Solution of equilibrium problems using extended mathematical programming

Youngdae Kim Michael C. Ferris

University of Wisconsin, Madison

2017 Bay Optimization Meeting, Davis, California, May 13, 2017

The good news!

- PATH solves rectangular VI

$$-F(x) \in N_{\mathcal{I}_1 \times \ldots \times \mathcal{I}_m}(x)$$

(feasible set is a Cartesian product of possibly unbounded intervals)

- PATHVI solves VI

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

by identifying

$$\mathcal{C} = \{x \in P : g(x) \in K\}$$

and reformulating as

$$\begin{aligned} x^* \text{ solves } \mathrm{VI}(F, \mathcal{C}) &\iff 0 \in F(x^*) + N_{\mathcal{C}}(x^*) \\ &\iff 0 \in \begin{bmatrix} F(x^*) + \nabla g(x^*)\lambda \\ -g(x^*) \end{bmatrix} + N_{P \times K^\circ}(x^*, \lambda) \end{aligned}$$

- Use Newton method, each step solves an affine variational inequality

Experimental results: AVI vs MCP

- Run PathAVI over AVI formulation.
- Run PATH over rectangular form (poorer theory as $\mathrm{rec}\mathcal{C}$ larger).
  - M is an $n \times n$ symmetric positive definite/indefinite matrix.
  - A has m randomly generated bounded inequality constraints.
- Structure knowledge leads to improved reliability

(m, n)	PathAVI		PATH		% negative eigenvalues
	status	# iterations	status	# iterations	
(180,60)	S	55	S	72	0
(180,60)	S	45	S	306	20
(180,60)	S	2	F	9616	60
(180,60)	S	1	F	10981	80
(360,120)	S	124	S	267	0
(360,120)	S	55	S	1095	20
(360,120)	S	2	F	10020	60
(360,120)	S	1	F	7988	80

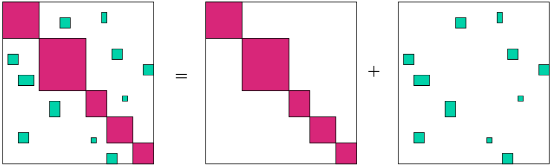
MOPEC

$$\min_{x_i} \theta_i(x_i, x_{-i}, \pi) \text{ s.t. } g_i(x_i, x_{-i}, \pi) \leq 0, \forall i$$

$$\pi \text{ solves } \mathrm{VI}(h(x, \cdot), C)$$

```
equilibrium
min theta(1) x(1) g(1)
...
min theta(m) x(m) g(m)
vi h pi cons
```

- (Generalized) Nash
- Reformulate optimization problem as first order conditions (complementarity)
- Use nonsmooth Newton methods to solve
- Solve overall problem using "individual optimizations"?



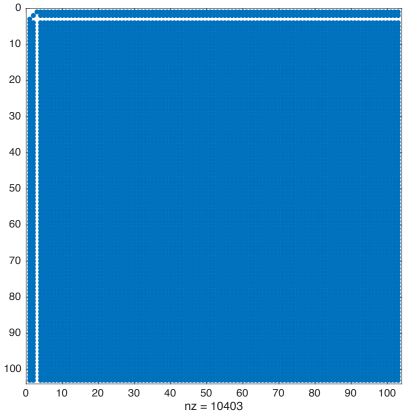
Bad news! Cournot Model (inverse demand function)

$$\max_{x_i} \; p\Big(\sum_j x_j\Big)^T x_i - c_i(x_i)$$

$$\text{s.t. } B_i x_i = b_i, x_i \geq 0$$

- Cournot model: $|\mathcal{A}| = 5$
- Size $n = |\mathcal{A}| * N_a$

Size (n)	Time (secs)
1,000	35.4
2,500	294.8
5,000	1024.6



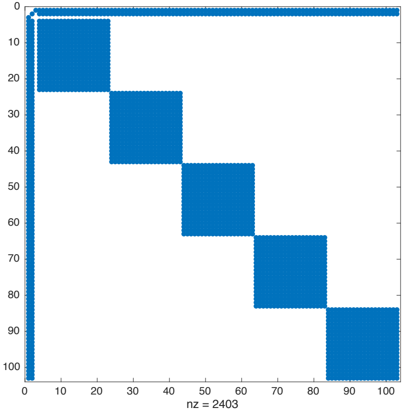
Jacobian nonzero pattern

$n = 100, \; N_a = 20$

Computation: implicit functions

- Use implicit fn: $z(x) = \sum_j x_j$
- Generalization to $F(z, x) = 0$ (via adjoints)
- empinfo: implicit z F

Size (n)	Time (secs)
1,000	2.0
2,500	8.7
5,000	38.8
10,000	> 1080



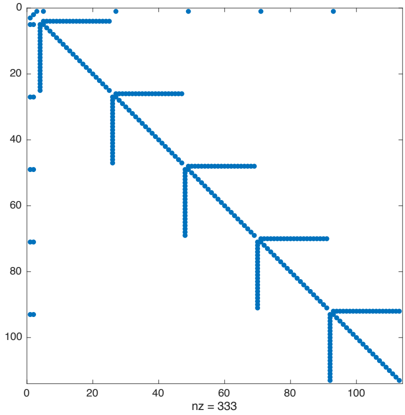
Jacobian nonzero pattern

$n = 100,\ N_a = 20$

Computation: implicit functions and local variables

- Use implicit fn: $z(x) = \sum_j x_j$ (and local aggregation)
- Generalization to $F(z, x) = 0$ (via adjoints)
- empinfo: implicit z F

Size (n)	Time (secs)
1,000	0.5
2,500	0.8
5,000	1.6
10,000	3.9
25,000	17.7
50,000	52.3



Jacobian nonzero pattern
$n = 100,\ N_a = 20$

Other specializations and extensions

$$\min_{x_i} \theta_i(x_i, x_{-i}, z(x_i, x_{-i}), \pi) \text{ s.t. } g_i(x_i, x_{-i}, z, \pi) \leq 0, \forall i, f(x, z, \pi) = 0$$

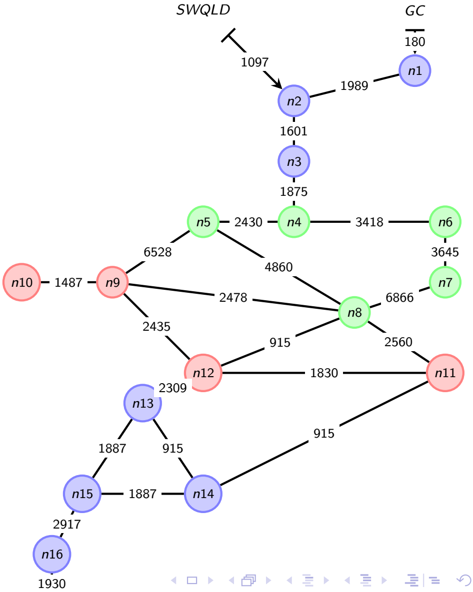
$$\pi \text{ solves } \mathrm{VI}(h(x, \cdot), C)$$

- NE: Nash equilibrium (no VI coupling constraints, $g_i(x_i)$ only)
- GNE: Generalized Nash Equilibrium (feasible sets of each players problem depends on other players variables)
- Implicit variables: $z(x_i, x_{-i})$ shared
- Shared constraints: f is known to all (many) players
- Force all shared constraints to have same dual variable (VI solution)
- Can use EMP to write all these problems, and convert to MCP form
- Use models to evaluate effects of regulations and their implementation in a competitive environment

A Simple Network Model

Load segments s represent electrical load at various instances

- d_n^s Demand at node n in load segment s (MWe)
- X_i^s Generation by unit i (MWe)
- F_L^s Net electricity transmission on link L (MWe)
- Y_n^s Net supply at node n (MWe)
- π_n^s Wholesale price ($ per MWhe)



Nodes n, load segments s, generators i, Ψ is node-generator map

$$\begin{aligned}
\max_{X,F,d,Y} \quad & \sum_s \left(W(d^s(\lambda^s)) - \sum_i c_i(X_i^s) \right) \\
\text{s.t.} \quad & \Psi(X^s) - d^s(\lambda^s) = Y^s \\
& 0 \le X_i^s \le \overline{X}_i, \quad \overline{G}_i \ge \sum_s X_i^s \\
& Y \in \mathcal{X}
\end{aligned}$$

where the network is described using:

$$\mathcal{X} = \left\{ Y : \exists F, F^s = \mathcal{H} Y^s, -\overline{F}^s \le F^s \le \overline{F}^s, \sum_n Y_n^s \ge 0, \forall s \right\}$$

- Key issue: decompose. Introduce multiplier π^s on supply demand constraint (and use $\lambda^s := \pi^s$)
- How different approximations of $\mathcal{X}$ affect the overall solution

Case $\mathcal{H}$: Loop flow model

$$
\begin{aligned}
\max_{d} \quad & \sum_{s} \left(W(d^s(\lambda^s)) - \pi^s d^s(\lambda^s)\right) \\
+\max_{X} \quad & \sum_{s} \left(\pi^s \Psi(X^s) - \sum_{i} c_i(X_i^s)\right) \\
\text{s.t.} \quad & 0 \le X_i^s \le \overline{X}_i, \quad \overline{G}_i \ge \sum_{s} X_i^s \\
+\max_{Y} \quad & \sum_{s} -\pi^s Y^s \\
\text{s.t.} \quad & \sum_{i} Y_i^s \ge 0, -\overline{F}^s \le \mathcal{H} Y^s \le \overline{F}^s
\end{aligned}
$$

$$
\pi^s \perp \Psi(X^s) - d^s(\lambda^s) - Y^s = 0
$$

Let $\mathcal{A}$ be the node-arc incidence matrix, $\mathcal{H}$ be the shift matrix, $\mathcal{L}$ be the loop constraint matrix. Standard results show:

$$\mathcal{X} = \{Y : \exists F, F = \mathcal{H}Y, F \in \mathcal{F}\}$$

$$\mathcal{X} = \left\{Y : \exists (F, \theta), Y = \mathcal{A}F, B\mathcal{A}^T\theta = F, \theta \in \Theta, F \in \mathcal{F}\right\}$$

$$\mathcal{X} = \{Y : \exists F, Y = \mathcal{A}F, \mathcal{L}F = 0, F \in \mathcal{F}\}$$

Loopflow model (using $\mathcal{A}, \mathcal{L}$)

$$
\begin{aligned}
\max_{d} \quad & \sum_{s} \left(W(d^s(\lambda^s)) - \pi^s d^s(\lambda^s) \right) \\
+ \max_{X} \quad & \sum_{s} \left(\pi^s \Psi(X^s) - \sum_{i} c_i(X_i^s) \right) \\
\text{s.t.} \quad & 0 \le X_i^s \le \overline{X}_i, \quad \overline{G}_i \ge \sum_{s} X_i^s \\
+ \max_{F,Y} \quad & \sum_{s} -\pi^s Y^s \\
\text{s.t.} \quad & Y^s = \mathcal{A} F^s, \mathcal{L} F^s = 0, -\overline{F}^s \le F^s \le \overline{F}^s
\end{aligned}
$$

$$\pi^s \perp \Psi(X^s) - d^s(\lambda^s) - Y^s = 0$$

Network model

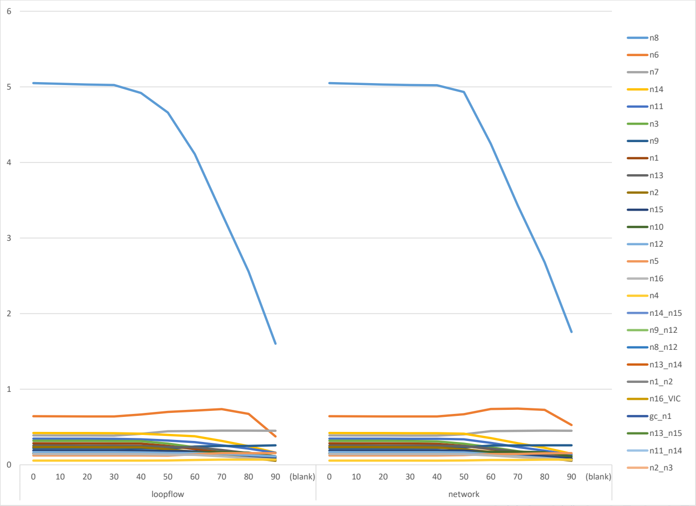
Drop loop constraints:

$$
\begin{aligned}
&\max_{d} && \sum_s \left(W(d^s(\lambda^s)) - \pi^s d^s(\lambda^s)\right) \\
+ &\max_{X} && \sum_s \left(\pi^s \Psi(X^s) - \sum_i c_i(X_i^s)\right) \\
&\text{s.t.} && 0 \le X_i^s \le \overline{X}_i, \quad \overline{G}_i \ge \sum_s X_i^s \\
+ &\max_{F,Y} && \sum_s -\pi^s Y^s \\
&\text{s.t.} && Y^s = \mathcal{A}F^s, -\overline{F}^s \le F^s \le \overline{F}^s
\end{aligned}
$$

$$\pi^s \perp \Psi(X^s) - d^s(\lambda^s) - Y^s = 0$$

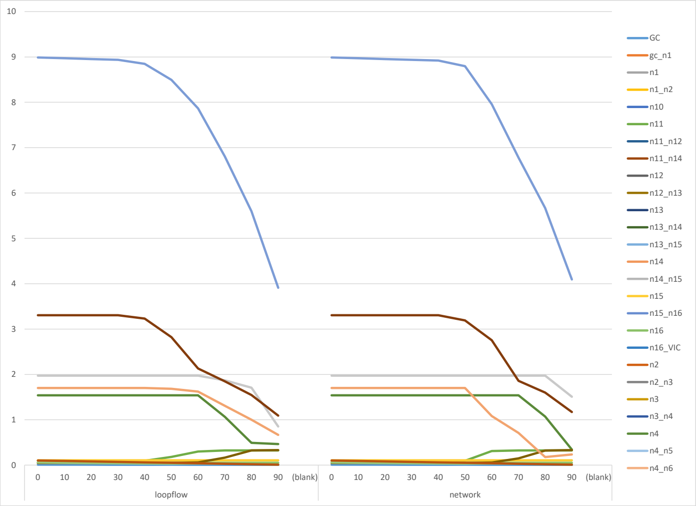
Comparing Network and Loopflow: Demand

Here we look at simulations which impose a proportional reduction in transmission across the network. The *network* and *loopflow* models demonstrate similar responses:



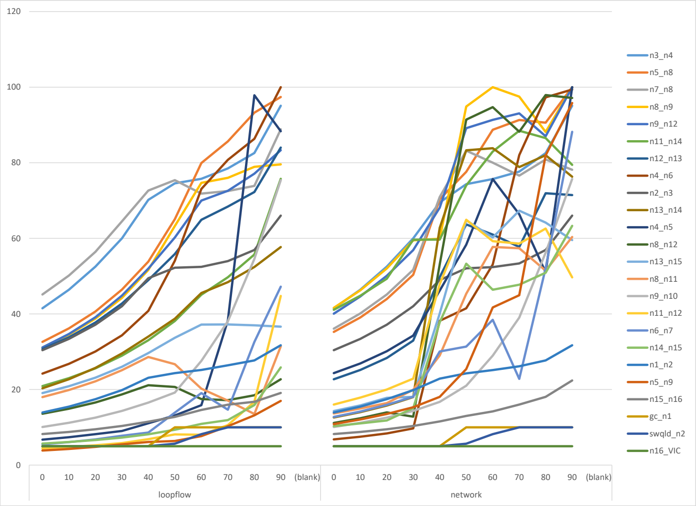
Comparing Network and Loopflow: Generation

Likewise, generation is similar in the two models:



Comparing Network and Loopflow: Transmission

Network transmission levels reveal that the two models are quite different:



The Game: update red, blue and purple components

$$\max_{d} \quad \sum_{s} \left(W(d^s(\lambda^s)) - \pi^s d^s(\lambda^s) \right)$$

$$+\max_{X} \quad \sum_{s} \left(\pi^s \Psi(X^s) - \sum_{i} c_i(X_i^s) \right)$$

$$\text{s.t.} \quad 0 \le X_i^s \le \overline{X}_i, \quad \overline{G}_i \ge \sum_{s} X_i^s$$

$$+\max_{Y} \quad \sum_{s} -\pi^s Y^s$$

$$\text{s.t.} \quad \sum_{i} Y_i^s \ge 0, -\overline{F}^s \le \mathcal{H} Y^s \le \overline{F}^s$$

$$\pi^s \perp \Psi(X^s) - d^s(\lambda^s) - Y^s = 0$$

Top down/bottom up

- $\lambda^s = \pi^s$ so use complementarity to expose (EMP: dualvar)
- Change interaction via new price mechanisms
- All network constraints encapsulated in (bottom up) NLP (or its approximation by dropping $\mathcal{L}F^s = 0$):

$$\begin{aligned} \max_{F,Y} \quad & \sum_s -\pi^s Y^s \\ \text{s.t.} \quad & Y^s = \mathcal{A}F^s, \mathcal{L}F^s = 0, -\overline{F}^s \leq F^s \leq \overline{F}^s \end{aligned}$$

- Could instead use the NLP over Y with $\mathcal{H}$
- Clear how to instrument different behavior or different policies in interactions (e.g. Cournot, etc) within EMP
- Can add additional detail into top level economic model describing consumers and producers

Pricing

Our implementation of the heterogeneous demand model incorporates three alternative pricing rules. The first is *average cost pricing*, defined by

$$P_{\mathrm{ACP}} = \frac{\sum_{jn \in \mathcal{R}_{\mathrm{ACP}}} \sum_s p_{jns} q_{jns}}{\sum_{jn \in \mathcal{R}_{\mathrm{ACP}}} \sum_s q_{jns}}$$

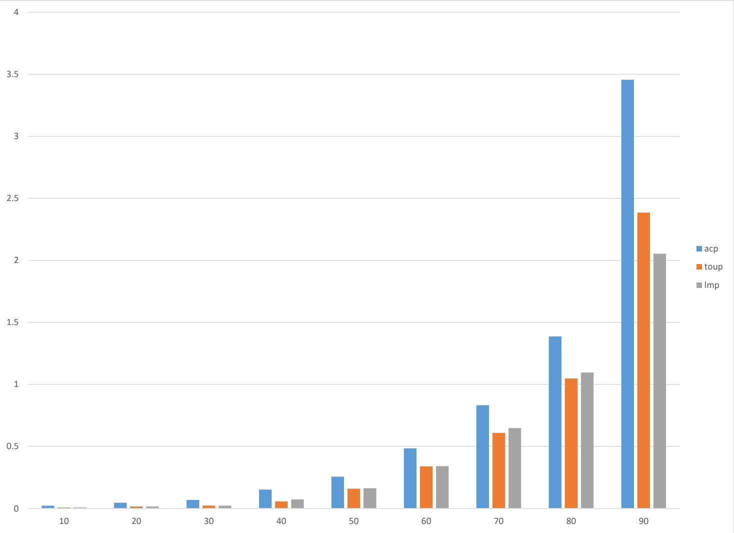
The second is *time of use pricing*, defined by:

$$P_s^{\mathrm{TOU}} = \frac{\sum_{jn \in \mathcal{R}_{\mathrm{TOU}}} p_{jns} q_{jns}}{\sum_{jn \in \mathcal{R}_{\mathrm{TOU}}} q_{jns}}$$

The third is *location marginal pricing* corresponding to the wholesale prices denoted P_{ns} above. Prices for individual demand segments are then assigned:

$$p_{jns} = \begin{cases} P_{\mathrm{ACP}} & (jn) \in \mathcal{R}_{\mathrm{ACP}} \\ P_s^{\mathrm{TOU}} & (jn) \in \mathcal{R}_{\mathrm{TOU}} \\ P_{ns} & (jn) \in \mathcal{R}_{\mathrm{LMP}} \end{cases}$$

Smart Metering Lowers the Cost of Congestion



Economic Application

- Model is a partial equilibrium, geographic exchange model.
- Goods are distinguished by region of origin.
- There is one unit of region r goods.
- These goods may be consumed in region r or they may be exported.
- Each region solves:

$$\min_{X,T_r} f_r(X,T) \text{ s.t. } H(X,T)=0,\ T_j=\bar{T}_j, j\neq r$$

  where $f_r(X,T)$ is a quadratic form and $H(X,T)$ defines X uniquely as a function of T.
- $H(X,T)$ defines an equilibrium; here it is simply a set of equations, not a complementarity problem
- Applications: Brexit, modified GATT, Russian Sanctions

Model statistics and performance comparison of the EPEC

MCP statistics according to the shared variable formulation		
Replication	Switching	Substitution
12,144 rows/cols 544,019 non-zeros 0.37% dense	6,578 rows/cols 444,243 non-zeros 1.03% dense	129,030 rows/cols 3,561,521 non-zeros 0.02% dense

PATH			Shared variable formulation (major, time)		
crash	spacer	prox	Replication	Switching	Substitution
✓		✓	7 iters 8 secs	20 iters 22 secs	20 iters 406 secs
		✓	24 iters 376 secs	22 iters 19 secs	21 iters 395 secs
	✓		8 iters 28 secs	8 iters 18 secs	8 iters 219 secs

Results

Gauss-Seidel residuals	
Iteration	deviation
1	3.14930
2	0.90970
3	0.14224
4	0.02285
5	0.00373
6	0.00061
7	0.00010
8	0.00002
9	0.00000

	Tariff revenue	
region	SysOpt	MOPEC
1	0.117	0.012
2	0.517	0.407
3	0.496	0.214
4	0.517	0.407
5	0.117	0.012

- Note that competitive solution produces much less revenue than system optimal solution
- Model has non-convex objective, but each subproblem is solved globally (lindoglobal)

Conclusions

- Equilibrium problems can be formulated naturally and modeler can specify who controls what
- It's available (in GAMS)
- Enables modelers to convey simple structures to algorithms and allows algorithms to exploit this
- New decomposition algorithms available to modeler (Gauss Seidel, Randomized Sweeps, Gauss Southwell, Grouping of subproblems)
- Can evaluate effects of regulations and their implementation in a competitive environment

MCP size of equilibrium problems containing shared variables by formulation strategy

Strategy	Size of the MCP
replication	$(n + 2mN)$
switching	$(n + mN + m)$
substitution (explicit)	$(n + m)$
substitution (implicit)	$(n + nm + m)$

$$F_i(z) = \begin{bmatrix} \nabla_{x_i} f_i(x,y) - (\nabla_{x_i} H(y,x))\mu_i \\ \nabla_{y_i} f_i(x,y) - (\nabla_{y_i} H(y,x))\mu_i \\ H(y_i, x) \end{bmatrix}, \quad z_i = \begin{bmatrix} x_i \\ y_i \\ \mu_i \end{bmatrix}.$$

Spacer steps

- Given (x, y, μ) during iterations
- Compute a unique feasible pair $(\tilde{y}, \tilde{\mu})$
- Evaluate the residual at $(x, \tilde{y}, \tilde{\mu})$
- Choose the point if it has less residual than the one of (x, y, μ)