Structural Estimation

Numerical Methods

Cezar Santos

FGV/EPGE

Acknowledgements

- This set of slides is heavily based on notes by Georg Dürnecker.

Estimation

- There are several approaches to this:
- Maximum likelihood estimation - either classical or using Bayesian methods
  [In-class exercise will be an application]
- **G**eneralized **M**ethod of **M**oments (GMM)
- Simulation-based methods: SMM, EMM, Indirect Inference
- Here we'll deal with the latter class of estimators

Suggested Readings

- Structural estimation of dynamic stochastic (GE) models
  - Canova, F., 2007, *"Methods for Applied Macroeconomic Research"*, Chap. 5, 6, 9, 11
  - Adda, J. and R. Cooper, 2003, *"Dynamic Economics"*, Chap. 4 + Applications
- Simulation-based estimation techniques
  - Gourieroux, C., A. Monfort and E. Renault, 1993, *"Indirect Inference"*, Journal of Applied Econometrics, 8, pp. S85-S118
  - Gallant, R. A. and Tauchen, G., 1996, *"Which Moments to Match?"*, Econometric Theory, 12, pp. 657-681.
  - Smith, A. A., 1993, *"Estimating Nonlinear Time Series Models Using Simulated Vector Autoregressions"*, J. of Applied Econometrics, 8, pp. 63-84
  - For a discussion of simulation-based estimators see Gourieroux, C. and Monfort, A., 1996, *"Simulation Based Econometric Methods"*, Oxford University Press

- Why shall we use simulation-based methods?
- ML: Requires the (closed-form) representation of the likelihood function. Often models are too complex $\Rightarrow$ ML is intractable *[Can use Simulated Maximum Likelihood instead]*
- GMM: Performs poorly in small sample (post-war quarterly data is considered small sample!). Can not deal with unobserved variables, Often problems with identification and choice of orthogonality conditions.

Simulation-Based Estimators

- Developed by Gourieroux and Montfort (1996), Pakes and Pollard (1989), McFadden (1989), ...
- Main idea: We choose the parameters of the model to **minimize the distance** between functions of actual and simulated data
- Those functions can be:
  - Moments such as variances, covariances, autocorrelations $\Rightarrow$ SMM
  - VAR coefficients, impulse response functions $\Rightarrow$ Indirect inference

Simulated Method of Moments

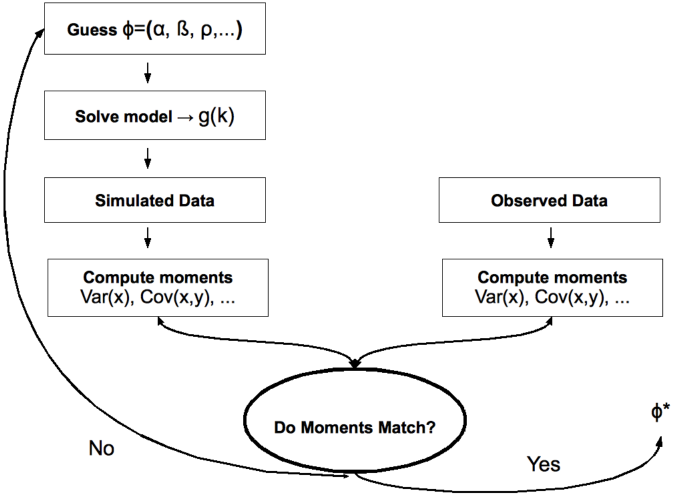
- Developed by McFadden (1989), Lee and Ingram (1991) ...
- Let θ denote the $K \times 1$ vector of structural parameters associated with our economic model
- In the DSGE case that would be $\theta = (\beta, \alpha, \rho, \delta, ...)$
- Let $\{y_t(\theta_0)\}_{t=1}^T$ be a sequence of **observed data** and let $\{\tilde{y}_t(\theta)\}_{t=1}^T$ be a set of **simulated data**, each of length T
- θ_0 denotes the "true" but unobserved set of parameters
- The simulations are done by fixing θ and by using T draws of the stochastic element

- Denote by $\mu[y_T]$ a $J \times 1$ vector of functions of the **observed data**
- ... and denote by $\mu[\tilde{y}_T(\theta)]$ the same $J \times 1$ vector of functions computed using the **simulated time series** $\{\tilde{y}_t(\theta)\}$, once the $K \times 1$ vector of parameters θ is chosen.
- The SMM estimator $\hat{\theta}_T^S(W)$ is defined as

$$\hat{\theta}_T^S(W) = \arg\min_{\theta} \left[\mu[y_T] - \frac{1}{S} \sum_{s=1}^{S} \mu[\tilde{y}_T^s(\theta)] \right]' \times W_T \times \left[\mu[y_T] - \frac{1}{S} \sum_{s=1}^{S} \mu[\tilde{y}_T^s(\theta$$

- where W_T is a $J \times J$, symmetric and positive definite weighting matrix
- For each θ we average the simulated moments over S draws $\frac{1}{S}\sum_{s=1}^{S} \mu[\tilde{y}_T^s(\theta)]$ to minimize the error due to changing random draws
- Difference to GMM: Instead of calculating the theoretical moments $\mu[y_T]$ we approximate them with simulations

Simulated Method of Moments



Properties of SMM

- Conditions needed:
  - y_t and $\tilde{y}_t(\theta)$ are stationary and ergodic
  - $\mu[y_T] \to^P \mu_y$ as $T \to \infty$ and $\mu[\tilde{y}_T^S(\theta)] \to^P \mu_{\tilde{y}}(\theta)$ as $S \to \infty$ (Consistency).
  - Under the null that the model is true, there exists a θ^* such that $\mu_y = \mu_{\tilde{y}}(\theta^*)$ (Identifiability)
- Then $\hat{\theta}_T^S(W)$ is consistent and asymptotically normal with

$$\sqrt{T}[\hat{\theta}_T^S(W) - \theta_0] \to N(0, Q_S(W))$$

- Where $Q_S(W))$ is the asymptotic variance-covariance matrix
- If $W_T = I$, $\hat{\theta}_T^S(W)$ consistent but inefficient. $\Rightarrow$ to gain on efficiency we use the optimal weighting matrix W_T^*
- We'll discuss shortly how to compute W_T^*

SMM Algorithm

1. Take the relevant data series and compute $\mu[y_T]$
2. Choose a θ^0 and solve the model for its equilibrium
3. Simulate the model S times and produce $\tilde{y}(\theta^0)$ and compute $\mu[\tilde{y}_T^S]$
4. Evaluate the objective function and get θ^1 using grid-search or gradient methods
5. Repeat the previous two steps until

$$\left[\mu[y_T] - \frac{1}{S}\sum_{s=1}^{S}\mu[\tilde{y}_T^s(\theta)]\right]' \times W_T \times \left[\mu[y_T] - \frac{1}{S}\sum_{s=1}^{S}\mu[\tilde{y}_T^s(\theta)]\right] \leq \varepsilon$$

$$\text{or} \qquad ||\theta^i - \theta^{i-1}|| \leq \varepsilon$$

or both with ε being small

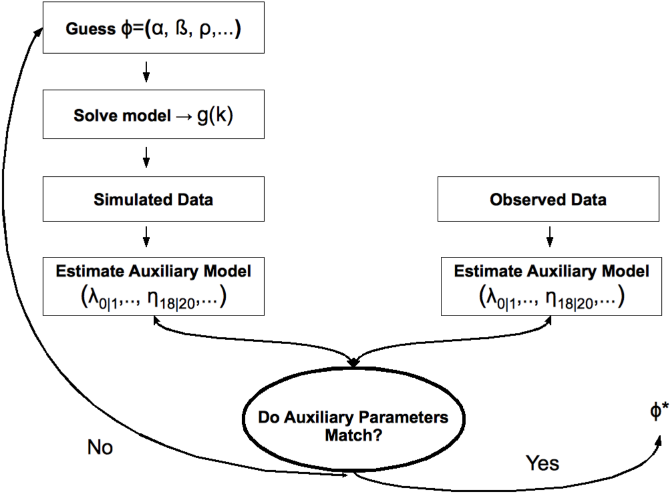
Some Issues

- In step 3 we need to choose a sequence of shocks for the model. **IMPORTANT:** Must use the same sequence of shocks during the iterations otherwise don't know if objective function changes because parameters change or shocks change
- How to get standard errors for $\hat{\theta}_T^S(W)$? Two alternatives:
  - Compute the Hessian of the objective function (will be dealt with later)
  - Use a Monte Carlo approach: Repeat algorithm for different sequences of the shocks, plot the histogram of the resulting θ and compute standard errors from this distribution

Indirect Inference (I.Inf)

- Developed by Gourioux et al. (1993) and Smith (1993)
- Indirect inference is a generalization of SMM where μ consists of generic continuous functions of the data
- I.Inf is particularly convenient when the model is complex $\Rightarrow$ likelihood is intractable
- The key feature of I.Inf is that it uses a simple *auxiliary* model to "measure the distance between the model and the data"
- Main idea: (1) Auxiliary model is estimated both on the actual and on simulated data $\rightarrow$ (2) I.Inf chooses the **structural parameters** so that the auxiliary parameters from the **simulated data** are as close as possible to those obtained on **actual data**.

Indirect Inference



- Choice of the auxiliary model is key! It has to be good!! More specifically:
  1. Has to provide a good statistical representation of the data
  2. Has to represent well the economic processes captured by the structural model
  3. Easy and quickly to compute
- Points (1)-(2) relate to the proper identification of structural parameters (more on that shortly)
- If (1)-(2) are not satisfied $\Rightarrow$ Objective function will not be responsive to variation in structural parameters
- Examples for auxiliary models:
  1. Structural impulse response functions (DSGE literature)
  2. Wage regressions and hazard rate models (Applied labor)
  3. Discrete choice models (IO)

- Suppose $\Theta(y_T, \beta)$ is the auxiliary model at hand, where β is a vector of auxiliary parameters and y_T is the actual data
- Let $\hat{\beta}_T$ be estimator of β computed from the observed data
- The way $\hat{\beta}_T$ is computed, clearly depends on the structure of the auxiliary model: Impulse responses: SVAR estimation / Hazard model: ML estimation / etc.
- Notice that the null hypotheses underlying the estimation is that the observed data are generated by the model at the true value of the parameter θ_0
- Given a set of structural parameters θ we solve and simulate the model to get artificial data $\tilde{y}_T(\theta)$
- The auxiliary model is then estimated out of the simulated data. The associated estimator of the auxiliary parameters is $\tilde{\beta}(\tilde{y}_T(\theta))$

- For the same θ we repeat the last step S times and compute

$$\hat{\beta}_S(\theta) = \frac{1}{S}\sum_{s=1}^{S} \tilde{\beta}(\tilde{y}_T^s(\theta))$$

- The averaging is done to minimize the error due to changing draws

- The indirect inference estimator of the model's structural parameters $\hat{\theta}_T^S(W)$ is the solution to

$$\hat{\theta}_T^S(W) = \arg\min_{\theta}\left[\hat{\beta}_T - \hat{\beta}_S(\theta)\right]' \times W_T \times \left[\hat{\beta}_T - \hat{\beta}_S(\theta)\right]$$

- where again W_T is a symmetric, non-negative definite weighting matrix