Generalized Method of Moments

Lecture 3

Alastair R. Hall
North Carolina State University

June 26, 2002

Lecture outline:

1. Introduction

2. GMM estimation in correctly specified nonlinear dynamic models - part I.

3. **GMM estimation in correctly specified nonlinear dynamic models - part II.**

4. Hypothesis testing.

5. Finite sample behaviour and moment selection.

Today:

- Two step and iterated GMM estimation

- Impact of transformations

- Continuous updating GMM estimator

1. Two - step and iterated GMM estimation

Recall that: $T^{1/2}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(0, MSM')$
where

- $M = (G_0' W G_0)^{-1} G_0' W$.

- $G_0 = E[\partial f(v_t, \theta_0)/\partial \theta']$

- $S = \lim_{T\to\infty} Var[T^{1/2} g_T(\theta_0)]$

If $p = q$ then $MSM' = (G_0' S^{-1} G_0)^{-1}$
– independent of W.

If $p < q$ then asymptotic variance depends on W.

Optimal choice of W? Choice that minimizes variance.

Optimal weighting matrix: $W = S^{-1}$

$\rightarrow MSM' = (G_0' S^{-1} G_0)^{-1}$

How can we obtain an estimator that has this variance?

Two–step procedure:

1. Estimate with sub–optimal $W_T \to \widehat{\theta}_T(1) \to \widehat{S}_T(1)$.

2. Estimate with $W_T = \widehat{S}_T(1)^{-1} \to \widehat{\theta}_T(1)$.

Can iterate further

1. $\widehat{\theta}_T(i-1) \to \widehat{S}_T(i-1)$

2. Estimate with $W_T = \widehat{S}_T(i-1)^{-1} \to \widehat{\theta}_T(i)$.

Continue until $\|\widehat{\theta}_T(i-1) - \widehat{\theta}_T(i)\| < \epsilon$ or $i = i_{max}$.

Numerical illustration: Hansen & Singleton (1982) CBAPM

$$\text{pmc} : \; E[e_t(\gamma, \delta)\, z_t] = 0$$

where

$$e_t(\gamma, \delta) \;=\; \delta(r_{t+1}/p_t)(c_{t+1}/c_t)^{\gamma-1} - 1$$

and

- Single asset = equally weighted NYSE index (EWR) or value weighted NYSE index (VWR)
- c_t = aggregate per capita consumption
- $z_t = (1, c_t/c_{t-1}, c_{t-1}/c_{t-2}, r_t/p_{t-1}, r_{t-1}/p_{t-2})'$
- Sample: 1960.1–1991.12

As before for first step use $W_T = 10^5 I_5$ and $(T^{-1}Z'Z)^{-1}$.

table 3.7

table 3.8

2. Impact of transformations

Consider five types of transformation:

- Units of measurement for v_t
- Reparameterization
- Normalization of the parameter vector
- Curvature altering transformations of the population moment condition
- Stationarity inducing transformations of $f(.)$

(i) Units of measurement for v_t:
In general, the GMM estimator is not invariant to changes in the units of measurement of v_t.

Example: $E[v_t] - \theta_0 = 0 \rightarrow \hat{\theta}_T = T^{-1} \sum_{t=1}^{T} v_t$.

Data are $x_t = cv_t$

$E[x_t] - \theta_0 = 0 \rightarrow \tilde{\theta}_T = T^{-1} \sum_{t=1}^{T} x_t$

So $\tilde{\theta}_T = c\hat{\theta}_T$

But interpretation of θ_0 has changed!

(ii) Reparameterization:

θ_0 satisfies:

- globally identified
- can be written as $\theta_0 = h(\gamma_0)$ where $h : \Re^p \rightarrow \Re^p$ is a continuous, differentiable bijective mapping.

pmc: $E[f(v_t, h(\gamma_0))] = E[f_\gamma(v_t, \gamma_0)] = 0$

What is relationship between:

- estimate γ_0 based on $E[f_\gamma(v_t, \gamma_0)] = 0$
- estimate of θ_0 based on $E[f(v_t, \theta_0)] = 0$?

The GMM estimator is invariant to reparameterization in the sense that the two parameterizations yield logically consistent estimators.

$Q_{\gamma,T}(\gamma) =$ GMM minimand associated with the reparameterized model, *i.e.* $Q_{\gamma,T}(\gamma) = Q_T(h(\gamma))$.

$\widehat{\gamma}_T = argmin\, Q_{\gamma,T}(\gamma)$

Can calculate $\widehat{\gamma}_T$ as follows.

- min $Q_T(h(\gamma))$ wrt $h(\gamma) \to \widehat{h}_T$

- $\widehat{h}_T = h(\widehat{\gamma}_T) \to \widehat{\gamma}_T$.

But $\widehat{h}_T = \widehat{\theta}_T$ and so by construction

$$\widehat{\theta}_T \;=\; h(\widehat{\gamma}_T)$$

Similar result does not extend to the estimated asymptotic standard errors.

It can be shown that

$$\widehat{V}_{\gamma,T} = [H(\widehat{\gamma}_T)]^{-1}\widehat{V}_{\theta,T}[H(\widehat{\gamma}_T)']^{-1}$$

where $H(.) = \partial h(.)/\partial\gamma'$.

So inferences may be sensitive.

This sensitivity is a potential source of concern, and motivates an alternative method for the construction of confidence intervals covered later.

"natural parameterization" argument

(iii) The GMM estimator and normalization of the parameter vector

In some cases, θ_0 may only be identified up to some scaling factor and so it is necessary to impose some normalization on θ_0, such as $\theta_{0,1} = 1$, in order to achieve identification.

In general, the GMM estimators associated with different normalizations of the parameter vector do not exhibit a logical consistency in finite samples. However, they do exhibit a logical consistency in the limit.

Example: Suppose that

$$e_t(\theta) = \theta_1 R_{1,t} + \theta_2 R_{2,t} + \theta_3 I_t + \theta_4 S_t$$

where

$$E[z_t e_t(\theta_0)] = 0$$

θ_0 is unidentified

Normalize to achieve identification:

- Divide $e_t(\theta_0)$ by $\theta_{0,1}$:

  $\tilde{e}_t(\psi_0) \;=\; R_{1,t} + \psi_{0,1} R_{2,t} + \psi_{0,2} I_t + \psi_{0,3} S_t$

  where $\psi_{0,i} = \theta_{0,i+1}/\theta_{0,1}$.

- Divide $e_t(\theta_0)$ by $\theta_{0,4}$:

  $\bar{e}_t(\phi_0) \;=\; \phi_{0,1} R_{1,t} + \phi_{0,2} R_{2,t} + \phi_{0,3} I_t + S_t$

  where $\phi_{0,i} = \theta_{0,i}/\theta_{0,4}$.

Both normalizations are logically consistent.

These normalizations lead to pmc's:

$$
\begin{aligned}
E[z_t \tilde{e}_t(\psi_0)] &= 0 \\
E[z_t \bar{e}_t(\phi_0)] &= 0
\end{aligned}
$$

Corresponding estimators are:

$$
\begin{aligned}
\widehat{\psi}_T &= \left[(T^{-1} \sum_{t=1}^{T} x_{1,t} z_t^{'}) W_T (T^{-1} \sum_{t=1}^{T} z_t x_{1,t}^{'}) \right]^{-1} \\
& \qquad \times (T^{-1} \sum_{t=1}^{T} x_{1,t} z_t^{'}) W_T (T^{-1} \sum_{t=1}^{T} z_t R_{1,t}) \\
\widehat{\phi}_T &= \left[(T^{-1} \sum_{t=1}^{T} x_{2,t} z_t^{'}) W_T (T^{-1} \sum_{t=1}^{T} z_t x_{2,t}^{'}) \right]^{-1} \\
& \qquad \times (T^{-1} \sum_{t=1}^{T} x_{2,t} z_t^{'}) W_T (T^{-1} \sum_{t=1}^{T} z_t S_t)
\end{aligned}
$$

$\widehat{\psi}_T$ and $\widehat{\phi}_T$ are not logically consistent but exhibit this property in the limit.

(iv) Curvature altering transformations of the population moment condition

Model implies $E[f(v_t, \theta_0)] = 0$ but estimation is based on $c(\theta_0)E[f(v_t, \theta_0)] = 0$.

GMM is invariant to curvature altering transformations if $p = q$ but only in the limit if $p < q$.

$p = q$: $\hat{\theta}_T$ solves $c(\hat{\theta}_T)T^{-1}\sum_{t=1}^{T} f(v_t, \hat{\theta}_T) = 0$.

$p < q$: FOC are

$$\left\{ \left[\frac{\partial c(\hat{\theta}_T)}{\partial \theta} \right] T^{-1} \sum_{t=1}^{T} f(v_t, \hat{\theta}_T)' + c(\hat{\theta}_T) G_T(\hat{\theta}_T)' \right\} \times$$

$$W_T T^{-1} \sum_{t=1}^{T} f(v_t, \hat{\theta}_T) = 0$$

(v) Stationarity inducing transformations of $f(.)$
Model implies $E[f(\tilde{v}_t, \theta_0)] = 0$ but $\tilde{v}_t$ is nonstationary.

Seek $H(\tilde{v}_{t-1}, \theta_0)$ such that:

- $H(\tilde{v}_{t-1}, \theta_0) f((\tilde{v}_t, \theta_0) = h(v_t, \theta_0)$ where v_t is stationary

- $E[h(v_t, \theta_0)] = 0$.

So that GMM estimation can be based on the scaled moment condition $E[h(v_t, \theta_0)] = 0$.

If one stationarity inducing transformation of $f(.)$ then many.

GMM estimator is sensitive to the choice of transformation in finite samples, but is consistent no matter which transformation is used.

Example: Consumption based asset pricing model

Recall that FOC

$$p_t c_t^{\gamma_0 - 1} = \delta_0 E[r_{t+1} c_{t+1}^{\gamma_0 - 1} | -\ _t]$$

Since $p_t c_t^{\gamma_0 - 1} \in -\ _t$, both sides of this equation were divided by $c_t^{\gamma_0 - 1} p_t$ to give

$$E[\delta_0 (r_{t+1}/p_t)(c_{t+1}/c_t)^{\gamma_0 - 1} - 1 | -\ _t] = 0$$

However FOC also implies

$$E[\delta_0 r_{t+1} c_{t+1}^{\gamma_0 - 1} - p_t c_t^{\gamma_0 - 1} | -\ _t] = 0$$

Take first approach because $x_{1,t+1} = c_{t+1}/c_t$ and $x_{2,t+1} = r_{t+1}/p_t$ are stationary but (c_t, r_t, p_t) are not.

Transformation not unique.

If $w_t \in$ - $_t$ is a stationary random variable, then division of FOC by $w_t r_t c_t^{\gamma_0 - 1}$ yields

$$E[\delta_0 w_t^{-1}(r_{t+1}/p_t)(c_{t+1}/c_t)^{\gamma_0-1} - w_t^{-1} | \text{-}\ _t] = 0 \quad (1)$$

Now consider pmc: $E[u_t(\theta_0)z_t] = 0$. If apply similar argument to (1) then obtain

$$E[u_t(\theta_0)\tilde{z}_t] = 0$$

where $\tilde{z}_t = w_t^{-1} z_t$.

3. Continuous updating GMM estimator

Recall optimal choice of weighting matrix $W = S^{-1}$.

Recall also that $S = \lim_{T\to\infty} Var[T^{1/2} g_T(\theta_0)]$ and so write $S = S(\theta_0)$.

Population analog to the GMM minimand is

$$Q_{pop}(\theta) \;=\; E[f(v_t,\theta)]' S(\theta)^{-1} E[f(v_t,\theta)]$$

In the iterated estimation,

$$Q_{iter,T}(\theta) \;=\; g_T(\theta)' \widehat{S}_T(i-1)^{-1} g_T(\theta)$$

Alternative is to minimize

$$Q_{cont,T}(\theta) \;=\; g_T(\theta)' S_T(\theta)^{-1} g_T(\theta)$$

The *continuous updating GMM estimator* is defined to be,

$$\widehat{\theta}_{cont,T} \;=\; argmin_{\theta\in\Theta}\, Q_{cont,T}(\theta)$$

Properties of Continuous updating GMM estimator:

- numerically diferent from two-step or iterated estimator
- has same limiting distribution as two-step or iterated estimator
- invariant to curvature altering transformations

Confidence sets that are invariant to reparameterization.

$$TQ_{cont,T}(\theta_0) \xrightarrow{d} \chi^2_q$$

Therefore an asymptotically valid $100(1-\alpha)\%$ confidence set for θ_0 is then given by

$$\{\,\theta : TQ_{cont,T}(\theta) < c_q(\alpha)\,\}$$

where $c_q(\alpha)$ is the $100(1-\alpha)\%$ percentile of χ^2_q distribution.