Lecture 2: Structural Estimation
using Simulated Method of Moments (SMM)

Tim D. Maurer[1]

Workshop on Life-cycle Models and Pensions

September 26, 2024

[1] Department of Economics, Norwegian School of Economics. **E-mail**: `tim.maurer@nhh.bo`.

Structural Estimation

- We know how to solve and simulate our life-cycle model
- How can we estimate it? We need
  - Data on (some) states and choices
- Two standard approaches
  1. Maximum likelihood (ML)
  2. General Method of Moments (GMM)
     - Requires that the model moment function is known analytically
- Simulated versions
  1. Maximum Simulated Likelihood (MSL)
  2. Method of Simulated Moments (SMM)

Example of our model in L1

- State: a_0^i
- Choices: $c_{j,t}^i$
- Parameter to estimate: $\theta = \{\beta\}$
- Calibration ("known"): ρ, μ, σ, w, r

SMM

- Let Λ^d be a vector of J moments in the data
  - Could be avg., var, cov, regression-coefs, etc.
- Let $\Lambda^m(\theta)$ be a vector of the **same** moments calculated using **simulated** data from the model solved with parameters θ
- Let $e(\theta)$ be the error vector as the $J \times 1$ vector of moment error functions $e_j(\theta)$ of the jth moment error, where

$$e_j(\theta) = \frac{\Lambda_j^d - \Lambda_j^m(\theta)}{\Lambda_j^d} \quad or \quad e_j(\theta) = \Lambda_j^d - \Lambda_j^m(\theta)$$

The SMM estimator is then defined by:

$$\hat{\theta} = \arg\min_{\theta} e(\theta)' W e(\theta),$$

where W is a weighting matrix. W is $J \times J$, where J is the number of moments.

Weighting matrix

Typical choices are

1. Theoretically optimal (see [Adda and Cooper, 2003] for formula)
2. Diagonal matrix with inverse of (bootstrapped) empirical variances of the moments
3. Freely chosen to focus on fitting some specific dimensions of the data
4. Identity matrix $W = I$ as your weighting matrix. This changes the criterion function to a simple sum of squared error functions:

$$\hat{\theta} = \arg\min_{\theta} e(\theta)' e(\theta),$$

   If the problem is well conditioned and well identified, then your SMM estimates will not be greatly affected by this simplest of weighting matrices.

Asymptotics

The SMM estimator is consistent and asymptotically normal under standard assumptions

$$\sqrt{N}\left(\hat{\theta}-\theta_0\right) \to \mathcal{N}\left(0, \left(1+S^{-1}\right) V\right)$$

where θ_0 are the true parameters and S the number of simulations.

Standard formulas for V is

$$V = \left(G'WG\right)^{-1} G'W\Omega W'G \left(G'WG\right)^{-1},$$

where $G = -\frac{\partial \Lambda^m(\theta)}{\partial \theta}$ is the Jacobian of the objective function and Ω is the variance-covariance matrix of the moments in the data.

$\Rightarrow$ Standard errors are large if large changes in θ imply small changes in the objective function

Identification and Simulation Pitfalls

- Is there enough variation in the data to identify θ?
  Very hard to prove anything if the model is strongly non-linear
- Requires at least the same number of moments as parameters
- Problems
  - The objective function might have multiple minima
  - The objective function could be very flat in some directions and steep in others (ill-conditioned)
- Graphical inspection is useful: Plot the objective function in the neighborhood of the found optimum
- Use more data
  1. Quantitatively: More agents, more time periods
  2. Qualitative: New data types, e.g., natural experiments

Simulation Pitfalls

- FIX the seed (or draws) to void unnecessary noise
- Ill-conditioned objective function
  - Gradient-based numerical optimization will likely fail $\Rightarrow$ use gradient-free optimization, e.g, Nelder-Mead

Example - Model from L1

Agent i born at time t, who starts to consume in period $t+1$, faces the following problem:

$$\max_{\left\{c_t^i, a_t^i\right\}} \quad U = \sum_{j=1}^{T} \beta^t u(c_{j,t}^i)$$

$$s.t.$$

$$c_{j,t}^i + a_{j,t}^i = w_{t+j} l_{j,t} + (1 + r_{t+j})\, a_{j-1,t}^i \tag{1}$$

$$\text{with}$$

$$l_{j,t} = \begin{cases} 1 & \text{if } j \leq T_r \\ 0 & \text{otherwise} \end{cases}$$

$$a_{0,t} \sim \text{Lognormal}\left(\mu, \sigma^2\right)$$

$$u(c) = \begin{cases} \dfrac{c^{1-\rho} - 1}{1-\rho} & \rho \geq 0, \rho \neq 1 \\ \log(c) & \rho = 1, \end{cases}$$

Closed-form Solution

We simply borrow the solution from Lecture 1 under $w_t = w$ and $r_t = r$:

$$c_{1,t}^{i,*} = \frac{w \sum_{j=1}^{T} \frac{l_{j,t}}{(1+r)^j} + \frac{a_{0,t}^i}{1+r}}{\sum_{j=1}^{T} \frac{[\beta(1+r)]^{\frac{j-1}{\rho}}}{(1+r)^j}} \tag{2}$$

$$c_{j,t}^{i,*} = c_{1,t}^{i,*} [\beta(1+r)]^{\frac{j-1}{\rho}} \quad \text{for } j > 1 \tag{3}$$

Insert the solution for consumption $c_{j,t}^*$ into the budget constraint (1) to back out the solution for savings:

$$a_{j,t}^* = w_{t+j} l_{j,t} + (1 + r_{t+j}) a_{j-1,t} - c_{j,t}^* \tag{4}$$

Goal and Data

Goal: We want to estimate $\theta = \{\beta\}$

- Rest of the parameters are known: $\rho, T, T_r, \mu, \sigma, w, r$

Data

1. We simulate the model with the "true" parameters.
2. The outcome is our "empirical" data set.
3. We therefore know exactly what our estimation should lead to.

Data moments Λ^d

We want to calculate three moments:

1. Mean consumption at age 5: $\Lambda_1^d = \frac{1}{N} \sum_{i=1}^{N^d} c_{5,t}^{i,d}$
2. Mean consumption at age 10: $\Lambda_2^d = \frac{1}{N} \sum_{i=1}^{N^d} c_{10,t}^{i,d}$
3. Mean consumption at age 15: $\Lambda_3^d = \frac{1}{N} \sum_{i=1}^{N^d} c_{15,t}^{i,d}$

where N^d is the number of agents in the data. Hence, the vector Λ^d stacking the individual data moments reads

$$\Lambda^d = \begin{pmatrix} \Lambda_1^d \\ \Lambda_2^d \\ \Lambda_3^d \end{pmatrix}$$

Note: In practice, it is tricky how to choose moments. Trial-and-error is required.

Simulated moments Λ^m and W

Run S number of simulation of N^d simulated agents

1. Mean consumption at age 5: $\Lambda_1^m(\theta) = \frac{1}{S}\sum_{s=1}^{S}\frac{1}{N^d}\sum_{i=1}^{N^d} c_{5,t}^{i,s,*}$
2. Mean consumption at age 10: $\Lambda_2^m(\theta) = \frac{1}{S}\sum_{s=1}^{S}\frac{1}{N^d}\sum_{i=1}^{N^d} c_{10,t}^{i,s,*}$
3. Mean consumption at age 15: $\Lambda_3^m(\theta) = \frac{1}{S}\sum_{s=1}^{S}\frac{1}{N^d}\sum_{i=1}^{N^d} c_{15,t}^{i,s,*}$

Hence, the vector Λ^m stacking the individual simulated moments reads

$$\Lambda^d = \begin{pmatrix} \Lambda_1^m(\theta) \\ \Lambda_2^m(\theta) \\ \Lambda_3^m(\theta) \end{pmatrix}$$

Note that we have a closed-form solution for $\mathbb{E}\left[c_{1,t}^{i,*}\right]$

$$\mathbb{E}\left[c_{1,t}^{i,*}\right] = \frac{w\sum_{j=1}^{T}\frac{l_{j,t}}{(1+r)^j} + \frac{e^{\mu+\frac{\sigma^2}{2}}}{1+r}}{\sum_{j=1}^{T}\frac{[\beta(1+r)]^{\frac{j-1}{\rho}}}{(1+r)^j}} \tag{5}$$

Example of SMM Estimator

Using $W = I$ estimator is given by:

$$\hat{\theta} = \arg\min_{\theta} \begin{pmatrix} e_1(\theta) \\ e_2(\theta) \\ e_3(\theta) \end{pmatrix}' \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e_1(\theta) \\ e_2(\theta) \\ e_3(\theta) \end{pmatrix} \tag{6}$$

where

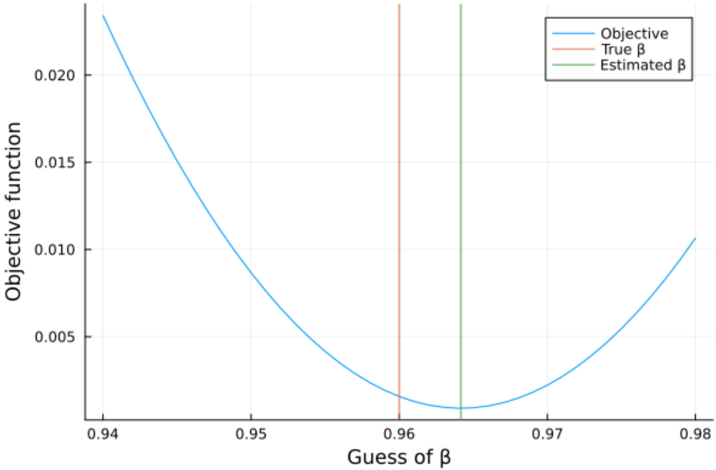
$$\begin{pmatrix} e_1(\theta) \\ e_2(\theta) \\ e_3(\theta) \end{pmatrix} = \begin{pmatrix} \Lambda_1^d - \Lambda_1^m(\theta) \\ \Lambda_2^d - \Lambda_2^m(\theta) \\ \Lambda_3^d - \Lambda_3^m(\theta) \end{pmatrix} \quad or \quad \begin{pmatrix} e_1(\theta) \\ e_2(\theta) \\ e_3(\theta) \end{pmatrix} = \begin{pmatrix} \frac{\Lambda_1^d - \Lambda_1^m(\theta)}{\Lambda_1^d} \\ \frac{\Lambda_2^d - \Lambda_2^m(\theta)}{\Lambda_2^d} \\ \frac{\Lambda_3^d - \Lambda_3^m(\theta)}{\Lambda_3^d} \end{pmatrix} \tag{7}$$

"Known parameters"

Parameter	Description	Value	Origin
Timing			
T	Maximum age of life	20	Model period $\approx$ 4 years
T_R	Retirement Age	15	
Prices			
r	Interest rate	0.13	To match an annual int. rate of 3%
w	Wage	1	Normalization
Preferences			
ρ	RRA / Inverse IES	2.0	Standard value
Distribution			
μ	Location	0	
σ	Scale	1	

Estimation Results

Based on "true" $\beta = 0.96$



References I

Adda, J. and Cooper, R. (2003).
Dynamic Economics: Quantitative Methods and Applications.
The MIT Press.

Variance of SMM Estimator

The variance of the estimator is

$$V = \left(G'WG\right)^{-1} G'W\Omega W'G \left(G'WG\right)^{-1},$$

where

$$G = -\begin{pmatrix} \frac{\partial \Lambda_1^m(\theta)}{\partial \beta} \\ \frac{\partial \Lambda_2^m(\theta)}{\partial \beta(\theta)} \\ \frac{\partial \Lambda_3^m(\theta)}{\partial \beta} \end{pmatrix} \tag{8}$$

is the Jacobian and

$$\Omega = \begin{bmatrix} Var\left(\Lambda_1^d\right) & Cov\left(\Lambda_1^d, \Lambda_2^d\right) & Cov\left(\Lambda_1^d, \Lambda_3^d\right) \\ Cov\left(\Lambda_2^d, \Lambda_1^d\right) & Var\left(\Lambda_2^d\right) & Cov\left(\Lambda_2^d, \Lambda_3^d\right) \\ Cov\left(\Lambda_3^d, \Lambda_1^d\right) & Cov\left(\Lambda_3^d, \Lambda_2^d\right) & Var\left(\Lambda_3^d\right) \end{bmatrix} \tag{9}$$

is the variance-covariance matrix of the moments in the data.

Approximation of Jacobian

The Jacobian G can be approximated using is a centered second-order finite difference numerical approximation of the derivatives of the function

$$\frac{\partial \Lambda_j^m(\theta)}{\partial \beta} \approx \frac{\Lambda_j^m(\beta - h) - \Lambda_j^m(\beta + h)}{2h} \tag{10}$$

with h being a small number.

Approximation of Ω

Bootstrapping

To bootstrap the variance-covariance matrix of the empirical moments:

1. Compute the empirical moments Λ^d from the original data.
2. Resample the original dataset with replacement B-times.
3. For each bootstrap sample, calculate the moments $\Lambda^{(b)}$.
4. Estimate Ω, the variance-covariance matrix, using the bootstrapped moments.

Formula for the Covariance Matrix Ω

Let Λ^d be the vector of empirical moments computed from the b-th bootstrap sample:

$$\Lambda^{(b)} = \begin{bmatrix} \Lambda_1^{(b)} \\ \Lambda_2^{(b)} \\ \vdots \\ \Lambda_J^{(b)} \end{bmatrix}$$

The mean of the bootstrapped moments is:

$$\bar{\Lambda} = \frac{1}{B} \sum_{b=1}^{B} \Lambda^{(b)}$$

The variance-covariance matrix Ω is estimated as:

$$\hat{\Omega} = \frac{1}{B-1} \sum_{b=1}^{B} \left(\Lambda^{(b)} - \bar{\Lambda}\right) \left(\Lambda^{(b)} - \bar{\Lambda}\right)^T$$