Introduction to Structural Econometrics

Lesson 4: Simulation-based Econometrics

Guillaume Wilemme

Aix-Marseille School of Economics - PhD course

Nov 2017

Table of contents

Plan

1. General questions
2. Estimation methods, a panorama
3. Maximum simulated likelihood (MSL) and method of simulated moments (MSM)
   - MSL
   - Dual interpretation
   - MSM
   - Practical advise
4. Indirect inference and limits

Schedule

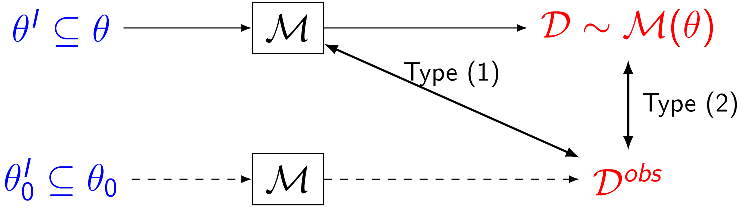
- October 3rd, room 23, 10-12am
  What is Structural Econometrics?
- October 10th, room 23, 10-12am
  Dynamic Programming and Recursive Models
- October 24th, room 23, 10-12am
  Monte-Carlo Methods
- November 7th, room 23, 10-12am
  Simulation-based Estimation
- November 14th, room 15, 10-12am
  State-Space Representation and Latent Variables
- November 21st, room 16, 10-12am
  Numerical tools

Plan

Two categories of estimation (II)

- Goal of the econometrician: recover θ_0^I (or $\mathcal{M}(\theta_0)$) using available information from $\mathcal{D}^{obs}$. $\mathcal{M}(\theta_0)$ is the distribution from which $\mathcal{D}^{obs}$ is drawn.
- 2 ways of constructing an estimation strategy $\bar{\theta}^I(\mathcal{D})$.
    1. Direct methods. The structure of $\mathcal{M}$ enables you to build an explicit criterion function to optimize. Examples: least squares, maximum likelihood, (generalized) method of moments.
    2. Indirect methods. The structure is not tractable enough. The estimation relies on the comparison between simulated samples and the observed dataset. Examples: simulated methods.

Two categories of estimation (II)



A categorization of estimation approaches

Direct: explicit objective function from the structure "$Obj(\theta^I, \mathcal{D}^{obs})$"			Indirect
Closed-form solution (first-order condition)	Numerical problem		
	Deterministic method	Stochastic method (Monte-Carlo)	
OLS formula, IV formula, FOC in ML	OLS objective, Max. likelihood (ML), Pseudo-ML, GMM	MSL, MSM, MSPL	Indirect inference, MSL, MSM 7.2%

Plan

Example: biprobit

$$
\begin{aligned}
y_{i1} &= \mathbf{1}(\beta x_{i1} + \epsilon_{i1} \geq 0),\\
y_{i2} &= \mathbf{1}(\beta x_{i2} + \epsilon_{i2} \geq 0),\\
(\epsilon_{i1}, \epsilon_{i2}) &\sim \mathcal{N}(0, \Sigma).
\end{aligned}
$$

Notations: $\mathcal{D} = \{y_i, x_i\}$ and $\theta = \{x_i, \beta, \Sigma\}$.

Likelihood function

- Define the likelihood of a dataset $\mathcal{D}$ as $\mathcal{L}(\mathcal{D}, \theta)$. If individual outcomes independent, then $\mathcal{L}(\mathcal{D}, \theta) = \prod_i \mathcal{L}_i(\mathcal{D}_i, \theta)$.
- Biprobit: an individual likelihood writes as a joint probability, for instance

$$\mathcal{P}(\beta x_{i1} + \epsilon_{i1} \geq 0; \beta x_{i2} + \epsilon_{i2} < 0))$$

- It is an integral that I have introduced when describing the GHK:

$$\int \mathbf{1}(\beta x_{i1} + \epsilon_{i1} \geq 0)\mathbf{1}(\beta x_{i2} + \epsilon_{i2} < 0) d\mathcal{F}_\Sigma(\epsilon_{i1}, \epsilon_{i2})$$

- To estimate β and Σ, the econometrician (or rather the econometrician's computer) has to evaluate this integral as a function of (β, Σ) for each individual.

When the likelihood involves an integral

- Quadrature methods
- MC method for likelihood

$$\mathcal{L}_i(\mathcal{D}_i, \theta) = \int \tilde{\mathcal{L}}_i(\mathcal{D}_i, \epsilon, \theta) d\mathcal{F}(\epsilon)$$

Using a sequence $(\epsilon_i^1, ..., \epsilon_i^S)$ that is distributed from $\mathcal{F}$, the integral can be apprimated by a Monte-Carlo simulator:

$$\frac{1}{S} \sum_{s=1}^{S} \tilde{\mathcal{L}}_i(\mathcal{D}_i, \epsilon_i^s, \theta)$$

Maximizing the likelihood computed this way is what is called the maximum simulated likelihood (or simulated maximum likelihood). The GHK algorithm enables in practice to estimate multi-probit models through MSL.

Two was of interpreting MSL

- MC as a tool (almost a black box) to compute an integral.
- MC as indirect inference: consider that, at each simulation s, the model is simulated.
  For each simulated dataset $\mathcal{D}^s$, we do not evaluate the likelihood $\mathcal{L}$ but the likelihood with the information of ϵ, $\tilde{\mathcal{L}}$. By simulating the model multiple times, we can recover the likelihood function.

MSM

- Extend the MSL method to the methods of moments.
- Pick some moments of the data, for instance the average products $\frac{1}{N}\sum_i y_{i1}x_{i1}$, $\frac{1}{N}\sum_i y_{i2}x_{i2}$ and $\frac{1}{N}\sum_i y_{i1}y_{i2}x_{i1}x_{i2}$.
- Methods of moments consist in minimizing the distance between theoretical moments and the empirical counterpart of these moments as defined above.

MSM (II)

- Biprobit: the theoretical moments are integrals like

$$\mathbb{E}(Y_1 Y_2 X_1 X_2) = X_1 X_2 \int \mathbf{1}(\beta X_1 + \epsilon_1 \geq 0)\mathbf{1}(\beta X_2 + \epsilon_2 \geq 0) d\mathcal{F}_\Sigma(\epsilon_{i1}, \epsilon_{i2}$$

  This theoretical moment is a function of $\{x_i\}$ and (β, Σ).
- The objective function to minimize is a (weighted) sum of distances between theoretical moments and empirical moments.
- Here again, there are three approaches: quadrature computation of the integral, Monte-Carlo method and model simulation.

Three pieces of advise advise

In estimating a model through simulation (or Monte-Carlo):

- Beware of low probabilities of events which result in a high variance of the estimated integral or objective function. Example of solution : importance sampling.
- Try to keep the same sequence of simulated random varaibles, $\epsilon^1, .., \epsilon^S$, along the estimation, instead of drawing new sequence at each evaluation. This method gives better asymptotic properties of the estimates.
- Try to obtain a differentiable objective function in the estimate. The GHK method offers such a differentiability condition.

Plan

Indirect inference

- Indirect inference consists in looking at samples generated by the structural model $\mathcal{M}(\theta)$.
- Generate samples from simulations $\mathcal{D}^1(\theta), .., \mathcal{D}^S(\theta)$. The econometrician will build an estimate of θ^I by comparing the sequence of simulated data at a given θ to the "true" observed data.

Indirect inference (II)

- In practice, this comparison is based on "statistics" of the dataset, say $\beta(\mathcal{D})$.
- β is called the auxiliary parameter and the transformation from $\mathcal{D}$ to β is called the auxiliary model.
- If we define $\tilde{\beta}(\theta) = \frac{1}{S}\sum_s \beta(\mathcal{D}^s(\theta))$, then an estimate of θ^I can be found by minimizing a distance between $\beta(\mathcal{D}^{obs})$ and $\tilde{\beta}(\theta)$.
- For instance the MSM is simply choosing β as the moments of interest. Analoguously, MSL consists in choosing $\beta(\mathcal{D})$ as the ML estimate of θ (stupid in that case).

Where is the limit?

- Unless your model cannot be simulated, simulation-based inference opens the doors of estimating any model.
- There is, however, a constraint: identification. If the parameters of interest θ^I are not identified, simulation-based inference cannot help you.
- In practice, identification is a threat of your estimation and something to keep in mind when an author presents a paper based on structural econometrics.