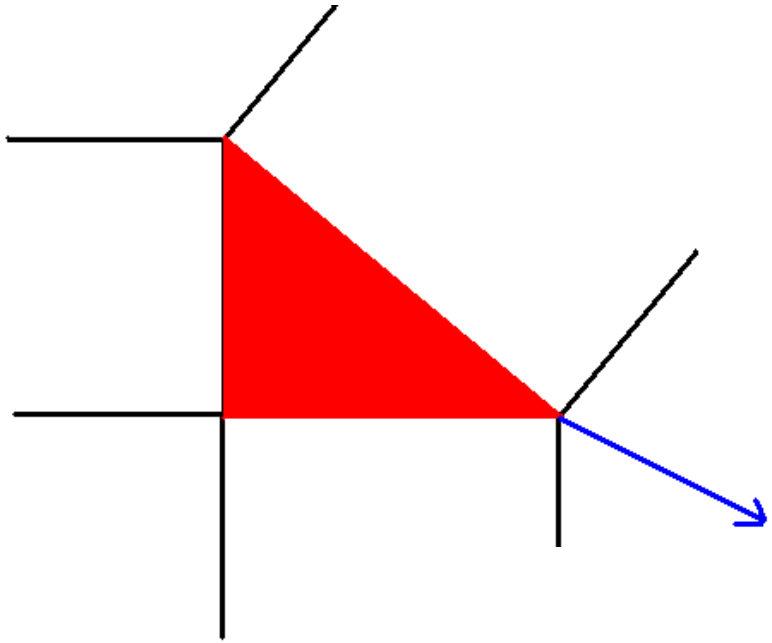
$$\langle F(z), y - z \rangle \geq 0, \quad \forall y \in C$$

- Many applications where F is not the derivative of some f
- model vi / F, g /;
  empinfo: vifunc F z
- Convert problem into complementarity problem by introducing multipliers on representation of C
- Can now do MPEC (as opposed to MPCC)!
- Projection algorithms, robustness (evaluate F only at points in C)

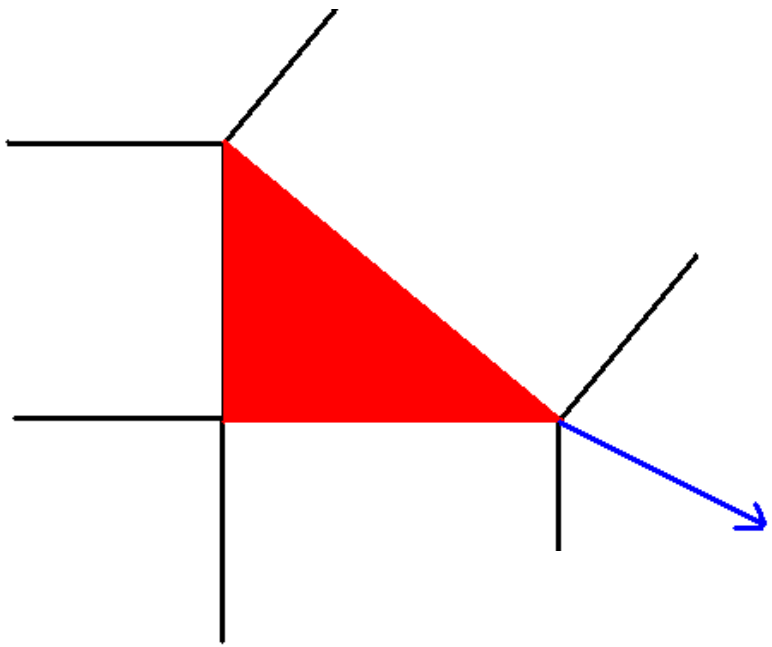
Normal map for polyhedral C

projection: $\pi_C(x)$

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



Normal map for polyhedral C



projection: $\pi_C(x)$

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

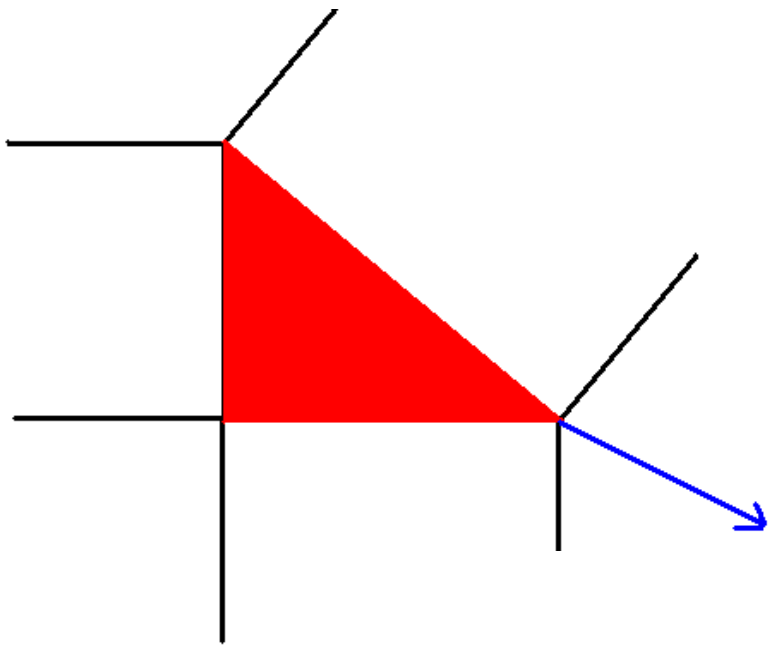
If $-M\pi_C(x) - q = x - \pi_C(x)$ then

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

so $z = \pi_C(x)$ solves

$$\langle -Mz - q, y - z \rangle \leq 0, \quad \forall y \in C$$

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

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

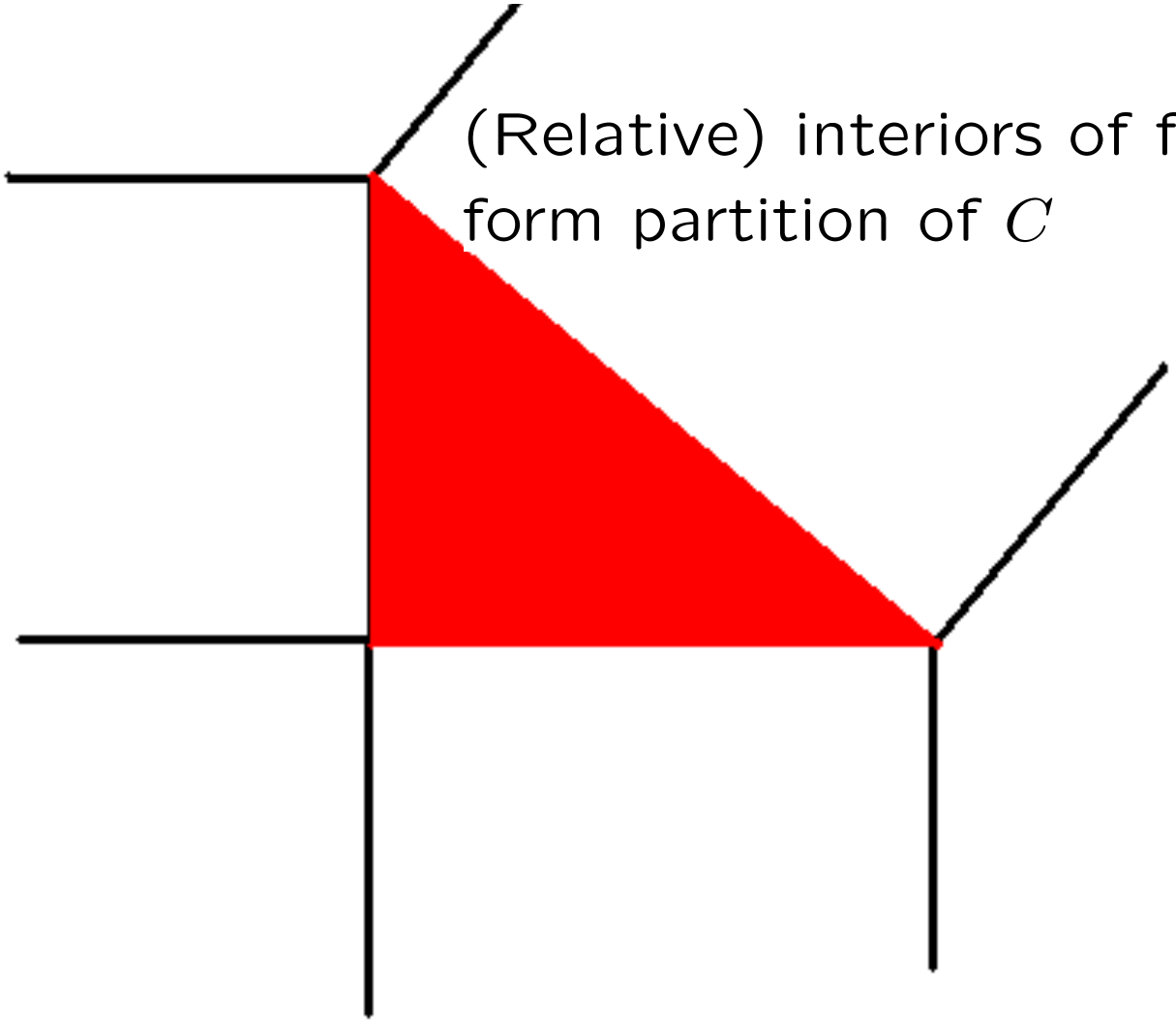
so $z = \pi_C(x)$ solves

$$\langle -Mz - q, y - z \rangle \leq 0, \quad \forall y \in C$$

Find x, a zero of the normal map:

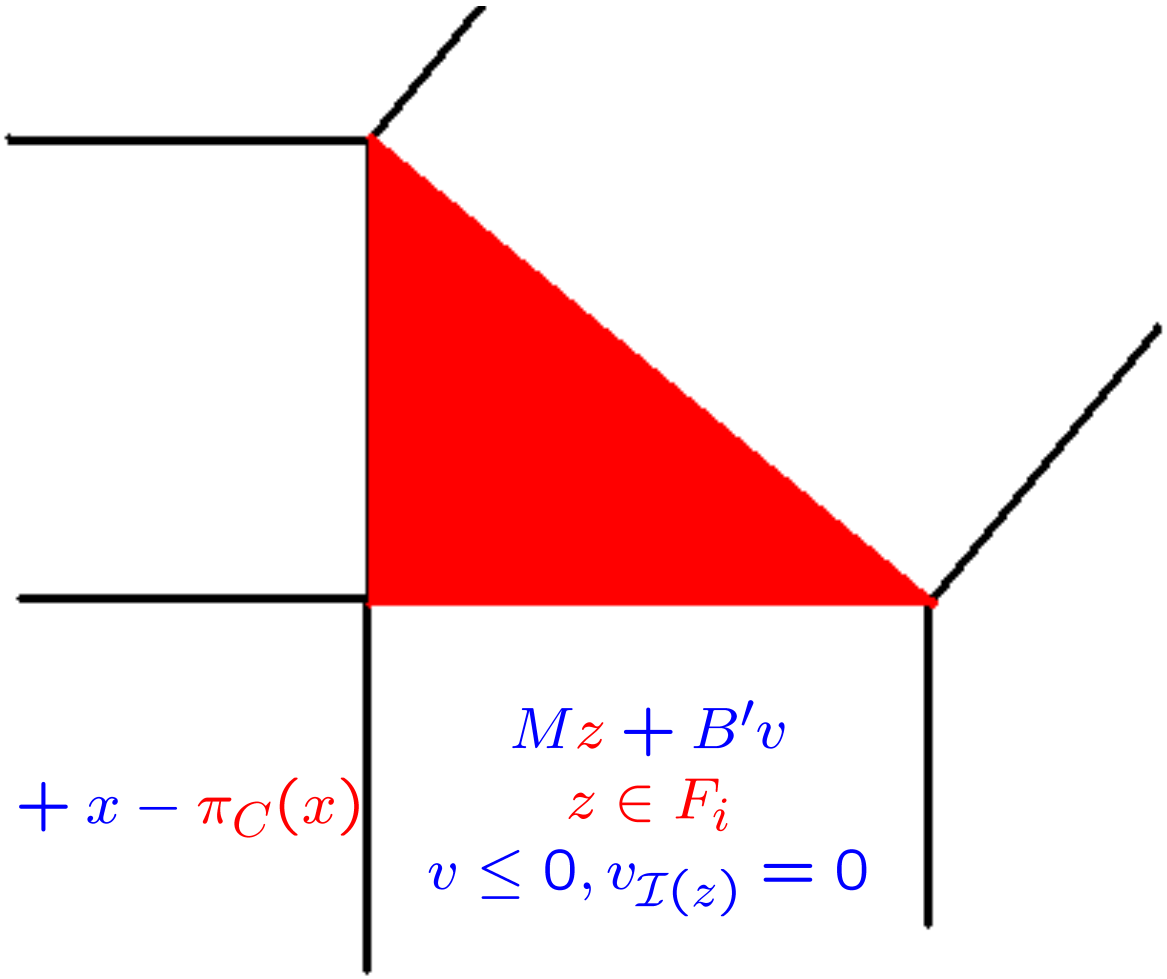
$$0 = M\pi_C(x) + q + x - \pi_C(x)$$

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



(Relative) interiors of faces F_i form partition of C

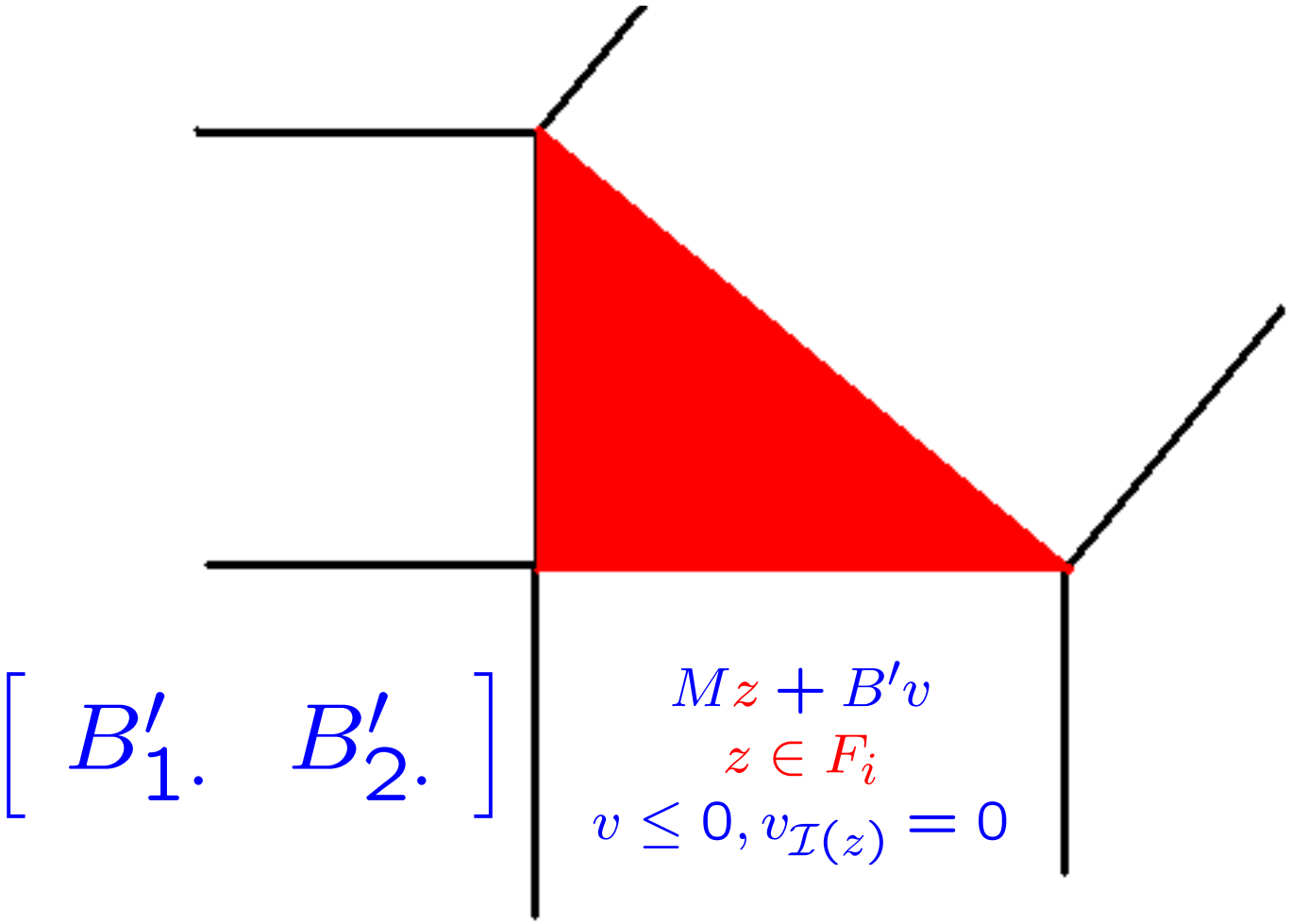
$$C = \{z | Bz \geq b\}, N_C(z) = \{B'v | v \leq 0, v_{\mathcal{I}(z)} = 0\}$$



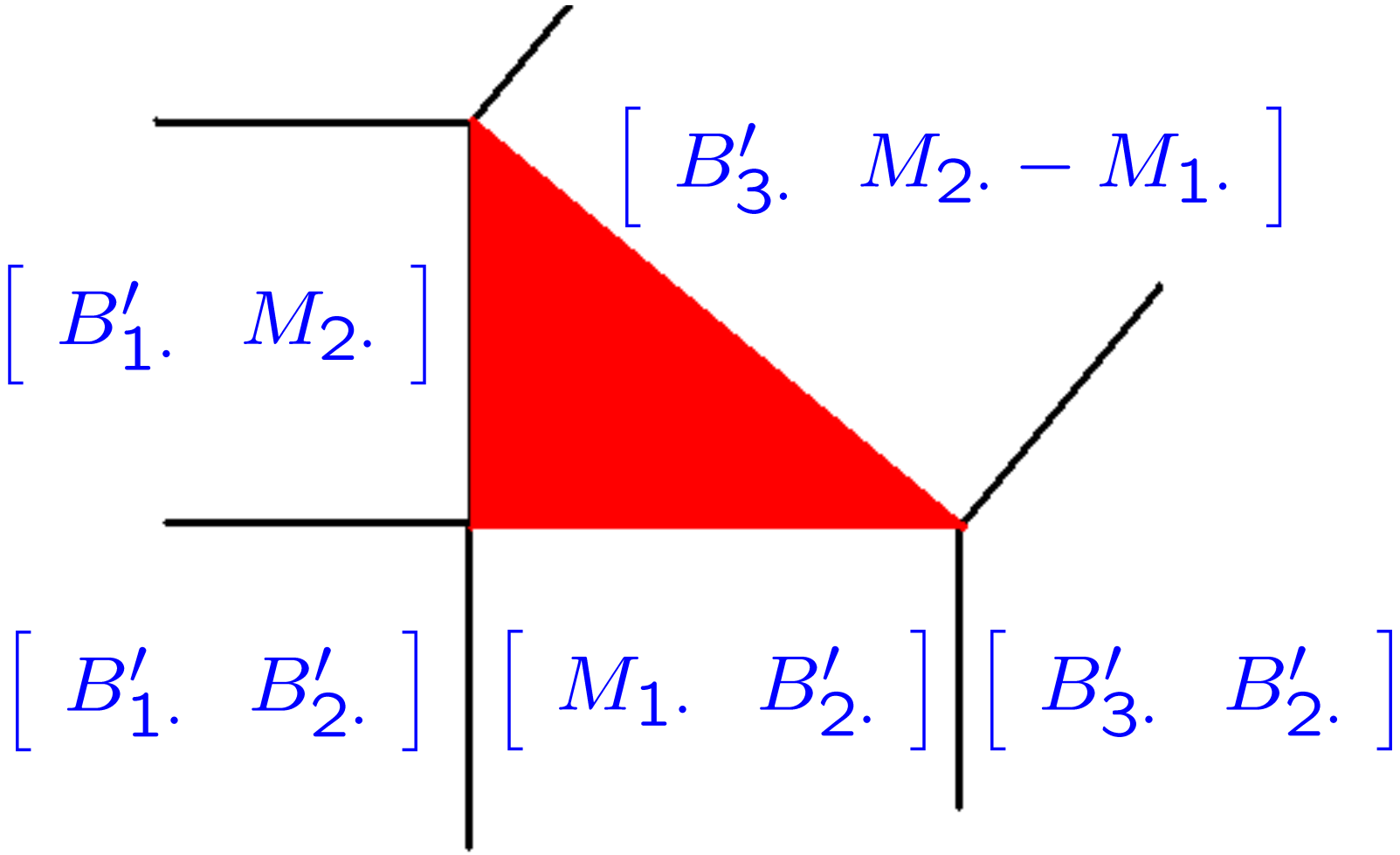
$$M\pi_C(x) + x - \pi_C(x)$$

$$Mz + B'v$$
$$z \in F_i$$
$$v \leq 0, v_{\mathcal{I}(z)} = 0$$

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

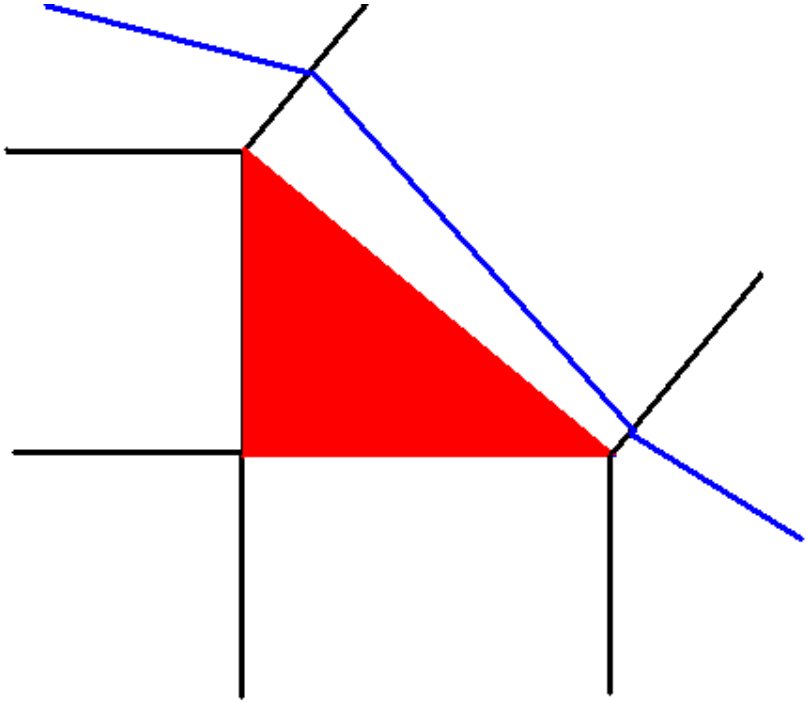


$C = \{z | Bz \geq b\}, N_C(z) = \{B'v | v \leq 0, v_{\mathcal{I}(z)} = 0\}$



Cao/Ferris Path (Eaves)

- Start in cell that has interior (face is an extreme point)
- 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)

Extensions and Computational Results

- Embed AVI solver in a Newton Method - each Newton step solves an AVI
- Compare performance of PathAVI with PATH on equivalent LCP
- PATH the most widely used code for solving MCP
- AVIs constructed to have solution with $M_{n\times n}$ symmetric indefinite

	PathAVI		*PATH*	
Size (m,n)	*Resid*	*Iter*	*Resid*	*Iter*
$(180, 60)$	3×10^{-14}	193	0.9	10176
$(360, 120)$	3×10^{-14}	1516	2.0	10594

- 2 - 10x speedup in Matlab using sparse LU instead of QR
- 2 - 10x speedup in C using sparse LU updates

Complementarity Problems in Economics (MCP)

- p represents prices, x represents activity levels
- System model: given prices, (agent) i determines activities x_i

$$G_i(x_i, x_{-i}, p) = 0$$

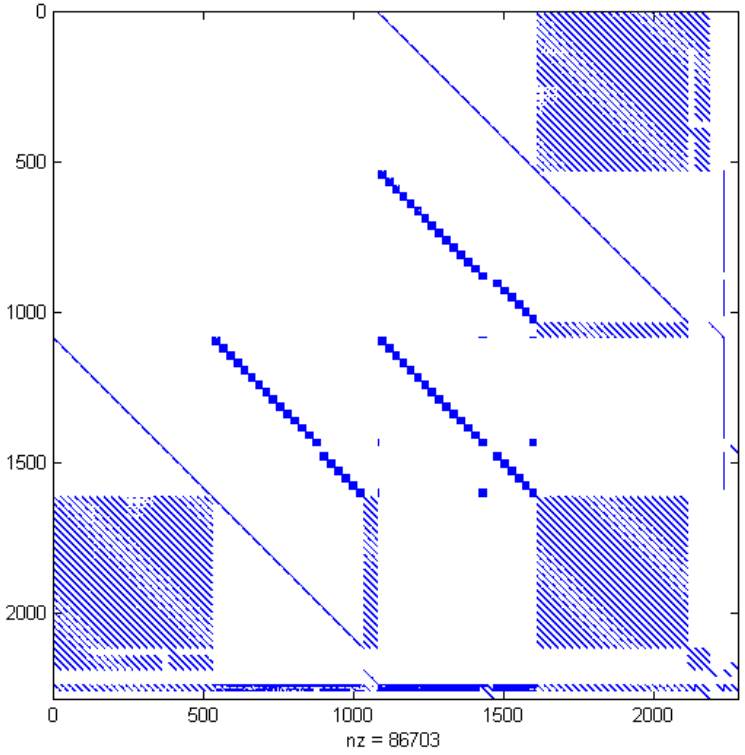
  x_{-i} are the decisions of other agents.
- Walras Law: market clearing

$$0 \leq S(x, p) - D(x, p) \ \perp \ p \geq 0$$

- Key difference: optimization assumes you control the complete system
- Complementarity determines what activities run, and who produces what

World Bank Project (Uruguay Round)

- 24 regions, 22 commodities
  - Nonlinear complementarity problem
  - Size: 2200 x 2200
- Short term gains $53 billion p.a.
  - Much smaller than previous literature
- Long term gains $188 billion p.a.
  - Number of less developed countries loose in short term
- Unpopular conclusions - forced concessions by World Bank
- Region/commodity structure not apparent to solver



EMP(ii): Embedded models

- Model has the format:

$$\text{Agent o:} \quad \min_x \quad f(x,y)$$
$$\text{s.t.} \quad g(x,y) \le 0 \quad (\perp \lambda \ge 0)$$

$$\text{Agent v:} \quad H(x,y,\lambda) = 0 \qquad (\perp y \text{ free})$$

- Difficult to implement correctly (multiple optimization models)
- Can do automatically - simply annotate equations
  empinfo: equilibrium
  min f x defg
  vifunc H y dualvar λ defg
- EMP tool automatically creates an MCP

$$\nabla_x f(x,y) + \lambda^T \nabla g(x,y) = 0$$
$$0 \le -g(x,y) \perp \lambda \ge 0$$
$$H(x,y,\lambda) = 0$$

Nash Equilibria

- Nash Games: x^* is a Nash Equilibrium if

$$x_i^* \in \arg\min_{x_i \in X_i} \ell_i(x_i, x_{-i}^*, q), \forall i \in \mathcal{I}$$

  x_{-i} are the decisions of other players.
- Quantities q given exogenously, or via complementarity:

$$0 \leq H(x, q) \ \perp \ q \geq 0$$

- empinfo: equilibrium
  min loss(i) x(i) cons(i)
  vifunc H q
- Applications: Discrete-Time Finite-State Stochastic Games. Specifically, the Ericson & Pakes (1995) model of dynamic competition in an oligopolistic industry.

Key point: models generated correctly solve quickly

Here S is mesh spacing parameter

S	Var	rows	non-zero	dense(%)	Steps	RT (m:s)
20	2400	2568	31536	0.48	5	0 : 03
50	15000	15408	195816	0.08	5	0 : 19
100	60000	60808	781616	0.02	5	1 : 16
200	240000	241608	3123216	0.01	5	5 : 12

Convergence for $S = 200$ (with new basis extensions in PATH)

Iteration	Residual
0	1.56(+4)
1	1.06(+1)
2	1.34
3	2.04(−2)
4	1.74(−5)
5	2.97(−11)

General Equilibrium models

$$(C) : \max_{x_k \in X_k} U_k(x_k) \text{ s.t. } p^T x_k \leq i_k(y, p)$$

$$(I) : i_k(y, p) = p^T \omega_k + \sum_j \alpha_{kj} p^T g_j(y_j)$$

$$(P) : \max_{y_j \in Y_j} p^T g_j(y_j)$$

$$(M) : \max_{p \geq 0} p^T \left(\sum_k x_k - \sum_k \omega_k - \sum_j g_j(y_j) \right) \text{ s.t. } \sum_l p_l = 1$$

General Equilibrium models

$$(C): \max_{x_k \in X_k} U_k(x_k) \text{ s.t. } p^T x_k \leq i_k(y,p)$$

$$(I): i_k(y,p) = p^T \omega_k + \sum_j \alpha_{kj} p^T g_j(y_j)$$

$$(P): \max_{y_j \in Y_j} p^T g_j(y_j)$$

$$(M): \max_{p \geq 0} p^T \left(\sum_k x_k - \sum_k \omega_k - \sum_j g_j(y_j) \right) \text{ s.t. } \sum_l p_l = 1$$

Can reformulate as embedded problem (Ermoliev et al):

$$\max_{x \in X, y \in Y} \sum_k \frac{t_k}{\beta_k} \log U_k(x_k)$$

$$\text{s.t. } \sum_k x_k \leq \sum_k \omega_k + \sum_j g_j(y_j)$$

$t_k = i_k(y,p)$ where p is multiplier on NLP constraint

Sequential Joint Maximization

$$\max_{x \in X, y \in Y} \quad \sum_k \frac{t_k}{\beta_k} \log U_k(x_k)$$

$$\text{s.t.} \quad \sum_k x_k \leq \sum_k \omega_k + \sum_j g_j(y_j)$$

$t_k = i_k(y, p)$ where p is multiplier on NLP constraint

- Embedded model often solves faster as an MCP than the original MCP from Nash game
- Can exploit structure to improve computational performance further

Sequential Joint Maximization

$$\max_{x \in X, y \in Y} \quad \sum_k \frac{t_k}{\beta_k} \log U_k(x_k)$$
$$\text{s.t.} \quad \sum_k x_k \leq \sum_k \omega_k + \sum_j g_j(y_j)$$

$t_k = i_k(y, p)$ where p is multiplier on NLP constraint

- Embedded model often solves faster as an MCP than the original MCP from Nash game
- Can exploit structure to improve computational performance further
- Can iterate (on m) $t_k^m = i_k(y^m, p^m)$, and solve sequence of NLP's

$$\max_{x \in X, y \in Y} \quad \sum_k \frac{t_k^m}{\beta_k} \log U_k(x_k)$$
$$\text{s.t.} \quad \sum_k x_k \leq \sum_k \omega_k + \sum_j g_j(y_j)$$

EMP(iii): Heirarchical models

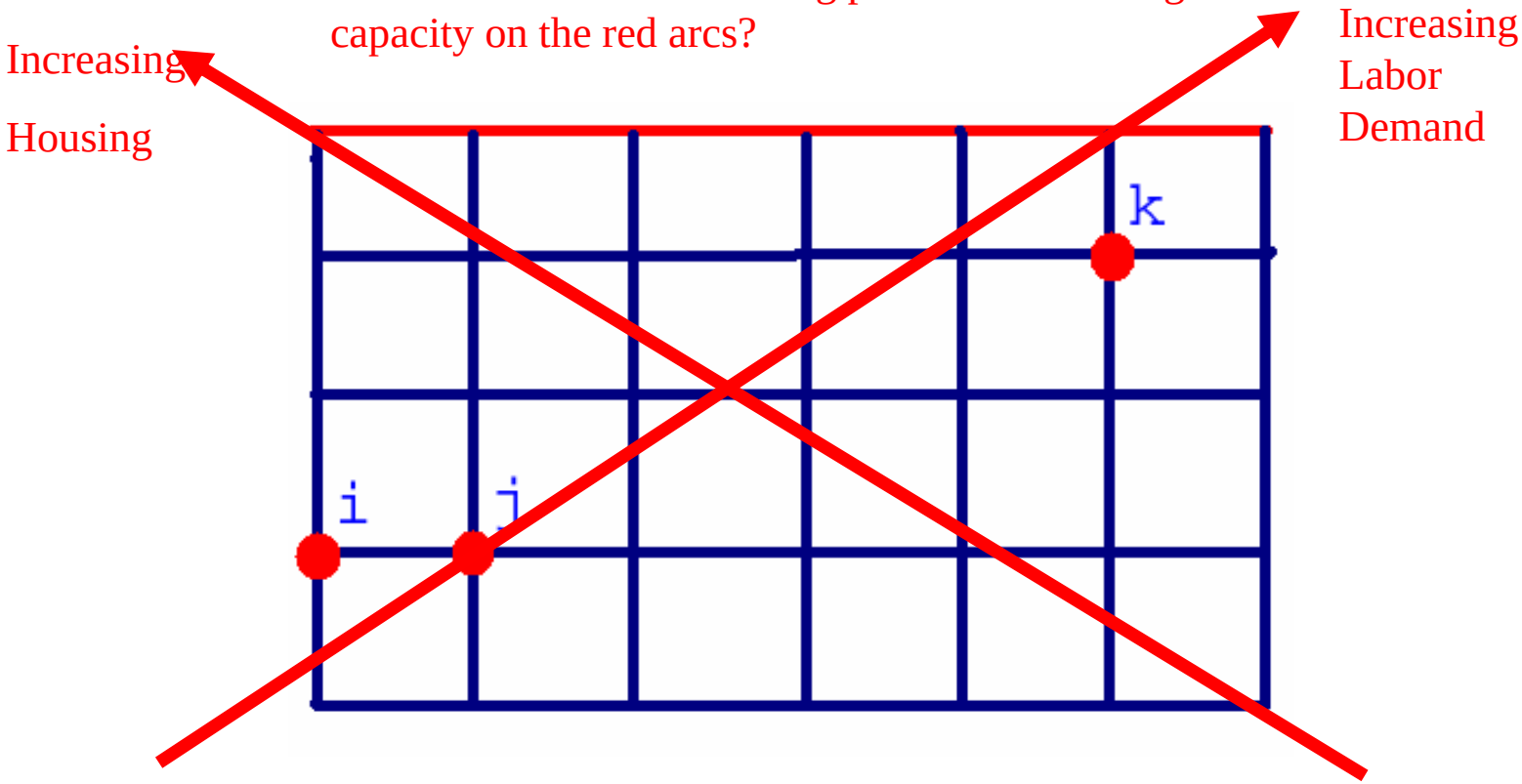
- Bilevel programs:

$$\begin{array}{ll} \min\limits_{x^*,y^*} & f(x^*,y^*) \\ \text{s.t.} & g(x^*,y^*) \leq 0, \\ & y^* \text{ solves } \min\limits_{y} v(x^*,y) \text{ s.t. } h(x^*,y) \leq 0 \end{array}$$

- model bilev /deff,defg,defv,defh/;
  empinfo: bilevel min v y defh
- EMP tool automatically creates the MPCC
- Note that heirarchical structure is available to solvers for decomposition approaches

Walras meets Wardrop

What is the effect on housing prices of increasing capacity on the red arcs?



Features

- We buy a house to "optimize" some measure
  - Price driven by market
  - We compete against each other
- Driver's choose routes to "optimize" travel time
  - Choices affect congestion
  - Your choice affects me!
- Production processes are "optimized"
- But the road designer does not control any of these!

$$\max_{x_k \in X_k} U_k(x_k, d)$$

$$H(x, p) = 0$$

$$\min_{t \in T} c(t, x, d)$$

$$\max_{y_j \in Y_j} p^T g_j(y_j)$$

$$\min_{d} f(d, x, p, t, y)$$

EMP(iv): Other new types of constraints

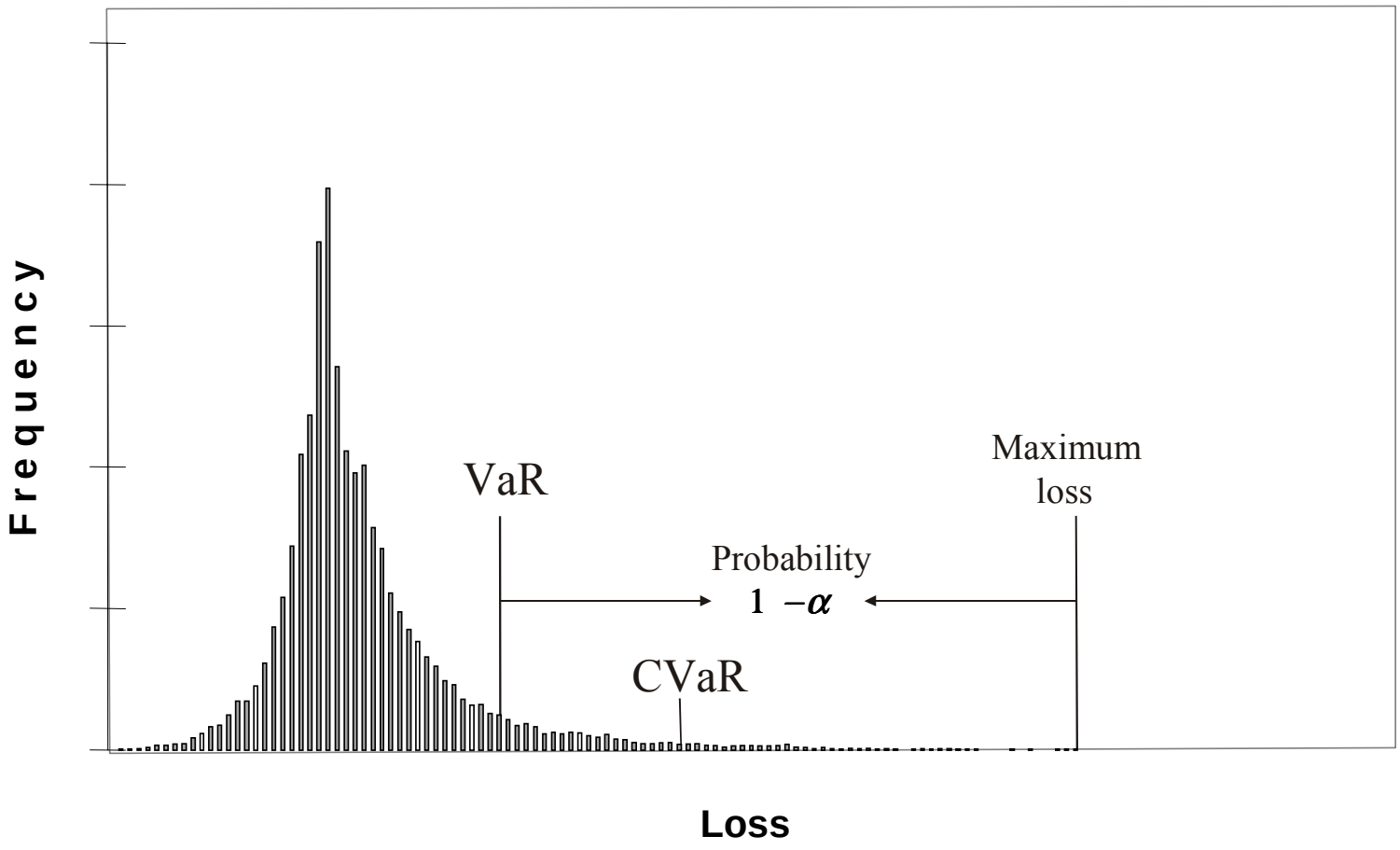
- range constraints $L \leq Ax - b \leq U$
- indicator constraints
- disjunctive programming
- soft constraints
- rewards and penalties
- robust programming (probability constraints, stochastics)

$$f(x, \xi) \leq 0, \forall \xi \in \mathcal{U}$$

- conic programming $a_i^T x - b_i \in K_i$

These constraints can be reformulated using EMP

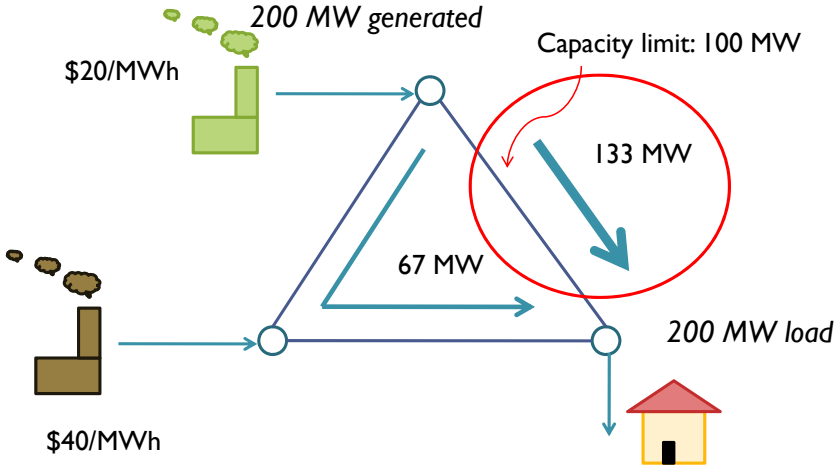
CVaR constraints: mean excess dose (radiotherapy)



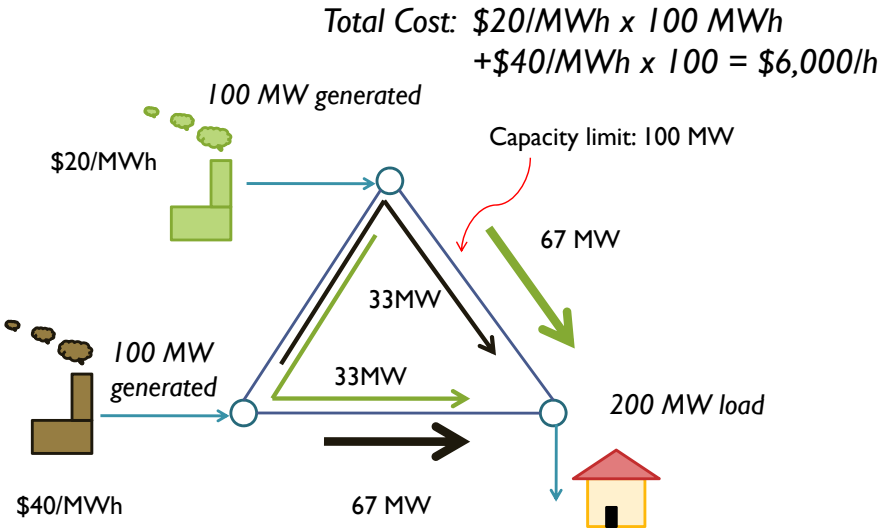
Move mean of tail to the left!

Transmission switching

Opening lines in a transmission network can reduce cost



(a) Infeasible due to line capacity



(b) Feasible dispatch

Need to use expensive generator due to power flow characteristics and capacity limit on transmission line

The basic model

$\min_{g,f,\theta}$	$c^T g$	generation cost
s.t.	$g - d = Af, f = BA^T\theta$	A is node-arc incidence
	$\bar{\theta}_L \leq \theta \leq \bar{\theta}_U$	bus angle constraints
	$\bar{g}_L \leq g \leq \bar{g}_U$	generator capacities
	$\bar{f}_L \leq f \leq \bar{f}_U$	transmission capacities

with transmission switching (within a smart grid technology) we modify as:

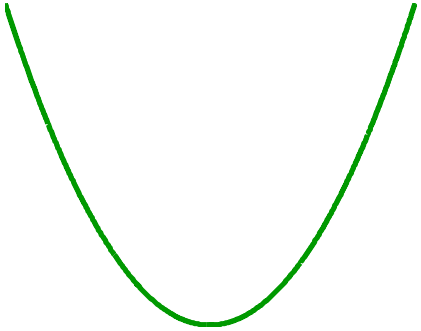
$\min_{g,f,\theta}$	$c^T g$	
s.t.	$g - d = Af$	
	$\bar{\theta}_L \leq \theta \leq \bar{\theta}_U$	
	$\bar{g}_L \leq g \leq \bar{g}_U$	
either	$f_i = (BA^T\theta)_i, \bar{f}_{L,i} \leq f_i \leq \bar{f}_{U,i}$	if i closed
or	$f_i = 0$	if i open

Use EMP to facilitate the disjunctive constraints (several equivalent formulations, including LPEC)

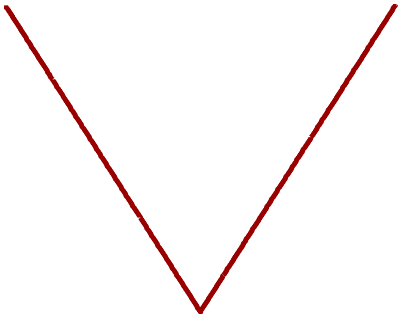
EMP(v): Extended nonlinear programs

$$\min_{x \in X} f_0(x) + \theta(f_1(x), \ldots, f_m(x))$$

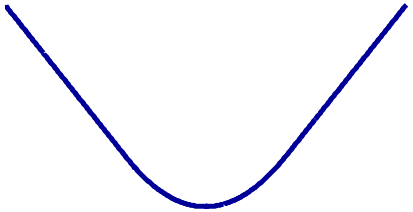
Examples of different θ



least squares,



absolute value,



Huber function

Solution reformulations are very different

Huber function used in robust statistics.

Key point for our work (Rockafellar)

- For many interesting choices of θ, the conjugate θ^* is of the form $k(y^*) + I_{Y^*}(y^*)$, where k is nice (e.g., C^2) and Y^* is closed convex, as is X; often these have simple structure
- Then we can deal with this problem by solving first-order conditions for a saddle point problem over $X \times Y^*$ rather than as a nonsmooth minimization problem
- The new feature here is implementation and solution within the GAMS modeling language framework, which produces a tool usable without advanced knowledge in convex analysis and without cumbersome "hand tailoring" to accommodate different penalizations [Ferris, Dirkse, Jagla, and Meeraus 2008]
- This makes the theoretical benefits accessible to users from a wide variety of different fields

Choices for θ

$$\inf_{x \in X} f_0 + \theta[f(x)], \quad \theta(u) = \sup_{y \in Y} \{y^T u - k(y)\}$$

θ is convex with values in $(-\infty, +\infty]$; may be nonsmooth

- L_2: $k(u) = \frac{1}{4\lambda} u^2$, $Y = (-\infty, +\infty)$
- L_1: $k(u) = 0$, $Y = [-\rho, \rho]$
- L_∞: $k(u) = 0$, $Y = \Delta$, unit simplex in $\mathbf{R}^m_+$
- Linear-quad (Huber 1981): $k(u) = \frac{1}{4\lambda} u^2$, $Y = [-\rho, \rho]$
- Second order cone constraint: $k(y) = 0$, $Y = C^\circ$

Solution Procedures

- Solution uses reformulation - one way: first order conditions

$$\mathrm{VI}\left(\begin{bmatrix} \nabla_x \mathcal{L}(x,y) \\ -\nabla_y \mathcal{L}(x,y) \end{bmatrix}, X \times Y\right)$$

based on extended form of the Lagrangian:

$$\mathcal{L}(x,y) = f_0(x) + \sum_{i=1}^{m} y_i f_i(x) - k(y)$$

- EMP: allows "annotation" of constraints to facilitate library of different θ functions to be applied
- EMP tool automatically creates an MCP (or a smooth NLP)
- Available!
- To do: extend solvers to exploit X and Y beyond simple bound sets

Conclusions

- Model is clearer, structure available to solver
- Large scale complementarity problems reliably solvable
- Complementarity and VI constraints within optimization problems
- Extended Mathematical Programming available within a modeling system
- System can easily formulate and solve second order cone programs, risk measures, robust optimization, soft constraints via piecewise linear penalization (with strong supporting theory)
- Embedded optimization models automatically reformulated for appropriate solution engine
- Enhance library of (implemented) θ functions (what do you want?)
- Exploit structure in solvers
- Extend application usage of complementarity solvers
- Extend complementarity solvers to VI solvers