Generalized Method of Moments with application to the SV model

Leopoldo Catania
Aarhus University and CREATES

leopoldo.catania@econ.au.dk

Introduction

- GMM has become one of the main statistical tools for the analysis of economic and financial data.
- GMM was first introduced by L.P. Hansen in 1982. Since then it has been widely applied to analyze economic and financial data.
- GMM has been applied to time series, cross sectional, and panel data.
- Optimality of MLE stems from its basis on the joint p.d. of the data. However, in some circumstances, this dependence becomes a weakness.

MLE

Problems of MLE:

- Sensitivity of statistical properties to the distributional assumptions
- The likelihood function is not always available...
- ... for instance if we have latent (unobserved) variables in the model (SV models).
- Computational burden.

The Method of Moments

- Population moments implied by the family of distributions are functions of the unknown parameter vector.
- Pearson (1895) proposed estimating the parameter vector by the value implied by the corresponding sample moments.
- Normal distribution with parameters μ_0 and σ_0^2

$$\begin{aligned} E[x_t] - \mu_0 &= 0 \\ E[x_t^2] - (\sigma_0^2 + \mu_0^2) &= 0 \end{aligned}$$

- Pearson's method involves replacing the population moments by the sample moments:

$$\begin{aligned} E_N[x_t] - \hat{\mu} &= 0 \\ E_N[x_t^2] - (\hat{\sigma}^2 + \hat{\mu}_0^2) &= 0 \end{aligned}$$

The Method of Moments: Orthogonality Condition and Regression

In linear regression model we assume the following condition:

$$E(x_t \epsilon_t) = 0 \quad t = 1, ..., T$$

Let β_0 denote the true value of β, where the latter denotes a generic value of the parameter. Then

$$E[x_t(y_t - x_t\beta_0)] = 0$$

The Method of Moments: Orthogonality Condition and Regression

The moment condition can be written as

$$g_t(\beta) = x_t(y_t - x_t\beta)$$

such that

$$E[g_t(\beta_0)] = 0$$

The empirical analogue of $E[g_t(\beta_0)]$ is

$$\bar{g}_t = \frac{1}{T}\sum_{t=1}^{T} x_t(y_t - x_t\beta) = \frac{1}{T}(X'y - X'X\beta).$$

Which implies

$$(X'y - X'X\hat{\beta}) = 0,$$

which are called *normal equations*.

The Method of Moments: the SV model

Consider the SV model

$$
\begin{aligned}
y_t &= \exp\left(h_t/2\right) u_t \\
h_{t+1} - \alpha &= \phi(h_t - \alpha) + \eta_t
\end{aligned}
$$

where

$$
\begin{pmatrix} u_t \\ \eta_t \end{pmatrix} \overset{iid}{\sim} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & \sigma_\eta^2 \end{pmatrix} \right).
$$

We want to estimate $\theta = (\alpha, \phi, \sigma_\eta^2)$, or the reparameterization $\theta = (\alpha, \phi, \beta^2)$ where $\beta^2 = \sigma_\eta^2/(1-\phi^2)$.

The Method of Moments: the SV model

Consider the population parameters α, ϕ, β^2. Three population moments are (for example):

$$E[|y_t|] = \sqrt{2/\pi} \exp\left(\frac{\alpha}{2} + \frac{1}{8}\beta^2\right)$$

$$E[y_t^2] = \exp\left(\alpha + \frac{1}{2}\beta^2\right)$$

$$E[y_t^4] = 3\exp\left(2\alpha + 2\beta^2\right)$$

The Method of Moments: the SV model

One possible solution for α and β^2 is obtained using the empirical counterparts of $E(y_t^2)$ and $E(y_t^4)$, that we indicate by $\hat{\mu}_2 = \frac{1}{T}\sum_{t=1}^T y_t^2$ and $\hat{\mu}_4 = \frac{1}{T}\sum_{t=1}^T y_t^4$, respectively. The MM estimator of (α, β^2), indicated by $(\hat{\alpha}, \hat{\beta})$, is given by:

$$\hat{\alpha} = \log\left(\hat{\mu}_2^2\sqrt{\frac{3}{\hat{\mu}_4}}\right), \qquad \hat{\beta}^2 = \log\left(\frac{\hat{\mu}_4}{3\hat{\mu}_2^2}\right)$$

Note that $\hat{\mu}_4$ is very sensitive to extreme observations!

The Method of Moments: the SV model

One second solution for α and β^2 is obtained using $E[|y_t|]$ and $E[y_t^2]$:

$$\hat{\alpha} = \log\left(\frac{\pi^2\hat{\kappa}^4}{4\mu_2}\right), \qquad \hat{\beta}^2 = \log\left(\frac{16\hat{\mu}_2^4}{\pi^4\hat{\kappa}^8}\right),$$

where $\hat{\kappa} = \frac{1}{T}\sum_{t=1}^{T}|y_t|$ is the empirical counterpart of $E[|y_t|]$.

- No simple method of moments estimator for ϕ
- Method of moments estimators are not unique!

The Generalized Method of Moments (GMM)

The idea of GMM is to optimally combine moment conditions to estimate population parameters.

Let $\{w_t\}$ be a covariance stationary and ergodic vector process representing the underlying data. Let the $p \times 1$ vector θ denote the population parameters. The moment conditions $g(w_t, \theta)$ are $K \geq p$ possibly nonlinear functions satisfying:

$$E[g(w_t, \theta_0)] = 0,$$

where θ_0 represents the true parameter vector.

The Generalized Method of Moments: identification

Global identification of θ_0 requires that:

$$
\begin{aligned}
E[g(w_t, \theta_0)] &= 0 \\
E[g(w_t, \theta)] &\neq 0 \quad \text{for} \quad \theta \neq \theta_0,
\end{aligned}
$$

Local identification requires that the $K \times p$ matrix

$$
G = E\left[\frac{\partial g(w_t, \theta_0)}{\partial \theta'}\right],
$$

has full column rank p.

The Generalized Method of Moments: estimation

The sample moment conditions for an arbitrary θ is:

$$g_T(\theta) = T^{-1} \sum_{t=1}^{T} g(w_t, \theta).$$

If $K = p$, then θ_0 is apparently just identified and the GMM objective function is:

$$J(\theta) = T g_T(\theta)' g_T(\theta),$$

which does not depend on a weight matrix. The corresponding GMM estimator is then:

$$\hat{\theta} = \arg\min_{\theta} J(\theta)$$

and satisfies $g_T(\hat{\theta}) = 0$.

The Generalized Method of Moments: estimation

If $K > p$, then θ_0 is apparently overidentified. We thus denote with $\hat{W}$ a $K \times K$ symmetric and positive definite weight matrix, possibly dependent on the data, such that $\hat{W} \rightarrow W$ as $T \rightarrow \infty$ with W symmetric and positive definite.

The GMM estimator of θ_0, denoted $\hat{\theta}(\hat{W})$, is defined as:

$$\hat{\theta} = \arg\min_{\theta} J(\theta, \hat{W}) = T g_T(\hat{\theta})' \hat{W} g_T(\hat{\theta}),$$

whose first order conditions are:

$$\begin{aligned}
\frac{\partial J(\hat{\theta}(\hat{W}), \hat{W})}{\partial \theta} &= 2 G_T(\hat{\theta}(\hat{W}))' \hat{W} g_T(\hat{\theta}(\hat{W})) = 0 \\
G_T(\hat{\theta}(\hat{W})) &= \frac{\partial g_T(\hat{\theta}(\hat{W}))}{\partial \theta'}
\end{aligned}$$

The Generalized Method of Moments: Asymptotic Properties

Under standard regularity conditions, it can be shown that:

$$\hat{\theta}(\hat{W}) \xrightarrow{p} \theta_0$$

$$\sqrt{T}(\hat{\theta}(\hat{W}) - \theta_0) \xrightarrow{d} N(0, avar(\hat{\theta}(\hat{W}))),$$

where

$$avar(\hat{\theta}(\hat{W}))) = (G'WG)^{-1}G'WSWG(G'WG)^{-1},$$

and

$$G = E\left[\frac{\partial g(w_t, \theta_0)}{\partial \theta'}\right]$$

$$S = avar(\sqrt{T}g_T(\theta_0))$$

The Generalized Method of Moments: About W

The efficient GMM estimator uses a weight matrix W that minimizes $avar(\hat{\theta}(\hat{W}))$. Hansen (1982) showed that the optimal weight matrix is $W = S^{-1}$, that is:

$$avar(\hat{\theta}(\hat{W})) = (G'S^{-1}G)^{-1},$$

if $\{g_t(w_t, \theta_0)\}$ is an ergodic stationarity martingale difference sequence then:

$$S = E[g_t(w_t, \theta_0)g_t(w_t, \theta_0)'],$$

and a consistent estimator of S takes the form:

$$\hat{S}_{HC} = T^{-1}\sum_{t=1}^{T} g_t(w_t, \hat{\theta})g_t(w_t, \hat{\theta})',$$

The Generalized Method of Moments: About W

If $\{g_t(w_t, \theta_0)\}$ is a mean–zero serially correlated ergodic stationary process then

$$S = \Gamma_0 + \sum_{j=1}^{\infty}(\Gamma_j + \Gamma_j'),$$
$$\Gamma_j = E[g_t(w_t, \theta_0)g_t(w_{t-j}, \theta_0)]$$

and a consistent estimator has the form:

$$S_{HAC} = \hat{\Gamma}_0(\hat{\theta}) + \sum_{j=1}^{q(T)} k\left(\frac{j}{q(T)+1}\right)(\hat{\Gamma}_j(\hat{\theta}) + \hat{\Gamma}_j(\hat{\theta})')$$
$$\hat{\Gamma}_j(\hat{\theta}) = \frac{1}{T-j}\sum_{t=j+1}^{T} g_t(w_t, \hat{\theta})g_t(w_{t-j}, \hat{\theta})',$$

for a proper kernel function $k(\cdot)$. The usual choice is the triangular kernel $k(x) = 1 - |x|$. We usually set $q(T) = \lfloor T^{1/3} \rfloor$

The Generalized Method of Moments: SV

Consider the alternative parameterization of the simple log-normal stochastic volatility (SV) model assuming:

$$\begin{aligned} y_t &= \exp(h_t/2)u_t \\ h_{t+1} &= \omega + \phi h_t + \eta_t, \end{aligned}$$

where

$$\begin{pmatrix} u_t \\ \eta_t \end{pmatrix} \overset{iid}{\sim} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & \sigma_\eta^2 \end{pmatrix}\right)$$

and in this case $\theta = (\omega, \phi, \sigma_\eta^2)$. Note that according to this model y_t is stationary and unconditional moments of all orders exists.

The Generalized Method of Moments: SV

The GMM estimation of the SV model is surveyed in Andersen and Sørensen (1996).

They recommended using moment conditions for GMM estimation based on lower-order moments of y_t, since higher-order moments tend to exhibit erratic finite sample behavior.

They considered a GMM estimation based on (subsets) of 24 moments considered by Jacquier et al. (1994). To describe these moment conditions, first define:

$$\alpha = \frac{\omega}{1-\phi}, \qquad \beta^2 = \frac{\sigma_\eta^2}{1-\phi^2},$$

and $\theta_h = (\alpha, \phi, \beta^2)$, which is just a reparameterization of $\theta = (\omega, \alpha, \sigma_\eta^2)$.

The Generalized Method of Moments: SV

The moment conditions, which follow from properties of the log-normal distribution and the Gaussian AR(1) model, are expressed as:

$$
\begin{aligned}
E[|y_t|] &= (2/\pi)^{1/2} E[\sigma_t] \\
E[y_t^2] &= E[\sigma_t^2] \\
E[|y_t^3|] &= 2\sqrt{2/\pi} E[\sigma_t^3] \\
E[y_t^4] &= 3E[\sigma_t^4] \\
E[|y_t y_{t-j}|] &= (2/\pi) E[\sigma_t \sigma_{t-j}], \quad j = 1, \dots, 10 \\
E[y_t^2 y_{t-j}^2] &= E[\sigma_t^2 \sigma_{t-j}^2], \quad j = 1, \dots, 10
\end{aligned}
$$

where for any positive integer j and positive constants p and s,

$$
E[\sigma_t^p] = \exp\left(\frac{p\alpha}{2} + \frac{p^2\beta^2}{8}\right)
$$

$$
E[\sigma_t^p \sigma_{t-j}^s] = E[\sigma_t^p] E[\sigma_t^s] \exp\left(\frac{ps\phi^j \beta^2}{4}\right)
$$

The Generalized Method of Moments: SV

We set $w_t = (|y_t|, y_t^2, |y_t^3|, y_t^4, |y_t y_{t-1}|, \dots, |y_t y_{t-10}|, y_t^2 y_{t-1}^2, \dots, y_t^2 y_{t-10}^2)'$, and define the 24×1 vector

$$g(w_t, \theta_h) = \begin{pmatrix} |y_t| - (2/\pi)^{1/2} \exp\left(\frac{\alpha}{2} + \frac{\beta^2}{8}\right) \\ y_t^2 - \exp(\alpha + \frac{\beta^2}{2}) \\ \vdots \\ y_t^2 y_{t-10}^2 - \exp\left(\alpha + \frac{\beta^2}{2}\right)^2 \exp(\phi^{10}\beta^2) \end{pmatrix}$$

Then, $E[g(w_t, \theta_{h0})] = 0$ is the population moment condition used for the GMM estimation of the model parameters $\theta_h = (\alpha, \phi, \beta^2)$.

Since the elements of w_t are serially correlated, the efficient weight matrix $S = avar(\sqrt{T} g_T(\theta_{h0}))$, must be estimated using an HAC estimator.

References I

Andersen, T. G. and Sørensen, B. E. (1996). Gmm estimation of a stochastic volatility model: a monte carlo study. *Journal of Business & Economic Statistics*, 14(3):328–352.

Jacquier, E., Polson, N. G., and Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics*, 20(1):69–87.