The Generalized Method of Moments

Emmanuel A. Cabral

Ateneo de Manila University

cabral@math.admu.edu.ph

05 August 2011

Motivation

Why GMM?

Ordinary least squares and generalized least squares are well-known methods of parameter estimation in financial and econometrics literature. The generalized method of moments (GMM) is a parameter estimation method that may not be as well-known but is currently becoming popular among researchers in financial economics and econometrics. This method may prove useful for practitioners especially in interest rate modeling and derivative pricing where parameters such as volatility are constantly being updated.

The Method of Moments (MOM)

Suppose we have data $Y_1, Y_2, \ldots, Y_T$ drawn from a student t-distribution with ν degrees of freedom and we want to estimate ν. It is well-known that for $\nu > 2$,

$$\begin{aligned} \mu_1 &= E(Y_t) = 0; \\ \mu_2 &= E(Y_t^2) = \frac{\nu}{\nu - 2}. \end{aligned}$$

Equating the sample second moment

$$\widehat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} Y_t^2$$

and the population second moment μ_2, we have

$$\widehat{\nu}^{(1)} = \frac{2\widehat{\sigma}^2}{\widehat{\sigma}^2 - 1}$$

MOM Continued

It is also known that the population 4th moment satisfies

$$\mu_4 = E(Y_t^4) = \frac{3\nu^2}{(\nu - 2)(\nu - 4)}$$

for $\nu > 4$. By equating sample and population 4th moments, we have

$$\widehat{\nu}^{(2)} = \frac{3\widehat{\mu}_4 \pm \sqrt{\widehat{\mu}_4(\widehat{\mu}_4 + 24)}}{3\widehat{\mu}_4 - 3}$$

Thus $\widehat{\nu}^{(1)}$ and $\widehat{\nu}^{(2)}$ are two different estimators of the degrees of freedom ν based on the sample statistics $\widehat{\sigma}^2$ and $\widehat{\mu}_4$, respectively. The two estimators are different with probability 1.

An Over-identified Model

There is no single $\widehat{\nu}$ that will satisfy the system

$$\begin{aligned} \widehat{\nu}^{(1)} &= \frac{2\widehat{\sigma}^2}{\widehat{\sigma}^2 - 1} \\ \widehat{\nu}^{(2)} &= \frac{3\widehat{\mu}_4 \pm \sqrt{\widehat{\mu}_4(\widehat{\mu}_4 + 24)}}{3\widehat{\mu}_4 - 3} \end{aligned}$$

The solution is to generalize the MOM approach by introducing a 2×2 matrix W that reflects our confidence in each moment condition above. Here, W is symmetric and positive-definite. This is done by writing

$$g(\nu) := \begin{bmatrix} \widehat{\sigma}^2 - \frac{\nu}{\nu-2} \\ \widehat{\mu}_4 - \frac{3\nu^2}{(\nu-2)(\nu-4)} \end{bmatrix}.$$

We then find ν that minimizes the objective function

$$Q(\nu) = g(\nu)' W g(\nu).$$

The GMM Formalization

In general, a $p \times 1$ parameter vector β_0 has to be estimated. Define the function

$$f(X_t, \beta_0) = \begin{bmatrix} f_1(X_t, \beta_0) \\ f_2(X_t, \beta_0) \\ \vdots \\ f_q(X_t, \beta_0) \end{bmatrix}$$

that satisfies q population moment conditions given by

$$E[f(X_t, \beta_0)] = 0.$$

Let

$$g_T(\beta) = \frac{1}{T} \sum_{t=1}^{T} f(X_t, \beta).$$

GMM Formalization

The idea behind GMM is to choose β so as to make $g(\beta)$ as close as possible to the population moment of zero, that is, β_{GMM} is the value of β that minimizes the scalar

$$Q_T(\beta) = g(\beta)' W_T g(\beta)$$

where $\{W_T\}_{T=1}^{\infty}$ is a sequence positive definite $q \times q$ matrices.
If $p = q$, then the model is exactly identified and $\widehat{\beta}$ can be found so that $g(\widehat{\beta}) = 0$. If $q > p$, how close the ith element of $g(\widehat{\beta})$ is to 0 depends on how much weight the ith orthogonality condition is given by the weighting matrix W_T.

Stationarity and Ergodicity of Time Series Data

GMM assumes that the underlying data X_t are stationary and ergodic.

1. Stationarity of X_t means
   1. $E(X_t) = \mu \;\; \forall t$;
   2. $E[(X_t - \mu)(X_{t-j} - \mu)'] = \gamma_j \;\; \forall t \;\; \forall j$.

   Stationarity is stronger than *identically distributed* but weaker than *i.i.d.*, since stationarity does not imply independence.
2. The data X_t are ergodic for the mean if

$$\frac{1}{T}\sum_{t=1}^{T} X_t \to E(X_t) \;\text{ as }\; T \to \infty.$$

   Stationarity and ergodicity together are strictly weaker than *i.i.d*.

Testing Stationarity and Ergodicity Empirically

The Optimal Weighting Matrix

Under the assumptions of stationarity and ergodicity, it can be shown that the optimal weighting matrix is given by

$$W = S^{-1}$$

where

$$S = \lim_{T \to \infty} T \cdot E[g(\beta_0) g(\beta_0)']$$

and β_0 is the only value of β that satisfies the population moment condition

$$E[f(X_t, \beta_0)] = 0.$$

Estimating W and β_0

In practice, we are estimating both W and β_0. But for very large T, we can use

$$W_T = S_T^{-1}$$

where

$$S_T = E[\sqrt{T}g(\beta_{GMM})\sqrt{T}g(\beta_{GMM})']$$

and β_{GMM} is the choice of β that minimizes

$$Q_T(\beta) = g(\beta)' W_T g(\beta))$$

The algorithm goes as follows:

An Algorithm to Estimate W and β

Step 1: Set $W = I$ and find $\beta^{(1)}$ so that $\beta^{(1)}$ minimizes

$$g(\beta)' W g(\beta).$$

Step 2: Use in $\beta^{(1)}$ to get

$$\begin{aligned} S_T^{(1)} &= E\left[\sqrt{T} g(\beta^{(1)}) \sqrt{T} g(\beta^{(1)})'\right] \\ &= E\left[\frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} f(X_t, \beta^{(1)}) f(X_t, \beta^{(1)})'\right] \\ W_T^{(1)} &= \left[S_T^{(1)}\right]^{-1}. \end{aligned}$$

Step 3: Get $\beta^{(2)}$ using $W_T^{(1)}$, and so on until

$$\beta^{(k+1)} \approx \beta^{(k)}.$$

Asymptotic Distribution of GMM Estimates

Proposition

The GMM estimator β_{GMM} is asymptotically normal, with

$$\sqrt{T}\,(\beta_{GMM} - \beta_0)) \sim N(0, V_{GMM})$$

where

$$\begin{aligned}
V_{GMM} &= \left(\Gamma' S_T^{-1} \Gamma\right)^{-1} \\
\Gamma &= E\left(\frac{\partial g(\beta)}{\partial \beta'}\right) = E\left(\frac{1}{T}\sum_{t=1}^{T}\frac{\partial f(X_t,\beta)}{\partial \beta'}\right) \\
S_T &= E\left[\sqrt{T} g(\beta)\sqrt{T} g(\beta)'\right] \\
&= E\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T} f(X_t,\beta) f(X_t,\beta)'\right]
\end{aligned}$$

References

Arnold T. and T. Falcon Crack (1999)
A Practical Guide to GMM (with Applications to Option Pricing)

Hamilton, J.D. (1994)
Time Series Analysis

Hansen, Lars Peter (1982)
Large Sample Properties of Generalized Method of Moments Estimators
Econometrica 50(4).

The End