

Structural Estimation: Simulated Method of Moments and Indirect Inference

Christine Braun

Last Time

- Last time we talked about GMM
 - we had closed for solutions to moments
 - this is rarely the case
- **Today:**
 - Simulated Method of Moments (SMM)
 - Indirect Inference

Simulated Method of Moments

- Same notation as GMM
- Y_t : n-dimensional vector of observations
 - t does not have to mean time, could be people
 - unemployment, wages, duration, observables characteristics, ect..
- θ_0 : vector of true parameters
- $g(Y_t, \theta)$: a vector valued function of data and parameters
 - such that $E[g(Y_t, \theta_0)] = 0$

Simulated Method of Moments

- Basic idea is the same:

$$E[g(Y_t, \theta)] \rightarrow \frac{1}{T} \sum_{t=1}^T g(Y_t, \theta)$$

- Last time we had

$$g(Y_i, \theta) = \begin{bmatrix} u_i - \frac{\delta}{\delta + \lambda[1 - F(w_R; \mu, \sigma)]} \\ tu_i - \frac{1}{\lambda[1 - F(w_R; \mu, \sigma)]} \\ te_i - \frac{1}{\delta} \\ w_i - E[w; \mu, \sigma] \\ w_i^2 - E[w^2; \mu, \sigma] \end{bmatrix}$$

- What do we do if we don't have a closed for $E[w^2; \mu, \sigma]$ or any other moment?

Simulated Method of Moments

- We will replace g with an estimate \hat{g}

$$E[g(Y_t, \theta)] \rightarrow \frac{1}{T} \sum_{t=1}^T g(Y_t, \theta) \rightarrow \frac{1}{T} \sum_{t=1}^T \hat{g}(Y_t, \theta)$$

- For a given θ we simulate $\{\hat{Y}_s\} = \{\hat{u}_s, \hat{t}u_s, \hat{t}e_s, \hat{w}_s\}$, then

$$\hat{g}(Y_i, \theta) = \begin{bmatrix} u_i - N_s^{-1} \sum_s \hat{u}_s \\ tu_i - N_{su}^{-1} \sum_s \hat{t}u_s \\ te_i - N_{se}^{-1} \sum_s \hat{t}e_s \\ w_i - N_{se}^{-1} \sum_s \hat{w}_s \\ w_i^2 - N_{se}^{-1} \sum_s \hat{w}_s^2 \end{bmatrix}$$

where N_s is the number of obs. in the simulated data.

Simulated Method of Moments

- The SMM estimate of θ_0 is

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \left(\frac{1}{T} \sum_{t=1}^T \hat{g}(Y_t, \theta) \right)' W \left(\frac{1}{T} \sum_{t=1}^T \hat{g}(Y_t, \theta) \right)$$

where W is the weighting matrix.

- The asymptotic distribution of SMM Estimator is

$$\sqrt{n}(\hat{\theta} - \theta_0) \rightarrow N(0, (J' W J)^{-1} J' W \Omega W J (J' W J)^{-1})$$

- $J = E[\nabla_{\theta} g(Y_t, \theta)]$: jacobian of g
- $\Omega = E[g(Y_t, \theta_0)g(Y_t, \theta_0)']$
- If we have $W = \Omega^{-1}$

$$\sqrt{n}(\hat{\theta} - \theta_0) \rightarrow N(0, (J' \Omega J)^{-1})$$

Simulated Method of Moments

- The general procedure is the same
 1. choose a weighting matrix
 2. estimate
 3. calculate $\hat{W} = \hat{\Omega}^{-1}$
 4. estimate
 5. repeat if necessary
- **What's New:** we need to simulate data

Model

- We will still use the simple search model
- We still have the same identification problem
 - $\hat{w}_R = \min_i w_i$
 - and set $r = 0.05$
- **Parameters to estimate:** $\lambda, \delta, \mu, \sigma$
- **Moments to match:**
 1. unemployment rate
 2. unemployment duration
 3. employment duration
 4. first moment of wages
 5. second moment of wages

Simulating Data

- What data do we need to simulate?
 - unemployment dummy
 - unemployment duration
 - wages
 - employment duration
- We will simulate the model in steady state
- How many observations should we simulate?
 - no perfect answer
 - I usually do the same as in the observed data
- **Very Important Note:** you must set a seed so that each simulation is created with the same random numbers. `rng(·)`

Simulating Data: Unemployment \hat{u}_s

- The steady state unemployment rate is

$$urate = \frac{\delta}{\delta + \lambda[1 - F(w_R)]} \in [0, 1]$$

- The probability of an individual being unemployed
- Simulation
 - N_s draws from $udraw_s \sim Unif[0, 1]$
 - Then create unemployment dummy

$$\hat{u}_s = \begin{cases} 1 & udraw_s < urate \\ 0 & udraw_s \geq urate \end{cases}$$

- check $N_s^{-1} \sum \hat{u}_s \approx urate$

Simulating Data: Unemployment duration $\hat{t}u_s$

- Unemployment duration follows an exponential dist.

$$G(tu) = 1 - \exp(-\lambda[1 - F(w_R)]tu)$$

- **Note:** in matlab all functions of the exponential dist. use as an input the mean rather than the rate

$$mean = \frac{1}{\lambda[1 - F(w_R)]}$$

- Simulation 1

- N_s draws from $tudraw_s \sim Unif[0, 1]$
- $\hat{t}u_s = \text{expinv}(tudraw_s, 1/(\lambda[1 - F(w_R)]))$

- or Simulation 2

- N_s draws from $\hat{t}u_s = \text{exprnd}(1/(\lambda[1 - F(w_R)]))$
- **check** $N_s^{-1} \sum \hat{t}u_s \approx 1/(\lambda[1 - F(w_R)])$

Simulating Data: Employment duration \hat{te}_s

- Employment duration follows an exponential dist.

$$G(te) = 1 - \exp(-\delta te)$$

- Simulation 1

- N_s draws from $tedraw_s \sim Unif[0, 1]$

- $\hat{te}_s = \text{expinv}(tedraw_s, 1/\delta)$

- or Simulation 2

- N_s draws from $\hat{te}_s = \text{exprnd}(1/\delta)$

- **check** $N_s^{-1} \sum \hat{te}_s \approx 1/\delta$

Simulating Data: Wage \hat{w}_s

- Wage offer distribution: $F(w) \sim \ln N(\mu, \sigma)$
- Accepted wage distribution

$$F^A(w) = \frac{f(w)}{1 - F(w_R)}$$

- We need to simulate from the accepted wage distribution
- Simulation
 - N_s draws from $wdraw_s \sim Unif[F(w_R), 1]$
 - $\hat{w}_s = \text{logniv}(wdraw_s, \mu, \sigma)$
- check

$$N_s^{-1} \sum \hat{w}_s \approx \exp\left(\mu + \frac{\sigma^2}{2}\right) \frac{\Phi\left(\frac{\mu + \sigma^2 - \ln(w_R)}{\sigma}\right)}{1 - \Phi\left(\frac{\ln(w_R) - \mu}{\sigma}\right)}$$

Matlab Estimation

- using data4.csv
- File 1: SE4_main
 - first we will work on simulating data, part 1
 - then estimate
- File 2: simulate_data.m
 - inputs ?
 - outputs ?
- File 3: g_function_sim.m
 - inputs ?
 - outputs ?
- File 4: SMM.m
 - inputs ?
 - outputs ?

Matlab Estimation: Part 1 Answer

- `guess = [0.7, 0.2, 1, 0.5]`
- `rng(7890)`
- $n_s = 10000$

Moment	Theoretical	Simulated Data
unemp. rate	0.9686	0.9680
unemp. dur.	154.3150	154.1840
emp. dur.	5.0000	4.7410
$E[w]$	10.5764	10.6417

Matlab Estimation: Part 2 & 3 Answer

Parameter	SMM $W = \hat{\Omega}^{-1}$		SMM $W = I/\text{mean}(\text{data})$	
	Estimate	Std. Err.	Estimate	Std. Err.
λ	0.2883	0.0112	0.2879	0.0112
δ	0.0216	0.0002	0.0216	0.0002
μ	2.2058	0.0196	2.2066	0.0195
σ	0.4021	0.0088	0.4018	0.0088

$$\hat{b} = 0.5670$$

SMM and Indirect Inference

- **Indirect Inference:** a simulation-based method for estimating parameters
 - useful when the likelihood function or moments are not analytically tractable or difficult to evaluate
 - example: models with latent variables
- **Auxiliary Model:** aspects of the data that can be calculated in the observed and simulated data
- **Main Idea:** Minimize the distance between the auxiliary model of the observed data and the simulated data
- SMM is a special case of Indirect Inference
 - auxiliary model: moments of the data

Indirect Inference

- Other examples of an auxiliary model
 - the likelihood function
 - coefficients from regressions
 - impulse response functions
 - coefficient of interest from an RCT
 - quantile regressions