A bit of GMM

FEI GAO

June 15, 2020

Layout of today

I want to share and examine what I did learn in last few days about

- What is GMM.
- Some simple examples (meaningless examples but good to know how to run GMM)
- Some practices examples

There are tons of good materials on Internet. Use wisely. The purpose of me today is fix the last mile gap

Some Good Reference

- Canova, F., (2008), Chapter 5: GMM and Simulation Estimators, Methods for Applied Macroeconomic Research, Princeton University Press.
- An open lecture slides from F.Canova available on `http://apps.eui.eu/Personal/Canova/Teachingmaterial/Gmm_eui2014.pdf`
- Fernández-Villaverde, J., Rubio-Ramírez, J.F., Schorfheide, F., (2016), Chapter 9 - Solution and Estimation Methods for DSGE Models, Handbook of Macroeconomics volume 2, Pages 527-724, Elsevier.
- Ruge- Murcia, F.J., (2013), Chapter 20: Generalized Method of Moments estimation of DSGE models*, Handbook of Research Methods and Applications in Empirical Macroeconomics, Pages 464-485, Edward Elgar Publishing.
- Hayashi, F. (2002), Econometrics, Princeton University Press.
- Kladivko, K., (2007), "The General Method of Moments (GMM) using MATLAB: The practical guide based on the CKLS interest rate model." Department of Statistics and

Why need to learn GMM?

In chapter 8 of Open Economy Macroeconomics, Uribe uses GMM method to estimate 4 parameters ϕ_m, ϕ_x, ϕ_n and ψ. GMM was popular among macroeconomist and is still a good tool to estimate parameter by partial equations from model without solve the DSGE model.

Basically, the advantages are

- Model doesn't need to be solved before estimation.
- Works well also on nonliear equations.
- Works well also on weak identification inference.

Disadvantages?

- Efficiency loss relative to maximum likelihood method.
- Frequentist perspective.

What is GMM?

Full name of GMM is Generalized Method of Moments. GMM is a general and powerful econometrics tool. It can be concluded by one equation

$$m(\theta_0) \equiv \mathbb{E}[g(Y_t, \theta_0)] = 0 \tag{1}$$

where $\mathbb{E}$ denotes expectation, Y_t is observation, θ_0 is true value of parameters of interest. $g(\cdot)$ is a generic function which differ from 0 if $\theta \neq \theta_0$.

Assume we have number T observations. Sample average implies the expectation, so we can write

$$\hat{m}(\theta) \equiv \frac{1}{T} \sum_{t=1}^{T} g(Y_t, \theta) \tag{2}$$

The GMM estimator is

$$\arg\min_{\theta \in \Theta} \hat{m}(\theta)^\intercal W \hat{m}(\theta) \tag{3}$$

where W is a weight matrix (we will see it later).

Example 1

$$\max_{\{C_t, a_{t+1}\}_{t=0}^{\infty}} \quad E_0 \sum_{t=0}^{\infty} \beta^t U(c_t) \tag{4}$$
$$c_t + a_{t+1} \leq w_t + (1+r)a_t$$

where w_t is labor income, a_{t+1} is saving which is predetermined at period t, and $r_t = r$ for all t. $U'_{c_t} \equiv \frac{\partial U(c_t)}{\partial c_t}$, λ_t is the constant lagrangian multiplier. FOCs are

$$\begin{aligned} U'_{c_t} &= \lambda_t \\ \lambda_t &= \beta \mathcal{E}_t \lambda_{t+1}(1+r) \end{aligned} \tag{5}$$

The Euler equation can be established by combine two equations above.

$$E_t[\beta(1+r)\frac{U'_{c_{t+1}}}{U'_{c_t}} - 1] = 0 \tag{6}$$

Example 1 (continue)

Assume we know $r = 0.04$ for all period, and the utility function is log from, i.e. $U(c_t) = \log(c_t)$.
The parameter of interst is β, we believe there is a true β denote as β_0 and we want to estimate β using GMM with T observations data.

$$m(\beta_0) = \mathbb{E}[\beta \frac{c_t}{c_{t+1}}(1+r) - 1] = 0 \tag{7}$$

$$\hat{m}(\beta) = \frac{1}{T-1-1} \sum_{t=1}^{T-1} [\beta \frac{c_t}{c_{t+1}}(1+r) - 1] = 0 \tag{8}$$

$$W = 1 \tag{9}$$

We minimize $\hat{m}(\beta)^{\intercal} W \hat{m}(\beta)$ by β. Notice that there is one equation and one parameter, i.e. just identified.

Example 2 (continue)& Example 2

We denote number of equations as k and number of unknown parameters as n. We have 3 cases

- $k < n$ Under identified. Unable to estimate. Need to increase equation.
- $k = n$ Just identified. Good to estimate.
- $k > n$ Over identification. To estimate drop some equations or use weight matrix.

To show over-identification case, see following example 2.
If we also know the production function of this economy such as $Y_t = A_t K_t^{\alpha} L_t^{1-\alpha}$. We have one more Euler equation (household can invest on capital instead of saving) such as

$$E_t[\beta \frac{U'_{c_{t+1}}}{U'_{c_t}}(1 + \alpha A_t K_t^{\alpha-1} L_t^{1-\alpha}) - 1] = 0 \tag{10}$$

Example 2(continue)

Hence we have other moment condition. Denote condition (7) in previous slides $m_1(\beta_0)$ with

$$m_2(\beta_0) = E_t[\beta \frac{U'_{c_{t+1}}}{U'_{c_t}}(1 + \alpha A_t K_t^{\alpha-1} L_t^{1-\alpha}) - 1] = 0 \tag{11}$$

together with same assumptions

$$\hat{m}(\beta) = \begin{bmatrix} \hat{m}_1(\beta) \\ \hat{m}_2(\beta) \end{bmatrix} = \frac{1}{T-1-1} \begin{bmatrix} \sum_{t=1}^{T-1}[\beta \frac{c_t}{c_{t+1}}(1+r) - 1] \\ \sum_{t=1}^{T-1}[\beta \frac{c_t}{c_{t+1}}(1 + \alpha A_t K_t^{\alpha-1} L_t^{1-\alpha}) - 1] \end{bmatrix} = 0 \tag{12}$$

In this case, we set W as

$$W = \begin{bmatrix} w_1 & 0 \\ 0 & w_2 \end{bmatrix} \tag{13}$$

Notice that if $w_2 = 0$, this case is as same as Example 1. We estimate with more information than case 1.

A further discussion on weight matrix

Intuition behind weight matrix.
If we have an unknown parameter x and two conditions.

$$m(x) = \begin{bmatrix} 5x - 1 \\ 3x - 2 \end{bmatrix} = 0 \tag{14}$$

There is no solution that solve $m(x)$ obviously. But we believe both equation should hold together somehow. The weight matrix allow us to transform the object function into

$$w_1(5x_1)^2 + w_2(3x - 2)^2 = 0 \tag{15}$$

By choosing right w_1 and w_2 we can solve x for the system.

A further discussion on weight matrix (continue)

How to choose W?

We estimate $\hat{\beta}$ and W in two stages

- First stage: Choose identity matrix I as W to estimate $\hat{\beta}$.
- Second stage: Estimate S and set $W = S^{-1}$.

In practice, we usually iterate two stages several times in estimation.

What is S? We use Newey West estimator to compute S to overcome autocorrelation.

$$S = \sum_{j=-\infty}^{\infty} \mathbb{E}[g(Y_t, \theta)g(Y_{t-j})(\theta)^\intercal]$$

$$\hat{S} = \hat{S}_0 + \sum_{j=1}^{k}(1 - \frac{j}{k+1})(\hat{S}_j + \hat{S}_j^\intercal) \tag{16}$$

$$\text{where } \hat{S}_j = \frac{1}{T}\sum_{t=j+1}^{T}[g(Y_t, \theta)g(Y_{t-j})(\theta)^\intercal]$$

Test

Need to compute variance of estimate parameter to conduct test.
GMM has the following property

$$\sqrt{T}(\hat{\theta} - \theta_0) \to^d N(0, (DS^{-1}D^\intercal)^{-1}) \tag{17}$$

where $D = \mathbb{E}\frac{\partial m(\theta)}{\partial \theta}$. We can compute $(DS^{-1}D^\intercal)^{-1}$ to have variance.

One more test...

Since we have 2 equations but 1 unknown parameters, we can conduct J test (also named Sargan-Hansen test). The test allow us to check whether the model's conditions match the data well or not.

- H_0: $m(\theta_0) = 0$
- H_1: $m(\theta) \neq 0$ for all θ.

$$J \equiv T * \hat{m}(\hat{\theta})^\intercal \hat{W} \hat{m}(\hat{\theta}) \to^d \chi^2_{k-n} \tag{18}$$

as before number of equations, k, and number of unknown parameters, n.

Talk is cheap, let me show some codes.

Examples in Practices

An example comes from practice. This is my favorite among all. It can be easily found in different materials. I refer Canova(2008) here. NKPC is known as

$$\pi_t = \mathbb{E}\beta\pi_{t+1} + \frac{(1-\xi)(1-\xi\beta)}{\xi}mc_t \tag{19}$$

where β is discount factor, π is inflation, ξ is the probability of Calvo pricing, $mc = \frac{wN}{GDP}$ is real marginal cost.

$$m(\xi_0, \beta_0) = \mathbb{E}(\pi_t - \beta\pi_{t+1} - \frac{(1-\xi)(1-\xi\beta)}{\xi}mc_t) = 0 \tag{20}$$

$$m(a_0, b_0) = \mathbb{E}(\pi_t - a\pi_{t+1} - bmc_t) = 0 \tag{21}$$

Uribe Chapter 8

Finally, we come to Uribe Chapter 8.
In MXN model, Uribe targets 3 moments such as stand deviation ratio between investment and real gdp, σ_i/σ_y, stand deviation between trade balance ratio and real gdp, σ_{tby}/σ_y,stand deviation ratio between investment in tradeable sector and nontradeable sector, i.e., $\sigma(ix+im)/\sigma_{in}$.
The unknown parameters are $\phi_{m,x,n}$ which are capital adjustment cost of investments in three sector, ψ the parameter in evolution of interest rate.
We let

$$x(\theta) = \mathbb{E}(g(\theta) - C) = 0$$

$$\text{where } \theta = \begin{bmatrix} \phi_m \\ \phi_x \\ \phi_n \\ \psi \end{bmatrix} \quad g(\theta) \text{ is a } 3x1 \text{ vector} \quad C = \begin{bmatrix} \frac{\sigma_i}{\sigma_y} \\ \frac{\sigma_{tby}}{\sigma_y} \\ \frac{\sigma(ix+im)}{\sigma_{in}} \end{bmatrix} \tag{22}$$

It seems like under identified because $k = 3 < n = 4$. But we can target four moment conditions instead of three (let $g(\theta)$ be a $4x1$ vector) and $C = \begin{bmatrix} \sigma_i \\ \sigma_{tby} \\ \sigma_y \\ \frac{\sigma(ix+im)}{\sigma_{in}} \end{bmatrix}$.

What will be $g(\theta)$?

Limited by my knowledge, I'd like to consider DSGE model itself is the function $g(\cdot)$.

In this way, no analytical solution, solve DSGE by RISE/ Uribe code, search Θ where θ lies to find a minimizer which minimizes the distance.

It sounds easy but I have problems....

Explain to you with some matlab codes in my hand.