# Discussion of “Constrained Optimization Approaches to Estimation of Structural Models” by Su and Judd

Andrew Ching
Rotman School of Management
University of Toronto

# Summary

- The key insight of the paper:
  - Likelihood is a function of the model, e.g., E[V(s')|s].
  - Solving E[V(s')|s] is time-consuming. But they can be treated as parameters. One parameter per state point.
  - The problem can be reformulated as a constrained maximization problems.
  - No. of parameters could increase dramatically. But it also eliminates non-linearity. With automatic differentiation software, one can get exact $1^{st}$ and $2^{nd}$ derivatives.
  - No. of $2^{nd}$ derivatives depend on the structure of the model.

# $2^{nd}$ derivatives of the Hessian matrix

- MPEC reduces computational burden more if the structure of model leads to many zeros in the $2^{nd}$ derivatives of the Hessian matrix (sparsity structure).
- Why? My guess is that the transition markov matrix for the state variables is sparse.
- Let $s'$ be the next period state var, and $s$ be the current period state var. $f(s'|s)$ restricts $s'$ to be in a small neighborhood around $s$.

# Potential limitation 1: $2^{nd}$ derivatives of the Hessian matrix

- An example where this may not be the case. Diermeier, Keane and Merlo (2005) study politicians career decisions. State variables includes political atmosphere (good for democrats, good for republican or neutral) at the state and district level. These markov matrices are not sparse at all. Another example could be price uncertainty.
- Additional computational costs of $2^{nd}$ derivatives would be involved. This could be a concern, in particular, if the size of the state space is large.

# Potential limitation 2: Models with complicated constraints

- Non-stationary dynamic problems, e.g., Keane and Wolpin (1997), Blau and Gillesekie (2008), Khwaja (2010), Yang and Ching (2010), Chen et al. (2009).
- The dynamic problems of these models are non-stationary because the value function depends on time. Typical way to solve these models is to use backward induction. Moreover, the structure of the problems could change across time.

# Potential limitation 2 (cont'd)

- Constraints: Let $T$ be the terminal period.

  $V_T(s_T)$ , for all $s_T$

  $V_{T-1}(s_{T-1}) = f(E\ [V_T(s_T)|s_{T-1}]\ )$, for all $s_{T-1}$

  $V_{T-2}(s_{T-2}) = f(E\ [V_{T-1}(s_{T-1})|s_{T-2}]\ )$, for all $s_{T-2}$

  .

  .

  .

  $V_1(s_1) = f(E\ [V_2(s_2)|s_1]\ )$, for all $s_1$

- Evaluating all the constraints once is the same as solving the model once.
- MPEC might still have an advantage because it reduces the non-linearity of the optimization. But it seems that the memory requirement could be very large.

# Potential limitation 3: Models w/o analytical derivatives

- Suppose we want to allow for random coefficients in a dynamic programming models.
- In principle, there is no analytical derivatives.
- But it maybe okay if one is only interested maximizing a simulated likelihood.

# Games with incomplete information

- It is good to know that MPEC is able to recover parameters even though the data is generated from an unstable equilibrium in the "best-reply" sense.
- How likely would the data be generated from such an unstable equilibrium?
- If one can show that the empirical games used in applied works are actually stable under other popular notions of stable equilibrium concepts, that will strengthen your results.
- You may also want to use a smaller sample size in your Monte Carlo exercises to show that 2-step method can really lead to biased results, but not for MPEC, even for data generated from a stable equilibrium.
- It is unclear to me why the likelihood generated from one equilibrium could be discontinuous.

# Some statements are unclear

- "The advantage of **not enforcing** the constraints to be satisfied until the last few iterations is that the algorithm has the flexibility to search over the **infeasible region** and, as a result, needs few **major** iterations to converge …"
  - No need to enforce the constraints at all, or require them to be loosely satisfied?
  - Why searching over infeasible region helps?
  - What does "major" mean?

# Final comment

- "for each theta considered in MPEC, the Bellman equation is evaluated once, not solved"
  - For a stationary dynamic programming (DP) problem, this sounds similar to Imai, Jain and Ching (2009).
  - For a non-stationary DP problem, this is equivalent to solving the entire model once.
- Can MPEC handles continuous state space?

# To summarize

- The authors show that MPEC could be very promising for estimating a structural model that satisfies certain structures, e.g., stationary problems, sparsity in transition matrix.
- But for more complicated models, I feel that it is still an open question whether MPEC could outperform other existing approaches.
- This method is not for Bayesian econometricians.