BYU-MCL Boot Camp
Structural Model Estimation: Simulated Method of Moments (SMM)
Professor Richard W. Evans

1 Structural Model Estimation

Structural model estimation is the estimation of model parameters θ to match the model to data in some sense. We already discussed generalized method of moments (GMM) estimation of a structural model in a Python lab. Define a dynamic economic model as a system of difference equations that defines how endogenous variables $\mathbf{x}_t$ progress over time,

$$\mathbf{F}\left(\mathbf{x}_{t+1},\mathbf{x}_t,\mathbf{z}_t;\theta\right)=\mathbf{0} \tag{1}$$

where $\mathbf{z}_t$ is the vector of exogenous state variables, and θ is a vector of parameters of the model. This structural model is the **data generating process (DGP)** for the variables $\mathbf{x}_t$ and $\mathbf{z}_t$.

As an example, the DGP or structural model for the Brock and Mirman (1972) stochastic growth model is the following.

$$(c_t)^{-1}-\beta E\left[r_{t+1}(c_{t+1})^{-1}\right]=0 \tag{2}$$

$$c_t+k_{t+1}-w_t-r_t k_t=0 \tag{3}$$

$$w_t-(1-\alpha)e^{z_t}\left(\frac{K_t}{L_t}\right)^{\alpha}=0 \tag{4}$$

$$r_t-\alpha e^{z_t}\left(\frac{L_t}{K_t}\right)^{1-\alpha}=0 \tag{5}$$

$$K_t-k_t=0 \tag{6}$$

$$L_t-1=0 \tag{7}$$

$$\Pr(z_t=z_L)=\frac{1}{2}\quad\text{and}\quad\Pr(z_t=z_H)=\frac{1}{2} \tag{8}$$

Note that the parameters of the model are (α,β,z_L,z_H), where z_L and z_H are the two possible values of z_t that control both the mean and the variance of the shock process. If we substitute equations (3) through (7) into (2), the DGP can be more simply summarized as a one-equation sequence of nonlinear second-order difference equations in $\mathbf{x}_t\equiv[k_t,z_t]$.

$$\Big(w_t+r_t k_t-k_{t+1}\Big)^{-1}-\beta E\left[r_{t+1}\Big(w_{t+1}+r_{t+1}k_{t+1}-k_{t+2}\Big)^{-1}\right]=0 \tag{9}$$

When you substitute in the known policy function of $k_{t+1}=\psi(k_t,z_t)=\alpha\beta e^{z_t}k_t^{\alpha}$, Euler equation (9) reduces to a zero function version of the policy function.

$$\alpha\beta\left(w_t+r_t k_t\right)-k_{t+1}=0 \tag{10}$$

This is the equation that generates the data, and is also sometimes called the characterizing equation (or characterizing equations when it is a system of equations). Equation (9) can be represented as $\mathbf{F}\left(\mathbf{x}_{t+2}, \mathbf{x}_{t+1}, \mathbf{x}_t; \theta\right) = \mathbf{0}$, where $\theta = [\alpha, \beta, z_L, z_H]$ is the vector of model parameters. Equation (10) can be represented as $\mathbf{F}\left(\mathbf{x}_{t+1}, \mathbf{x}_t; \theta\right) = \mathbf{0}$.

The question of structural model estimation is how do we choose values for the parameter vector θ of the DGP. Once we know θ and an initial value for the system $\mathbf{x}_0 \equiv [k_{2,0}, k_{3,0}]$ we can simulate the model or generate data. But different parameter values θ will generate different data.

2 When you have the data: GMM and MLE

When the data represented in the model $(c_t, K_t, Y_t, r_t, w_t)$ are available, we can use generalized method of moments (GMM)[1] or maximum likelihood estimation (MLE) to estimate the parameters of the model. We covered GMM as an application in the Python optimization lab. We will cover MLE in the lecture on distribution fitting. Both methods use different assumptions about the DGP to estimate the parameters. Both methods have strengths and weaknesses.[2]

MLE requires the modeler to assign a specific functional form to the distribution of the shocks in the model. Once this is done, the modeler can use the distribution of the shocks and the data to create a likelihood function for the observed data. The MLE estimate of the parameter vector $\hat{\theta}_{MLE}$ is the one that maximizes the likelihood function of the observed data. GMM remains completely agnostic as to the distribution of the shocks. The GMM estimate is simply the parameter vector $\hat{\theta}_{GMM}$ that minimizes the errors in the characterizing equations.

3 When you don't have the data: Calibration and SMM

Sometimes, all the data necessary to evaluate the characterizing equations of the model (or DGP) are not available. In the Brock and Mirman (1972) model above, equation (10) shows that k_t, w_t, r_t are a minimum set of variables necessary to evaluate the characterizing equations, assuming that the productivity shocks z_t are not observable. An even weaker condition is that you might have data, but you are not confident that all of it matches up well with the data concepts represented by the variables in your model. Examples include latent variables and censored variables. In these cases, you need a way to estimate the parameters of your model without all the data in the DGP.

Edward Prescott was one of the earliest supporters of a method—called calibration—for structurally estimating the parameters of a model.[3] In its simplest form, cali-

[1] GMM was first formalized by Hansen (1982).

[2] Fuhrer et al. (1995) find evidence that MLE estimators dominate GMM estimators in a certain class of DSGE models.

[3] See Prescott and Candler (2008) and Kydland and Prescott (1982).

bration involves taking parameters from other studies (particularly, microeconomic studies), and using those parameter values in the macroeconomic model. In its most general form, calibration involves choosing parameters to make simulated output of the model match the corresponding output from the data. Simulated method of moments (SMM) is the method of estimating model parameters that most closely match simulated model moments to empirical moments.[4]

The description here of SMM follows that of Adda and Cooper (2003, p. 87). Let $\{x(z_t,\theta_0)\}_{t=1}^T$ be a sequence of observed data generated by the true shocks z_t and the true parameter vector θ_0. Let $\{x(z_t^s,\theta)\}$ for $t=1,...T$ and $s=1,...S$ be a set of S simulated time series of the data with each series being T periods long and conditional on the parameter vector θ of i parameters.

Define $\mathbf{m}(x)$ as a vector of j empirical moments that are each functions of the data $\{x(z_t,\theta_0)\}_{t=1}^T$. Let $\mathbf{m}\big(x(z_t^s,\theta)\big)$ be the sth vector of corresponding moments from the simulated data.

$$\hat{\theta}_{SMM} = \arg\min_{\theta}\left[\mathbf{m}(x) - \frac{1}{S}\sum_{s=1}^{S}\mathbf{m}\big(x(z_t^s,\theta)\big)\right]' W^{-1}\left[\mathbf{m}(x) - \frac{1}{S}\sum_{s=1}^{S}\mathbf{m}\big(x(z_t^s,\theta)\big)\right] \quad (11)$$

The SMM estimator $\hat{\theta}_{SMM}$ minimizes the distance between the vector of empirical moments $\mathbf{m}(x)$ and the average of the S simulated vectors of model moments $\frac{1}{S}\sum_{s=1}^{S}\mathbf{m}\big(x(z_t^s,\theta)\big)$. Each moment vector is $j\times 1$.

The $j\times j$ matrix W^{-1} between the two moment difference vectors in (11) is an optimal weighting matrix. This matrix provides more efficient estimation of $\hat{\theta}_{SMM}$ because some of the simulated moments may be measured more precisely than others. For this reason, the optimal weighting matrix is simply the inverse of the variance-covariance matrix of the simulated moments. In other words, moments with a higher variance across simulations will have a lower weight in the criterion function used to estimate $\hat{\theta}_{SMM}$. However, for this section, we will use the identity matrix for our weighting matrix. Although not optimal, it is consistent. It simply weights each moment equally.

A few comments on identification in SMM are important. First, for estimating i parameters, you must have at least i moments to match. That is, $i\leq j$, where j is the number of moments. Second, you want to choose moments that are closely associated with certain parameters you are estimating. For example, a discount factor influences how you trade off consumption today with consumption tomorrow, so a couple of good moments for identifying the discount factor might be $\text{corr}(c_t,c_{t+1})$ or $\text{corr}(k_t,k_{t+1})$.

Below are the steps to executing an SMM estimation.

1. Draw S vectors of T shocks each for the S simulations of the model time series. It is important to only draw these shocks once and then use them for each choice of parameters θ as you converge to $\hat{\theta}_{SMM}$.

[4] Davidson and MacKinnon (2004, pp. 383-392) call this approach the method of simulated moments (MSM). Adda and Cooper (2003, pp. 87-89) has a clean, simple explanation of SMM. In addition to the early calibration paper Kydland and Prescott (1982), SMM was further developed by McFadden (1989), Lee and Ingram (1991), and Duffie and Singleton (1993).

2. Write a constrained minimization function that takes a guess of a parameter vector θ, a matrix of S time series of T periods of shocks, a starting value for k_1, and a vector of empirical moments $\mathbf{m}(x)$ and minimizes the distance between the average of the simulated model moments and the empirical moments.

4 Exercises

Exercise 1. Estimate the four parameters of the Brock and Mirman (1972) model $(\alpha, \beta, z_L, z_H)$ described by equations (2) through (8) by SMM. Choose the four parameters to match the following six moments from the 66 periods of empirical data $\{Y_t, k_t, c_t\}_{t=1}^{66}$ in `smmdata.txt`: $\text{mean}(Y_t)$, $\text{mean}(c_t)$, $\text{var}(Y_t)$, $\text{var}(c_t)$, $\text{corr}(k_t, Y_t)$, and $\text{corr}(k_t, k_{t+1})$. In your simulations of the model, set $T = 66$ and $S = 10,000$. Start each of your simulations from $k_1 = \text{mean}(k_t)$ from the `smmdata.txt` file. Use the `scipy.optimize.minimize` constrained minimizer command with the method set to `method='TNC'` and the tolerance set to `tol=1e-10`. Input the bounds to be $\alpha, \beta \in [\varepsilon, 1-\varepsilon]$, $z_L \in [-2, 0]$, and $z_H \in [1, 3]$, and where $\varepsilon = 1e-10$. Report your solution $(\hat{\alpha}, \hat{\beta}, \hat{z}_L, \hat{z}_H)$, the vector of moment differences, the sum of squared moment differences, and the computation time.

Once you have successfully estimated the parameters of a model by SMM, the question remains of how good the estimates $\hat{\theta}_{SMM}$ are. How can you check the accuracy? The first way is to see how close your simulated average of your model moments came to their target empirical moments that you were trying to match. However, this is not sufficient because you chose the parameters to minimize that distance. The best SMM estimations match well the moments that they used for the estimation, and they match some important moments that were not used in the estimation. These "outside" moments are a key piece of evidence that your model and estimation are good.

References

Adda, Jérôme and Russell Cooper, *Dynamic Economics: Quantitative Methods and Applications*, MIT Press, 2003.

Brock, William A. and Leonard Mirman, "Optimal Economic Growth and Uncertainty: the Discounted Case," *Journal of Economic Theory*, June 1972, *4* (3), 479–513.

Davidson, Russell and James G. MacKinnon, *Econometric Theory and Methods*, Oxford University Press, 2004.

Duffie, Darrell and Kenneth J. Singleton, "Simulated Moment Estimation of Markov Models of Asset Prices," *Econometrica*, July 1993, *61* (4), 929–952.

Fuhrer, Jeffrey C., George R. Moore, and Scott D. Schuh, "Estimating the Linear-quadratic Inventory Model: Maximum Likelihood versus Generalized Method of Moments," *Journal of Monetary Economics*, February 1995, *35* (1), 115–157.

Hansen, Lars Peter, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, July 1982, *50* (4), 1029–1054.

Kydland, Finn E. and Edward C. Prescott, "Time to Build and Aggregate Fluctuations," *Econometrica*, November 1982, *50* (6), 1345–1370.

Lee, Bong-Soo and Beth Fisher Ingram, "Simulation Based Estimation of Time-series Models," *Journal of Econometrics*, February 1991, *47* (2-3), 197–205.

McFadden, Daniel, "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica*, 1989 1989, *57* (5), 995–1026.

Prescott, Edward C. and Graham V. Candler, "Calibration," in Steven N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics*, 2nd ed., Palgrave Macmillan, 2008.