GMM, Indirect Inference and Bootstrap

Indirect Inference

Willi Mutschler

TU Dortmund

Winter 2015/2016

Improve

- Notation: β vs. θ vs tilde hat etc
- Slide for binding function: with analytical example and with monte carlo simulation
- Comparison with GMM
- Visualize idea
- Slide for weighting matrix and example!

Anthony Smith, Jr. (New Palgrave Dictionary of Economics):

> *Indirect inference is a simulation-based method for estimating the parameters of economic models. Its hallmark is the use of an auxiliary model to capture aspects of the data upon which to base the estimation. The parameters of the auxiliary model can be estimated using either the observed data or data simulated from the economic model. Indirect inference chooses the parameters of the economic model so that these two estimates of the parameters of the auxiliary model are as close as possible.*

- Economic model

$$y_t = G\left(y_{t-1}, x_t, u_t; \beta\right), \quad t = 1, \ldots, T$$

- Exogenous variables x_t and endogenous variables y_t
- Random errors u_t, i.i.d. with cdf F
- Parameter vector β of dimension K
- Let standard estimation methods for β be intractable
- It must be possible (and easy) to simulate $y_1, \ldots, y_T$ given y_0 (assumed to be known), $x_1, \ldots, x_T$ and β

- The true model is too complicated for estimation of β (but we can simulate data)
- Instead estimate an auxiliary model with parameter vector θ
- The dimension L of θ must be at least as large as the dimension K of β
- The auxiliary model must be
  - "suitable" (but is allowed to be misspecified): Existence of one-to-one-binding function: maps parameters of economic model into parameters of auxiliary model $\theta = h(\beta)$, binding function usually unknown, must be evaluated by Monte-Carlo simulations
  - easy and fast to estimate
- Often, the auxiliary model is a standard time series model, best possible approximation or some flexible form that encompasses approximations to a wide range of models or distributions within class of interest
- what we call moments in GMM, the "moments" that guide estimation are parameters of auxiliary model

- For given β (and $y_0, x_1, \dots, x_T$), the auxiliary model's parameters θ are estimated
  1. from the observed data $x_1, \dots, x_T, y_1, \dots, y_T$, resulting in estimator $\hat{\theta}$
  2. from H simulated datasets $x_1, \dots, x_T, \tilde{y}_1^{(h)}, \dots, \tilde{y}_T^{(h)}$ for $h = 1, \dots, H$, resulting in estimators $\tilde{\theta}^{(h)}(\beta)$
- Define

$$\tilde{\theta}(\beta) = \frac{1}{H} \sum_{h=1}^{H} \tilde{\theta}^{(h)}(\beta)$$

  this is the approximation of unknown binding function.
- $\tilde{\theta}^{(h)}(\beta)$: set of statistics that capture or summarize certain features of data, we want to reproduce this set of statistcs as losely as possible

- Compute the difference between the vectors $\hat{\theta}$ and $\tilde{\theta}(\beta)$

$$Q(\beta) = \left(\hat{\theta} - \tilde{\theta}(\beta)\right)' W \left(\hat{\theta} - \tilde{\theta}(\beta)\right)$$

where W is a positive definite weighting matrix

- The indirect inference estimator of β is

$$\hat{\beta} = \arg\min Q(\beta)$$

Indirect inference

Remarks

- The simulations have to be done with the same set of random errors
- Indirect inference is similar to GMM: the auxiliary parameters are the "moments"
- The asymptotic distribution of $\hat{\beta}$ can be derived (see Gourieroux et al., 1993)
- The weighting matrix W can be chosen optimally such that the estimator becomes asymptotically efficient

- Consider the $MA(1)$ process

$$y_t = \varepsilon_t - \beta \varepsilon_{t-1}$$

  with $\varepsilon_t \sim N(0,1)$ and $\beta = 0.5$ for $t = 1, \ldots, 250$
- The maximum likelihood estimator $\hat{\beta}_{ML}$ is not trivial
- Indirect inference estimator $\hat{\beta}_{II}$ of β ?
- Auxiliary model: $AR(3)$ with parameters θ
- No weighting, the matrix W is the identity matrix

Compare the distribution of $\hat{\beta}_{ML}$ and $\hat{\beta}_{II}$

- Step 1: Simulate a time series $y_1, \ldots, y_{250}$
- Step 2: Compute $\hat{\beta}_{ML}$
- Step 3: Estimate $\hat{\theta}$ from $y_1, \ldots, y_{250}$
- Step 4: For given β, simulate 10 paths $\tilde{y}_1^{(h)}, \ldots, \tilde{y}_{250}^{(h)}$
- Step 5: Estimate $\tilde{\theta}(\beta)$ from the simulated paths
- Step 6: Repeat steps 4 and 5 for different β until the difference between $\hat{\theta}$ and $\tilde{\theta}(\beta)$ is minimized
- Step 7: Save $\hat{\beta}_{II}$ and start again at step 1