Solving equilibrium problems using extended mathematical programming

Michael C. Ferris
(Joint work with Youngdae Kim, Andy Philpott and Jiajie Shen)

Jacques-Louis Lions Chair, and Stephen Kleene Professor of Computer Science
Computer Sciences Department and
Wisconsin Institute for Discovery, University of Wisconsin, Madison

Fields Workshop on Dynamics, Optimization and Variational Analysis in Applied Games, Toronto, May 1, 2020

What we can do?

- Equilibrium $\equiv$ complementarity ($\approx$ coupling)
- PATH solver for large scale mixed complementarity problems

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

- Nonsmooth Newton method, efficient linear algebra, available in modeling systems: GAMS, MPSGE, AMPL, AIMMS, Julia, Pyomo
- Used in models such as PIES, MERGE, VEMOD, MARKAL, TIMES, KAPSARC, ISEEM, MESSAGE, TEA, TIGER, Gemstone
- Models of Tobin, Nordhaus, Romer
- Frequently used in Computable General Equilibrium (CGE) analyses (GTAP data available), traffic, structural analysis
- Policy analyses such as Uruguay round, NAFTA, USMCA, Brexit

Equilibrium = the first-order optimality conditions (KKTs)

An equilibrium of a single optimization (a single agent) under CQs

$$
\begin{array}{lll}
\underset{x}{\text{minimize}} & f(x), & \nabla f(x) - \nabla g(x)^T \lambda - \nabla h(x)^T \mu = 0, \\
\text{subject to} & g(x) \leq 0, \;\; (\Rightarrow) & 0 \geq g(x) \perp \lambda \leq 0, \\
 & h(x) = 0, & 0 = h(x) \perp \mu,
\end{array}
$$

- Mixed complementarity problem $\text{MCP}([l, u], F) : l \leq z \leq u \;\perp\; F(z)$

Geometric first-order optimality conditions for a closed convex set K

$$
\underset{x \in K}{\text{minimize}} \quad f(x), \;\; (\Rightarrow) \quad 0 \in \nabla f(x) + N_K(x)
$$

$$
\text{i.e. } \mathrm{VI}(K, \nabla f(x))
$$

- Variational inequality $\mathrm{VI}(K, F) : \langle F(x), y - x \rangle \geq 0, \forall y \in K$

Generalizing to N agents: NEP

Nash equilibrium problem: $x = [x_i]_{i=1}^N$

$$\begin{array}{lll} \underset{x_i}{\text{minimize}} & f_i(x_i, x_{-i}), & \nabla_{x_i} f_i(x_i, x_{-i}) - \nabla g_i(x_i)\lambda_i - \nabla h_i(x_i)\mu_i = 0, \\ \text{subject to} & g_i(x_i) \leq 0, \ (\Rightarrow) & 0 \geq g_i(x_i) \perp \lambda_i \leq 0, \\ & h_i(x_i) = 0, & 0 = h_i(x_i) \perp \mu_i. \end{array}$$

- $x_{-i} := (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N)^T$.
- Equilibrium: satisfy the KKT conditions of all agents simultaneously.
- Interactions occur only in objective functions.
- Example of an interaction: $f_i(x_i, x_{-i}) = c_i(x_i) - x_i p\left(\sum_{j=1}^N x_j\right)$

NEP + non-disjoint feasible regions: GNEP

Generalized Nash equilibrium problem: $x = [x_i]_{i=1}^N$

$$\begin{aligned} \underset{x_i}{\text{minimize}} \quad & f_i(x_i, x_{-i}), \\ \text{subject to} \quad & g_i(x_i, x_{-i}) \le 0, \ (\Rightarrow) \\ & h_i(x_i, x_{-i}) = 0, \end{aligned}$$

$$\begin{aligned} \nabla_{x_i} f_i(x) - \nabla_{x_i} g_i(x)\lambda_i - \nabla_{x_i} h_i(x)\mu_i = 0, \\ 0 \ge g_i(x) \perp \lambda_i \le 0, \\ 0 = h_i(x) \perp \mu_i. \end{aligned}$$

- Interactions occur in both objective functions and constraints.
- Non-disjoint feasible region:
  $K_i(x_{-i}) = \{x_i \in \mathbb{R}^{n_i} \mid g_i(x_i, x_{-i}) \le 0, h_i(x_i, x_{-i}) = 0\}$.
  - $K_i : \mathbb{R}^{n-n_i} \rightrightarrows \mathbb{R}^{n_i}$ a set-valued mapping
  - e.g., shared resources among agents: $\sum_{i=1}^N x_i \le b$, or strategic interactions
  - Quasi-variational inequality

(G)NEP + VI agent: MOPEC

Multiple optimization problems with equilibrium constraints:
$x = [x_i]_{i=1}^N, \pi$

$$\begin{array}{lll} \underset{x_i}{\text{minimize}} & f_i(x_i, x_{-i}, \pi), & \nabla_{x_i} f_i(x,\pi) - \nabla_{x_i} g_i(x,\pi)\lambda_i - \nabla_{x_i} h_i(x,\pi)\mu_i = 0, \\ \text{subject to} & g_i(x_i, x_{-i}, \pi) \le 0, & 0 \ge g_i(x,\pi) \perp \lambda_i \le 0, \\ & h_i(x_i, x_{-i}, \pi) = 0, & 0 = h_i(x,\pi) \perp \mu_i, \end{array}$$

$$\pi \quad \in \mathrm{SOL}(K, F), \qquad \pi \in K(x), \langle F(\pi, x), y - \pi \rangle \ge 0, \ \forall y \in K(x).$$

- No hierarchy between agents, c.f., MPECs and EPECs
- An example of a VI agent: market clearing conditions

$$0 \le \text{supply} - \text{demand} \quad \perp \quad \text{price} \ge 0$$

Specifying (G)NEPs and MOPECs in modeling languages

- Existing method
  1. Compute an MCP function F using the KKT conditions.

$$\begin{array}{lll} \underset{x_i}{\text{minimize}} & f_i(x_i, x_{-i}), & \Longrightarrow \quad F_i(x, \lambda_i) = \begin{bmatrix} \nabla_{x_i} f_i - \nabla_{x_i} g_i \lambda_i \\ g_i \end{bmatrix}, \\ \text{subject to} & g_i(x_i, x_{-i}) \leq 0, & \\ & \text{for } i = 1, \ldots, N, & \quad \text{for } i = 1, \ldots, N. \end{array}$$

  2. Specify the complementarity relationship.

$$\begin{array}{ll} F \ \texttt{complements} \ (x, \lambda) & \text{in AMPL}, \\ F \ \perp \ (x, \lambda) & \text{in GAMS}. \end{array}$$

  3. Solve the resultant MCP$((x, \lambda), F)$ using the PATH solver.

  - Cons
    - Prone to errors as we require users to compute derivatives by hand
    - Not easy to modify the problem: a lot of derivative recomputations
    - Agent information is lost in the MCP function F.

The EMP framework

- Automates all the previous steps: no need to derive MCP by hand.
- Annotate equations and variables in an `empinfo` file.
- The framework automatically transforms the problem into another computationally more tractable form.

$$
\begin{aligned}
\underset{x_i}{\text{minimize}} \quad & f_i(x_i, x_{-i}, \pi), \\
\text{subject to} \quad & g_i(x_i, x_{-i}, \pi) \leq 0, \\
& h_i(x_i, x_{-i}, \pi) = 0, \\
& \text{for } i = 1, \dots, N, \\
\hline
\pi \quad & \in \mathrm{SOL}(K, F).
\end{aligned}
$$

```
equilibrium
min f('1') x('1') g('1') h('1')
...
min f('N') x('N') g('N') h('N')
vi F pi K
```

An example of using the EMP framework

- An oligopolistic market equilibrium problem formulated as a NEP:

$$q_i^* \in \quad \text{argmax}_{q_i \geq 0} \quad q_i p \left(\sum_{j=1, j \neq i}^{5} q_j^* + q_i \right) - c_i(q_i), \text{ for } i = 1, \ldots, 5.$$

```
variables obj(i); positive variables q(i);
equations defobj(i);
defobj(i)..  obj(i) =E= ...;
model m / defobj /;

file info / '%emp.info%' /;
put info 'equilibrium' /;
loop(i, put 'max', obj(i), q(i), defobj(i) /;);
putclose;
solve m using emp;
```

Special features I: supporting shared constraints

- Shared constraints: agents have shared resources.
- g is a *shared constraint*:

$$\begin{array}{ll} \underset{x_i}{\text{minimize}} & f_i(x_i, x_{-i}), \\ \text{subject to} & g(x_i, x_{-i}) \leq 0. \end{array}$$

- Examples:
  - Network capacity: $\sum_{i=1}^{N} x_i \leq b$
    Agents send packets through the same network channel.
  - Total pollutants: $\sum_{i=1}^{N} a_i x_i \leq c$
    Agents throw pollutants in the river. The maximum pollutants that can be thrown are set by a policy.

Different types of solutions for shared constraints

- A GNEP equilibrium: replicate g and assign a separate multiplier

$$\begin{aligned} \underset{x_i}{\text{minimize}} \quad & f_i(x_i, x_{-i}), \\ \text{subject to} \quad & g(x_i, x_{-i}) \leq 0, \quad (\perp \quad \mu_i \leq 0). \end{aligned}$$

- A variational equilibrium: force use of a single g and a single μ

$$\frac{\underset{x_i}{\text{minimize}} \quad f_i(x_i, x_{-i}) - \mu^T g,}{0 \geq g(x) \quad \perp \quad \mu \leq 0.}$$

- Syntactic enhancement

```
equilibrium
visol g
min f('1') x('1') g
...
min f('N') x('N') g
```

Special features II: supporting shared variables

- Shared variables: agents have shared states.
- y is a *shared variable*:

$$\begin{array}{ll} \underset{y, x_i}{\text{minimize}} & f_i(y, x_i, x_{-i}), \\ \text{subject to} & h(y, x_i, x_{-i}) = 0. \end{array}$$

  - For each x, the value of y is implicitly determined by h.
- Syntactic enhancement

```
equilibrium
implicit y h
min f('1') x('1') y
...
min f('N') x('N') y
```

MCP formulation strategies for shared variables

- Replication

$$F_i(x,y,\mu) = \begin{bmatrix} \nabla_{x_i} f_i - \nabla_{x_i} h \mu_i \\ \nabla_{y_i} f_i - \nabla_{y_i} h \mu_i \\ h \end{bmatrix} \quad \perp \quad \begin{bmatrix} x_i \\ y_i \\ \mu_i \end{bmatrix}$$

- Switching

$$\begin{array}{lcl} F_i(x,y,\mu) = \begin{bmatrix} \nabla_{x_i} f_i - \nabla_{x_i} h \mu_i \\ \nabla_y f_i - \nabla_y h \mu_i \end{bmatrix} & \perp & \begin{bmatrix} x_i \\ \mu_i \end{bmatrix} \\ \hline F_h(x,y,\mu) = \begin{bmatrix} h \end{bmatrix} & \perp & \begin{bmatrix} y \end{bmatrix} \end{array}$$

- Substitution: eliminate $\mu_i \leftarrow [\nabla_y h]^{-1} \nabla_y f_i$

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

Experimental results: improving sparsity

- Replace $p\left(\sum_{i=1}^{N} x_i\right)$ with $p(y)$ in oligopolistic market problem.
  - 1 ISO agent and 5 energy-producing agents
  - Each energy-producing agent has a fixed number of plants: $n/5$.

n	Original		Switching	
	Size	Density (%)	Size	Density (%)
2,500	2,502	99.92	2,508	0.20
5,000	5,002	99.96	5,008	0.10
10,000	10,002	99.98	10,008	0.05
25,000	-	-	25,008	0.02
50,000	-	-	50,008	0.01

n	Original		Switching	
	(Major,Minor)	Time (secs)	(Major,Minor)	Time (secs)
2,500	(2,2639)	57.78	(1,2630)	1.30
5,000	(2,5368)	420.92	(1,5353)	5.83
10,000	-	-	(1,10517)	22.01
25,000	-	-	(1,26408)	148.08
50,000	-	-	(1,52946)	651.14

Experimental results: modeling mixed behavior

- Revisiting the oligopolistic market equilibrium problem:

$$\underset{q_i \geq 0}{\text{maximize}} \quad q_i p\left(\sum_{j=1, j\neq i}^{5} q_j + q_i\right) - c_i(q_i), \text{ for } i = 1, \ldots, 5.$$

- Introduce a shared variable $y = p(q)$.
  - If an agent declares y as its decision variable, it is a price-maker.
  - Otherwise, it is a price-taker.

Profit	Compet.	Oligo1	Oligo12	Oligo123	Oligo1234	Oligo12345
Firm 1	123.83	125.51	145.59	167.02	185.958	199.93
Firm 2	195.31	216.45	219.63	243.59	264.469	279.72
Firm 3	257.81	278.98	306.17	309.99	331.189	346.59
Firm 4	302.86	322.51	347.48	373.46	376.697	391.28
Firm 5	327.59	344.82	366.54	388.97	408.308	410.36
Total	1207.41	1288.27	1385.42	1483.02	1566.62	1627.875
Soc./wf.	39063.82	39050.19	39034.58	39022.47	39016.37	39015.125

Optimal Value Functions

Problem type

Objective function	or	Constraint
$\min_{x \in X} \theta(x) + \rho(F(x))$		$\min_{x \in X} \theta(x) \text{ s.t. } \rho(F(x)) \leq \alpha$

- Special case is a Quadratic Support Function

$$\rho(y) = \sup_{u \in U} \langle u, By + b \rangle - \frac{1}{2} \langle u, Mu \rangle$$

- Dual representation (of coherent r.m.) in terms of risk sets

$$\rho(Z) = \sup_{\mu \in \mathcal{D}} \mathbb{E}_{\mu}[Z]$$

- If $\mathcal{D} = \{p\}$ then $\rho(Z) = \mathbb{E}[Z]$
- If $\mathcal{D}_{\alpha,p} = \{\lambda \in [0, p/(1-\alpha)] : \langle \mathbb{1}, \lambda \rangle = 1\}$, then $\rho(Z) = \overline{CVaR}_{\alpha}(Z)$

The transformation to MOPEC

- EMP allows any Quadratic Support Function to be defined and facilitates model transformations to tractable forms for solution
- `empinfo` file: OVF cvarup F(x) rho .9

$$\min_{x \in X} \theta(x) + \rho(F(x))$$

$$\rho(y) = \sup_{u \in U} \left\{ \langle u, y \rangle - \frac{1}{2} \langle u, Mu \rangle \right\}$$

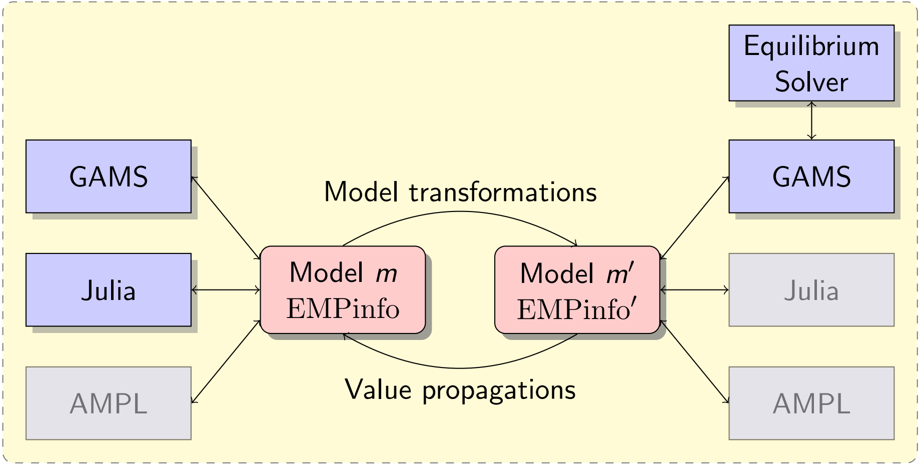
$$0 \in \partial\theta(x) + \nabla F(x)^T \partial\rho(F(x)) + N_X(x)$$

$$0 \in \partial\theta(x) + \nabla F(x)^T u + N_X(x)$$

$$0 \in -u + \partial\rho(F(x)) \iff 0 \in -F(x) + Mu + N_U(u)$$

- This is a MOPEC, and we have a copy of this construct for each agent

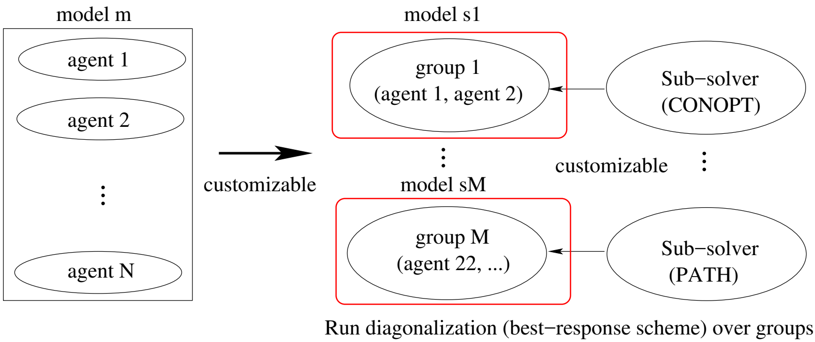
EMP framework



The model representation inside the EMP solver is independent of any model language

Selkie

- Selkie
  - Generates submodels for sub-solvers and decomposition.
  - Supports various decomposition methods.
  - Can compute a solution in an adaptable and flexible way.
  - ex) Selkie on equilibrium problems:



An example of using SELKIE for group diagonalization

- An oligopolistic market equilibrium problem:

$$\underset{q_i \geq 0}{\text{maximize}} \quad q_i p \left(\sum_{j=1, j\neq i}^{5} q_j + q_i \right) - c_i(q_i), \text{ for } i = 1, \ldots, 5.$$

Group	Iterations			
	Jacobi	GS	GSW	GS(RS)
{{1},{2},{3},{4},{5}}	155	45	28	50
{{1,2},{3,4},{5}}	57	21	22	30
{{1..3},{4,5}}	28	14	14	18
{{1..4},{5}}	22	12	12	16
{{1..5}}	1			

  - GS: Gauss-Seidel
  - GSW: Gauss-Southwell
  - GS(RS): Gauss-Seidel with random sweep

- An automatic detection of independent groups is supported.

Multistage MOPEC with risk averse agents

Deterministic equivalent under dual risk measure representation:

$$
\begin{aligned}
\min_{x_{i\cdot},\theta_{i\cdot}} \quad & f_{i0}(x_{i0}) - p_0^T g_{i0}(x_{i0}) + \theta_{i0} \\
\text{s.t.} \quad & \theta_{im} \geq \sum_{n \in m+} \pi_{in}^k \cdot \{ f_{in}(x_{in}) - p_n^T g_{in}(x_{in}) + \theta_{in} \}, \qquad && \forall m \notin L, \ \ k \in K_{im} \\
& x_{in} = H_{in} x_{im} + \omega_{in}, && \forall m \notin L, \ \ n \in m+ \\
& x_{in} \in [l_{in}, u_{in}], && \forall n
\end{aligned}
$$

with equilibrium constraint

$$
0 \leq \sum_a g_{in}(x_{in}) \perp p_n \ \ \geq 0, \quad \forall n
$$

- Here n represents the node of the scenario tree, L is the set of the nodes that are the leaves of the scenario tree, $n+$ represents the set of children nodes of the node n, K_{im} is the set of extreme points of risk set of agent i at node m.

Alternative method: Primal - dual method for solving the multistage MOPEC with risk averse agents

- Previous distributed methods (Gauss-Seidel and Jacobi) fail to solve the MOPEC
  - No implicit function $\pi = h(x)$ from the constraint $0 \leq H(x, \pi) \perp \pi \geq 0$
  - The subproblem is not solvable or unbounded without $H(x, \pi) \geq 0$
- In the risk-averse case, the corresponding reformulated mixed complementarity problem will lose monotonicity even when the risk-neutral case is monotone
- PATH fails to solve, even with informed choices of starting point
- Use Penalty (Augmented Lagrangian) of the constraint $H(x, \pi) \geq 0$ in each primal agents' problem and dual update in each major iterations.
- Performance depends on the choice of γ.

Algorithm 1 Primal-dual for multistage MOPEC with risk-averse agents

set $k = 0$, choose a starting point (x^0, p^0), parameter $\gamma > 0, 0 < \mu \leq 1$
while stopping criterion not met **do**
  for each agent a **do**

$$
\begin{aligned}
x_{i\cdot}^{k+1}, \theta_{i\cdot}^{k+1}, y_{i\cdot}^{k+1} = \arg \min_{x_{i\cdot}, \theta_{i\cdot}, y_{i\cdot}} & \; f_{i0}(x_{i0}) - (p_0^k)^T g_{i0}(x_{i0}) + (p_0^k)^T y_{i0} + \theta_{i0} \\
& + \frac{\gamma}{2} \sum_n \| g_{in}(x_{in}) + \sum_{j<i}^{N} \left(g_{jn}(x_{jn}^{k+1}) - y_j^{k+1} \right) + \sum_{j>i}^{N} \left(g_{jn}(x_{jn}^{k}) - y_j^{k} \right) - b_n \|^2 \\
\text{s.t. } & \; \theta_{im} \geq \sum_{n \in m+} \pi_{in}^k \cdot \{ f_{in}(x_{an}) - p_n^T g_{in}(x_{in}) + p_n^T y_{in} + \theta_{in} \}, \qquad m \notin L, \;\; k \in K_{im} \\
& \; x_{in} = H_{in} x_{im} + \omega_{an}, \quad m \notin L, \;\; n \in m+ \\
& \; x_{in} \in [l_{in}, u_{in}], \;\; y_{in} \geq 0
\end{aligned}
$$

  end for
  $p_n^{k+1} = p_n^k - \mu\gamma \cdot \left(\sum_{j=1}^N g_{jn}(x_{jn}^{k+1}) - b_n - \sum_{j=1}^N y_{jn}^{k+1} \right)$
  $k = k + 1$
end while

Comparison between PATH and Primal-Dual method

- A 4-agent example with 5 stochastic stages, where n represents the dimension of the corresponding MCP
- Risk measure: $\rho(X) = (1-\lambda)\cdot\mathbb{E}(X) + \lambda\cdot CVaR_{0.95}(X)$

Risk averse

Size: $n \times n$	λ	PATH	Primal-Dual		
		Final merit	# Iter	Final merit	Time(secs)
2680 × 2680	0.1	1.15e+03	6887	9.06e-07	7107
2680 × 2680	0.2	1.83e+02	7156	9.35e-05	7200
2680 × 2680	0.3	1.66e+02	7073	3.32e-03	7200
2680 × 2680	0.4	2.71e+02	7083	2.62e-02	7200

Risk averse: use previous solution as initial point

Size: $n \times n$	λ	Primal-Dual		
		# Iter	Final merit	Time(secs)
2680 × 2680	0.1	5179	9.51e-07	4401
2680 × 2680	0.2	5706	9.95e-07	7000
2680 × 2680	0.3	7049	2.52e-06	7200
2680 × 2680	0.4	6967	6.43e-05	7200

Conclusions

- Competition naturally modeled via complementarity
- Solvers exist for medium to large scale problems
- Frameworks (EMP) exist to streamline model transformations
- empinfo: dualvar, bilevel, equilibrium, vi, OVF
- Very large scale models (many agents with many instruments acting strategically) with risk are hard
- Decomposition/diagonalization methods (SELKIE) are effective when sensitivity information is exploited
- New algorithms enable solution of more detailed, authentic problems that address underlying policy questions
- Evaluation via simulation computations and out-of-sample testing