Introduction to Structural Econometrics

Lesson 5: Models with a state-space representation

Guillaume Wilemme

Aix-Marseille School of Economics - PhD course

Nov 2017

Table of contents

Plan

1. General questions

2. Reminder of simulation methods
   - Maximum simulated likelihood
   - Methods of Moments

3. Models with a state-space representation

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

1. General questions

2. Reminder of simulation methods
   - Maximum simulated likelihood
   - Methods of Moments

3. Models with a state-space representation

MSL

- Take the likelihood $\mathcal{L}(y, x, \theta^I)$ (or $\mathcal{L}(\mathcal{D}, \theta^I)$)
- If the likelihood writes as an integral (or expectation),

$$\mathcal{L}(y, x, \theta^I) = \int \tilde{\mathcal{L}}(y, x, \epsilon, \theta^I) d\mathcal{F}(\epsilon),$$

- then it can be simulated, meaning we can use an approximation of the likelihood as

$$\frac{1}{S} \sum_s \tilde{\mathcal{L}}(y, x, \epsilon^s, \theta^I)$$

- A MSL estimator is

$$\hat{\theta}_{MSL,1} = \text{argmax} \sum_s \tilde{\mathcal{L}}(y, x, \epsilon^s, \theta^I)$$

- Another could be

$$\hat{\theta}_{MSL,2} = \text{argmin} \left(\sum_s \frac{\partial \tilde{\mathcal{L}}(y, x, \epsilon^s, \theta^I)}{\partial \theta^I} \right)^2$$

MM

- First, we note Y for the random variable, and y for the realization. By assumption, x is, most of the time, taken as given.
- MM build on moment conditions or constraints that enables us to recover θ^I:
$$\mathbb{E}[g(x, \theta^I)] = 0$$
  - expectation can be understood as $\mathbb{E}_Y$ or $\mathbb{E}_\epsilon$ (as the model gives a bijection between Y and ϵ)
- The dimension of g is the number of moments

MM (II)

- Example with the linear regression
  - identifying moment condition is the exogeneneity condition $\mathbb{E}[x\epsilon] = 0$
- Interpretation of MM as distance between "theoretical" (according to the model) moments and empirical (purely data-based) moments
- as if $g(x, \theta^I) = m^{th}(x, \theta^I) - m^{emp}(y, x)$
  - example: the theoretical moment $\mathbb{E}[x\epsilon] = 0$ must be equal to its empirical equivalent meaning $\frac{1}{N}\sum_i x_i(y_i - \beta x_i)$
  - gives the OLS formula
  - an equivalent approach is to choose the theoretical moment as $\mathbb{E}[xY] = \frac{1}{N}\beta x'x$, with the empirical counterpart $\frac{1}{N}\sum_i x_i y_i$
- Max. likelihood using the derivative of likelihood (or the score) as the moment to match with 0

Difference between MM and GMM

- In the MM, θ^I is just-identified, meaning $dim(\theta) = dim(g)$.
  - example: with the linear model, θ^I has a closed-form which is the OLS formula
- If $dim(\theta) < dim(g)$, then you can use GMM
  - It is very unlikely that there exists a θ^I such that $\mathbb{E}[g(x, \theta^I)] = 0$
  - Idea is to get these moment conditions as close to zero as possible
  - in practice, use a distance measure with possible weights

Estimation with simulation

- As for maximum likelihood, you can use simulation to avoid computing integrals

$$m^{th}(y, x, \theta^l) = \int \tilde{m}^{th}(y, x, \epsilon, \theta^l) d\mathcal{F}(\epsilon)$$

- use the simulator

$$\frac{1}{S} \sum_s \tilde{m}^{th}(y, x, \epsilon^s, \theta^l)$$

Plan

Definition

$$M_t = F(S_t, u_t)$$
$$S_t = G(S_{t-1}, e_t)$$

- Two types of variables: measurement and state variables.
- Two different interpretation for the error terms, supposed independent.
- Some variables are latent. In economics there are many variables that we do not observe but that interest us. They can be interpreted as "fundamentals" (shadow price, potential output, individual ability, natural unemployment rate)

Intro to Kalman filters

- Take the simple linear case with gaussian errors

$$M_t = AS_t + u_t$$
$$S_t = BS_{t-1} + e_t$$

- Start with S_0 (possibly random with a known variance)
- We have first an estimate of S_1 as BS_0
- After observing M_1, we deduce an estimate of the prediction error $M_1 - ABS_0$
- M_1 brings additional information to estimate S_1
- The Kalman filter consists in updating the estimate of S_t using this information (using a formula such that the error variance of the estimate is minimum)
- On top of that, A, B and the covariance matrices of u and e can be estimated through maximum likelihood

Likeli

Intro to Kalman filters (I)

- On top of that, A, B and the covariance matrices of u and e can be estimated through maximum likelihood
- The likelihood writes recursively:

$$\mathbb{P}(M_1, ..., M_T) = \mathbb{P}(M_1)\mathbb{P}(M_2|M_1), .., \mathbb{P}(M_T|M_1, .., M_{T-1}).$$

- Example of a paper: Krusell, Ohanian, Rios-Rull and Violante, ECTA, 2000
  - parameters of the production function (efficiency of skilled/unskilled labor), and relative price of capital are latent and evolve over time
  - to study skill-bias technological change