Constrained Optimization Approach to Estimation of Structural Models (2012, Ecta)

Che-Lin Su, Kenneth L. Judd

Presenter: Jun Zhang

February 12, 2014

Motivation

- Computational burden of estimating structural models is a problem.
- Existing methods:
  - Nested fixed point (NFXP): Rust (1987).
  - Two-step estimator: Hotz and Miller (1993), and others.
  - Nested pseudo-likelihood (NPL): Aguirregabiria and Mira (2002, 2007).
  - Modified NPL: Kasahara and Shimotsu (2012).
- New computational method: mathematical program with equilibrium constraints (MPEC).
  - Computational ease: constraints need not to be satisfied until the final iteration .

Outline

1. Define MPEC.
2. Apply MPEC to single-agent dynamic discrete-choice model in Rust (1987).
3. Compare MPEC and NFXP in Monte Carlo experiments.
4. (If time allows) Apply MPEC to dynamic discrete-choice games of incomplete information: Egesdal, Lai and Su (2013).

MPEC

- Equilibrium constraint: $h(\theta, \sigma) = 0$.
  - θ, structural parameter: costs, transition probs, random shocks, etc.
  - σ, endogenous variable: expected value function, equilibrium response, etc.

- Data: $X = \{x_i, d_i\}_{i=1}^{M}$.
  - x_i: observed state variable.
  - d_i: observed equilibrium outcome.

- Examples:
  - single-agent dynamic discrete-choice: Rust (1987).
  - dynamic discrete-choice game of incomplete information: Aguirregairia and Mira (2007).

MPEC cont'd

- Let $\Sigma(\theta) := \{\sigma : h(\theta, \sigma) = 0\}$.
- If $\Sigma(\theta)$ is single-valued, then $\hat{\sigma}(\theta) = \Sigma(\theta)$ and

$$\hat{\theta} = \underset{\theta}{argmax} \frac{1}{M} \mathcal{L}(\theta, \hat{\sigma}(\theta); X). \quad (1)$$

- If $\Sigma(\theta)$ is multi-valued, then

$$\hat{\theta} = \underset{\theta}{argmax} \frac{1}{M} \left\{ \max_{\hat{\sigma}(\theta) \in \Sigma(\theta)} \mathcal{L}(\theta, \hat{\sigma}(\theta); X) \right\}. \quad (2)$$

- NFXP:
  - Outer loop: search over θ.
  - Inner loop: find $\Sigma(\theta)$ and $\hat{\sigma}(\theta)$.
- MPEC:

$$\max_{(\theta, \sigma)} \quad \frac{1}{M} \mathcal{L}(\theta, \sigma; X) \quad (3)$$

$$s.t. \quad h(\theta, \sigma) = 0.$$

MPEC cont'd

Proposition 1

Let $\hat{\theta}$ be the estimator defined in (2) and $(\bar{\theta}, \bar{\sigma})$ be a solution of (4). Define $\hat{\sigma}^*(\theta) = \underset{\hat{\sigma}(\theta)}{\text{argmax}} L(\theta, \hat{\sigma}(\theta))$. Then $L(\hat{\theta}, \hat{\sigma}^*(\theta)) = L(\bar{\theta}, \bar{\sigma})$. If the model is identified, then $\hat{\theta} = \bar{\theta}$.

Proof: Since $\bar{\sigma} \in \Sigma(\bar{\theta})$, hence, $L(\hat{\theta}, \hat{\sigma}^*(\theta)) \geq L(\bar{\theta}, \bar{\sigma})$. Conversely, since $\hat{\sigma}^*(\theta) \in \Sigma(\hat{\theta})$, so $L(\bar{\theta}, \bar{\sigma}) \geq L(\hat{\theta}, \hat{\sigma}^*(\theta))$.

Single-Agent Dynamic Discrete-Choice: Rust (1987)

- NFXP:

$$EV(x,d) = \sum_{x'} log\left\{\sum_{d'\in\{0,1\}} exp[v(x',d';\theta_1,RC) + \beta EV(x',d')]\right\} p_3(dx'|x,d,\theta_3). \tag{4}$$

$$P(d|x;\theta) = \frac{exp[v(x,d;\theta_1,RC) + \beta EV(x,d)]}{\sum_{d'\in\{0,1\}} exp[v(x,d';\theta_1,RC) + \beta EV(x,d')]}. \tag{5}$$

$$\max_{\theta} \frac{1}{M}\left(\sum_{i=1}^{M}\sum_{t=2}^{T} log[P(d_t^i|x_t^i;\theta)] + \sum_{i=1}^{M}\sum_{t=2}^{T} log[p_3(x_t^i|x_{t-1}^i,d_{t-1}^i;\theta_3)]\right). \tag{6}$$

- MPEC:

$$\max_{\theta,EV} \frac{1}{M}\left\{\sum_{i=1}^{M}\sum_{t=2}^{T} log\left(\frac{exp[v(x_t^i,d_t^i;\theta_1,RC) + \beta EV(x_t^i,d_t^i)]}{\sum_{d'\in\{0,1\}} exp[v(x_t^i,d';\theta_1,RC) + \beta EV(x_t^i,d')]}\right) + \right.$$

$$\left. \sum_{i=1}^{M}\sum_{t=2}^{T} log[p_3(x_t^i|x_{t-1}^i,d_{t-1}^i;\theta_3)]\right\} \tag{7}$$

s.t. (4).

Monte Carlo Experiments: Rust (1987)

- True values: $RC^0 = 11.7257, \theta_{11}^0 = 2.4569, \theta_3^0 = (.0937, .4475, .4459, .0127, .0002)$.
- See Su and Judd (2012).

Dynamic Discrete-Choice Games of Incomplete Information: Egesdal, Lai and Su (2013)

- Periods $t = 1, 2, ..., \infty$, players $i \in \mathcal{I} = \{1, 2, ..., N\}$, and market sizes $s^t \in \{s_1, s_2, ..., s_L\}$.

- At the beginning of t, i observes state variables $x^t = (s^t, a^{t-1}) \in \mathcal{X}$ and private shocks ϵ_i^t, and chooses to be active or not, $a_i^t \in \mathcal{A} = \{0, 1\}$. $a^t = (a_1^t, a_2^t, ..., a_N^t)$, $\epsilon_i^t = \{\epsilon_i^t(a_i^t)\}_{a_i^t \in \mathcal{A}}$.

- Per-period payoff: $\tilde{\Pi}_i(a_i^t, a_{-i}^t, x^t, \epsilon_i^t; \theta) = \Pi_i(a_i^t, a_{-i}^t, x^t; \theta) + \epsilon_i^t(a_i^t)$.

- $s^{t+1} \sim f_S(s^{t+1}|s^t)$, $\epsilon_i^t(a_i^t)$ follows type-I extreme value dist. and is i.i.d across actions, players and periods. Assume
$$p[x^{t+1} = (s', a'), \epsilon_i^{t+1} | x^t = (s, \tilde{a}), \epsilon_i^t, a^t] = f_S(s'|s)\mathbf{1}\{a' = a^t\}g(\epsilon_i^{t+1}).$$

- Assume Markov perfect equilibrium. Let $P_i(a_i|x)$ be conditional choice prob, then expected payoff from Π_i is
$$\pi_i(a_i|x, \theta) = \sum_{a_{-i} \in \mathcal{A}^{N-1}} \left\{ \left[\prod_{a_j \in a_{-i}} P_j(a_j|x) \right] \Pi_i(a_i, a_{-i}, x; \theta) \right\}.$$

- $V_i(x)$, expected value function at state x. $P = \{P_i(a_i|x)\}_{a_i \in \mathcal{A}, i \in \mathcal{I}, x \in \mathcal{X}}$, $V = \{V_i(x)\}_{i \in \mathcal{I}, x \in \mathcal{X}}$.
- A MPE is a tuple (V, P) satisfying following two conditions:
  - Bellman Optimality:

$$V_i(x) = \sum_{a_i \in \mathcal{A}} P_i(a_i|x)[\pi_i(a_i|x,\theta) + e^P(a_i,x)] + \beta \sum_{x' \in \mathcal{X}} V_i(x') f_{\mathcal{X}}^P(x'|x)$$
$$= \Psi_i^V(x; V, P, \theta), \text{ where } e_i^P(a_i, x) = \text{Euler's constant} - \sigma log[P_i(a_i|x)].$$

  - Bayes-Nash Equilibrium:

$$v_i(a_i|x) = \pi_i(a_i|x,\theta) + \beta \sum_{x' \in \mathcal{X}} V_i(x') f_i^P(x'|x, a_i),$$
$$P_i(a_i = j|x) = Pr\Big[\epsilon_i | v_i(a_i = j|x) + \epsilon_i(a_i = j) > \max_{k \in \mathcal{A} \setminus j}\{v_i(a_i = k|x) + \epsilon_i(a_i = k)\}\Big].$$

    Following type-I EVD assumption,

$$P_i(a_i = j|x) = \frac{exp[v_i(a_i = j|x)]}{\sum_{k \in \mathcal{A}} exp[v_i(a_i = k|x)]}$$
$$= \Psi_i^P(a_i = j|x; V, P, \theta).$$

- Define $\Psi^V(x; V, P, \theta) = \{\Psi_i^V(x; V, P, \theta)\}$ and $\Psi^P(x; V, P, \theta) = \{\Psi_i^P(a_i = j|x; V, P, \theta)\}$. A MPE is characterized by

$$V = \Psi^V(V, P, \theta), \tag{8}$$

$$P = \Psi^P(V, P, \theta). \tag{9}$$

Estimation

- Data: M independent markets over T periods. $Z = \{\bar{a}^{mt}, \bar{x}^{mt}\}_{m \in \mathcal{M}, t \in \mathcal{T}}$ are generated from only one Markov perfect equilibrium (V^0, P^0) at the true parameter value θ^0.
- MPEC:

$$\max_{(\theta,P,V)} \frac{1}{M} \mathcal{L}(Z; V, P, \theta) = \frac{1}{M} \left(\sum_{i=1}^{N} \sum_{m=1}^{M} \sum_{t=1}^{T} log \Psi_i^P (\bar{a}_i^{mt} | \bar{x}^{mt}; V, P, \theta) \right) \tag{10}$$

$$\begin{aligned} \text{s.t.} \quad & V = \Psi^V(V, P, \theta), \\ & P = \Psi^P(V, P, \theta). \end{aligned}$$

- Two-Step Pseudo-Maximum Likelihood:
  - Step one: Nonparametrically estimate P^0 from data Z, denoted $\hat{P}$.
  - Step two:

$$\max_{(\theta,V)} \frac{1}{M} \mathcal{L}(Z; V, \hat{P}, \theta) \tag{11}$$

$$s.t. \quad V = \Psi^V(V, \hat{P}, \theta). \tag{12}$$

  - By (12) we get $V = \Gamma(\theta, \hat{P})$. So equivalently, $\theta^{2S-PML} = \underset{\theta}{argmax} \frac{1}{M} \mathcal{L}(Z; \Gamma(\theta, \hat{P}), \hat{P}, \theta)$.
  - $(\theta^{2S-PML}, \hat{P})$ may not be BNE, that is, not satisfying $P = \Psi^P(V, P, \theta)$.

- Nested Pseudo-Likelihood:
  - An NPL fixed point:

$$\tilde{\theta} = argmax_{\theta} \frac{1}{M} \mathcal{L}(Z; \Gamma(\theta, \tilde{P}), \tilde{P}, \theta),$$

$$\tilde{P} = \Psi^{P}(\Gamma(\tilde{\theta}, \tilde{P}), \tilde{P}, \tilde{\theta}).$$

  - NPL algorithm: Start with $\tilde{P}_0$,

    Step one: Given $\tilde{P}_{K-1}$, solve $\tilde{\theta}_K = argmax_{\theta} \frac{1}{M} \mathcal{L}(Z; \Gamma(\theta, \tilde{P}_{K-1}), \tilde{P}_{K-1}, \theta)$,

    Step two: Given $\tilde{\theta}_K$, $\tilde{P}_K = \Psi^{P}(\Gamma(\tilde{\theta}_K, \tilde{P}_{K-1}), \tilde{P}_{K-1}, \tilde{\theta}_K)$.

    Maximum number of iterations, $\bar{K}$.

- Modified NPL (NPL-Λ):
  Alter step two of NPL to

$$\tilde{P}_K = \left(\Psi^{P}(\Gamma(\tilde{\theta}_K, \tilde{P}_{K-1}), \tilde{P}_{K-1}, \tilde{\theta}_K)\right)^{\lambda} \left(\tilde{P}_{K-1}\right)^{1-\lambda}, \lambda \in [0, 1].$$

Monte Carlo Experiments

See Egesdal, Lai and Su (2013).