Econometricks: Short guides to econometrics

Trick 07: The Generalized Method of Moments

Davud Rostam-Afschar (Uni Mannheim)

Content

1. How to choose from too many restrictions?
2. Get the sampling error (at least approximately)
3. The econometric model
4. Consistency
5. Asymptotic normality
6. Asymptotic efficiency

Asymptotic properties of the GMM estimator

1. How to choose from too many restrictions?
2. Get the sampling error (at least approximately)
3. The econometric model
4. Consistency
5. Asymptotic normality
6. Asymptotic efficiency

Minimize the quadratic form

The overidentified GMM estimator $\hat{\theta}_{GMM}(W_n)$ for K parameters in θ identified by $L > K$ moment conditions is a function of the weighting matrix W_n for a sample of $i = 1, ..., n$ observations:

$$\hat{\theta}_{GMM}(W_n) = \underset{\theta}{\text{argmin}} q_n(\theta),$$

where the quadratic form $q_n(\theta)$ is the criterion function and is given as a function of the sample moments $\bar{m}_n(\theta)$

$$q_n(\theta) = \bar{m}_n(\theta)' W \bar{m}_n(\theta).$$

The sample moments are a function

$$\bar{m}_n(\theta) = 1/n \sum_{i=1}^{N} m(X_i, Z_i, \theta_0)$$

of the model variables X_i, the instruments Z_i, and the true parameters θ_0.

$$\underset{1\times 1}{q_n(\theta)} = \underset{1\times L}{\bar{m}_n(\theta)'} \underset{L\times L}{W} \underset{L\times 1}{\bar{m}_n(\theta)}.$$

Quadratic form criterion function $q_n(\theta) \geq 0$ is a scalar!

Weighting matrix W is symmetric (and positive definite that is $x'Wx > 0$ for all non-zero x)!

Asymptotic properties of the GMM estimator

1. How to choose from too many restrictions?
2. Get the sampling error (at least approximately)
3. The econometric model
4. Consistency
5. Asymptotic normality
6. Asymptotic efficiency

First order Taylor expansion of sample moments $\bar{m}_n(\hat{\theta}_{GMM})$ around $\bar{m}_n(\theta_0)$ at true parameters gives:

$$\bar{m}_n(\hat{\theta}_{GMM}) \approx \bar{m}_n(\theta_0) + \bar{G}_n(\bar{\theta})(\hat{\theta}_{GMM} - \theta_0),$$

where $\bar{G}_n(\bar{\theta}) = \frac{\partial \bar{m}_n(\bar{\theta})}{\partial \bar{\theta}'}$ and $\bar{\theta}$ is a point between $\hat{\theta}_{GMM}$ and θ_0.

Check the dimensions

First order Taylor expansion of sample moments $\bar{m}_n(\hat{\theta}_{GMM})$ around $\bar{m}_n(\theta_0)$ at true parameters gives:

$$\underset{L\times 1}{\bar{m}_n(\hat{\theta}_{GMM})} \approx \underset{L\times 1}{\bar{m}_n(\theta_0)} + \underset{L\times K}{\bar{G}_n(\bar{\theta})}\underset{K\times 1}{(\hat{\theta}_{GMM} - \theta_0)},$$

where $\bar{G}_n(\bar{\theta}) = \frac{\partial \underset{L\times 1}{\bar{m}_n(\bar{\theta})}}{\partial \underset{1\times K}{\bar{\theta}'}}$ and $\bar{\theta}$ is a point between $\hat{\theta}_{GMM}$ and θ_0, because of the

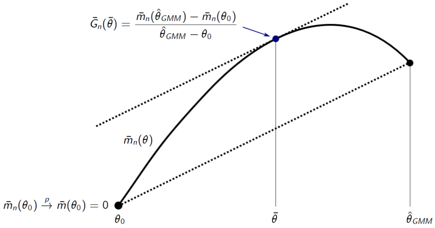
Mean value theorem...

Approximation introduced $\bar{\theta}$

...where $\bar{G}_n(\bar{\theta}) = \frac{\partial \bar{m}_n(\bar{\theta})}{\partial \bar{\theta}'}$ and $\bar{\theta}$ is a point between $\hat{\theta}_{GMM}$ and θ_0.

Mean value theorem:

$$\bar{G}_n(\bar{\theta}) = \frac{\bar{m}_n(\hat{\theta}_{GMM}) - \bar{m}_n(\theta_0)}{\hat{\theta}_{GMM} - \theta_0} \text{ for } \theta_0 < \bar{\theta} < \hat{\theta}_{GMM}.$$



Do the minimization

To minimize the quadratic form criterion, we take the first derivative of

$$q_n(\theta) = \bar{m}_n(\theta)' W \bar{m}_n(\theta)$$

$$\frac{\partial q_n(\hat{\theta}_{GMM})}{\partial \hat{\theta}_{GMM}} = 2\bar{G}_n(\hat{\theta}_{GMM})' W_n \bar{m}_n(\hat{\theta}_{GMM}) = 0.$$

Express as much as possible asymptotically

$$\frac{\partial q_n(\hat{\theta}_{GMM})}{\partial \hat{\theta}_{GMM}} = 2\bar{G}_n(\hat{\theta}_{GMM})' W_n \bar{m}_n(\hat{\theta}_{GMM}) = 0,$$

Plug in the approximation from before

$$\bar{m}_n(\hat{\theta}_{GMM}) \approx \bar{m}_n(\theta_0) + \bar{G}_n(\bar{\theta})(\hat{\theta}_{GMM} - \theta_0)$$

to obtain

$$\bar{G}_n(\hat{\theta}_{GMM})' W_n \bar{m}_n(\theta_0) + \bar{G}_n(\hat{\theta}_{GMM})' W_n \bar{G}_n(\bar{\theta})(\hat{\theta}_{GMM} - \theta_0) \approx 0$$

which we rearrange to get the very useful

$$\hat{\theta}_{GMM} \approx \theta_0 - (\bar{G}_n(\hat{\theta}_{GMM})' W_n \bar{G}_n(\bar{\theta}))^{-1} \bar{G}_n(\hat{\theta}_{GMM})' W_n \bar{m}_n(\theta_0).$$

So the estimate $\hat{\theta}_{GMM}$ is approximately the true parameter θ_0 plus a sampling error that depends on the sample moment $\bar{m}_n(\theta_0)$.

Quickly check dimensions

Useful approximation

$$\underset{K\times 1}{\hat{\theta}_{GMM}} \approx \underset{K\times 1}{\theta_0} - (\underset{K\times L}{\bar{G}_n(\hat{\theta}_{GMM})'}\underset{L\times L}{W_n}\underset{L\times K}{\bar{G}_n(\bar{\theta})})^{-1}\underset{K\times L}{\bar{G}_n(\hat{\theta}_{GMM})'}\underset{L\times L}{W_n}\underset{L\times 1}{\bar{m}_n(\theta_0)}.$$

So the estimate $\hat{\theta}_{GMM}$ is approximately the true parameter θ_0 plus a sampling error that depends on the sample moment $\bar{m}_n(\theta_0)$.

Asymptotic properties of the GMM estimator

1. How to choose from too many restrictions?
2. Get the sampling error (at least approximately)
3. The econometric model
4. Consistency
5. Asymptotic normality
6. Asymptotic efficiency

Three assumptions: moment conditions

Definition

GMM1: Moment Conditions and Identification.

$$\bar{m}(\theta_a) \neq \bar{m}(\theta_0) = E[m(X_i, Z_i, \theta_0)] = 0.$$

Identification implies that the probability limit of the GMM criterion function is uniquely minimized at the true parameters.

Three assumptions: law of large numbers

Definition

GMM2: Law of Large Numbers Applies.

$$\bar{m}_n(\theta) = 1/n \sum_{i=1}^{N} m(X_i, Z_i, \theta_0) \xrightarrow{p} E[m(X_i, Z_i, \theta_0)].$$

The data meets the conditions for a law of large numbers to apply, so that we may assume that the empirical moments converge in probability to their expectation.

Three assumptions: central limit theorem

Definition

GMM3: Central Limit Theorem Applies.

$$\sqrt{n}\bar{m}_n(\theta) = \sqrt{n}/n \sum_{i=1}^{N} m(X_i, Z_i, \theta_0) \xrightarrow{d} N[0, \Phi].$$

The empirical moments obey a central limit theorem. This assumes that the moments have a finite asymptotic covariance matrix $E[m(X_i, Z_i, \theta_0) m(X_i, Z_i, \theta_0)'] = \Phi$.

Asymptotic properties of the GMM estimator

1. How to choose from too many restrictions?
2. Get the sampling error (at least approximately)
3. The econometric model
4. Consistency
5. Asymptotic normality
6. Asymptotic efficiency

Consistency

Recall the useful approximation of the estimator:

$$\hat{\theta}_{GMM} \approx \theta_0 - (\bar{G}_n(\hat{\theta}_{GMM})' W_n \bar{G}_n(\bar{\theta}))^{-1} \bar{G}_n(\hat{\theta}_{GMM})' W_n \bar{m}_n(\theta_0).$$

Assumption GMM2 implies that

$$\bar{m}_n(\theta_0) = 1/n \sum_{i=1}^{N} m(X_i, Z_i, \theta_0) \xrightarrow{p} E[m(X_i, Z_i, \theta_0)] = \bar{m}(\theta_0).$$

That is, the sample moment equals the population moment in probability. Assumption GMM1 implies that

$$\bar{m}(\theta_0) = 0.$$

Then

$$\bar{m}_n(\theta_0) \xrightarrow{p} \bar{m}(\theta_0) = 0,$$

such that

$$\hat{\theta}_{GMM} \xrightarrow{p} \theta_0 \text{ for } N \to \infty$$

That is, by GMM1 and GMM2 the GMM estimator is consistent.

Asymptotic properties of the GMM estimator

1. How to choose from too many restrictions?
2. Get the sampling error (at least approximately)
3. The econometric model
4. Consistency
5. Asymptotic normality
6. Asymptotic efficiency

Asymptotic normality

Recall the useful approximation of the estimator:

$$\hat{\theta}_{GMM} \approx \theta_0 - (\bar{G}_n(\hat{\theta}_{GMM})' W_n \bar{G}_n(\bar{\theta}))^{-1} \bar{G}_n(\hat{\theta}_{GMM})' W_n \bar{m}_n(\theta_0).$$

Rewrite to obtain

$$\sqrt{n}(\hat{\theta}_{GMM} - \theta_0) \approx -(\bar{G}_n(\hat{\theta}_{GMM})' W_n \bar{G}_n(\bar{\theta}))^{-1} \bar{G}_n(\hat{\theta}_{GMM})' W_n \sqrt{n}\bar{m}_n(\theta_0),$$

The right hand side has several parts for which we made assumptions on what happens when $N \to \infty$. Under the central limit theorem (GMM3)

$$\sqrt{n}\bar{m}_n(\theta_0) \stackrel{d}{\to} N[0, \Phi]$$

$$plim W_n = W$$

$$plim \bar{G}_n(\hat{\theta}_{GMM}) = plim \bar{G}_n(\bar{\theta}) = plim \frac{\partial m(X_i, Z_i, \theta_0)}{\partial \theta_0'} = E\left[\frac{\partial \bar{m}(\theta_0)}{\partial \theta_0'}\right] = \Gamma(\theta_0)$$

Asymptotic normality

With $plim W_n = W$ and

$$plim \bar{G}_n(\hat{\theta}_{GMM}) = plim \bar{G}_n(\bar{\theta}) = \Gamma(\theta_0)$$

the expression

$$\sqrt{n}(\hat{\theta}_{GMM} - \theta_0) \approx -(\bar{G}_n(\hat{\theta}_{GMM})' W_n \bar{G}_n(\bar{\theta}))^{-1} \bar{G}_n(\hat{\theta}_{GMM})' W_n \sqrt{n} \bar{m}_n(\theta_0)$$

becomes

$$\sqrt{n}(\hat{\theta}_{GMM} - \theta_0) \approx -(\Gamma(\theta_0)' W \Gamma(\theta_0))^{-1} \Gamma(\theta_0)' W \sqrt{n} \bar{m}_n(\theta_0)$$

from which we get the variance V. So

$$\sqrt{n}(\hat{\theta}_{GMM} - \theta_0) \xrightarrow{d} N[0, V]$$

with

$$\underset{K \times K}{V} = 1/n[\Gamma(\theta_0)' W \Gamma(\theta_0)]^{-1} [\Gamma(\theta_0)' W \Phi W' \Gamma(\theta_0)] [\Gamma(\theta_0)' W \Gamma(\theta_0)]^{-1}$$

That is by GMM1, GMM2, and GMM3 the GMM estimator is asymptotic normal.

Asymptotic properties of the GMM estimator

1. How to choose from too many restrictions?
2. Get the sampling error (at least approximately)
3. The econometric model
4. Consistency
5. Asymptotic normality
6. Asymptotic efficiency

Asymptotic efficiency

Which weighting matrix W gives us the smallest possible asymptotic variance of the GMM estimator $\hat{\theta}_{GMM}$.
The variance of the GMM estimator V depends on the choice of W

$$V = 1/n[\Gamma(\theta_0)' W \Gamma(\theta_0)]^{-1}[\Gamma(\theta_0)' W \Phi W' \Gamma(\theta_0)][\Gamma(\theta_0)' W \Gamma(\theta_0)]^{-1}$$

So let us minimize V to get the *optimal* weight matrix. Try from GMM3

$$\underset{n\to\infty}{plim} W_n = W = \Phi^{-1}$$

$$V_{GMM,optimal} = 1/n[\Gamma(\theta_0)'\Phi^{-1}\Gamma(\theta_0)]^{-1}[\Gamma(\theta_0)'\Phi^{-1}\Phi\Phi^{-1'}\Gamma(\theta_0)][\Gamma(\theta_0)'\Phi^{-1}\Gamma(\theta_0)]^{-1}$$

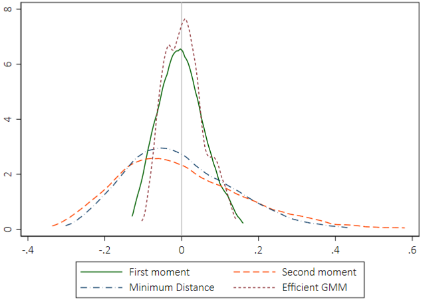
Which can be simplified to

$$V_{GMM,optimal} = 1/n[\Gamma(\theta_0)'\Phi^{-1}\Gamma(\theta_0)]^{-1}$$

Asymptotic efficiency

$$V_{GMM,optimal} = 1/n[\Gamma(\theta_0)'\Phi^{-1}\Gamma(\theta_0)]^{-1}$$

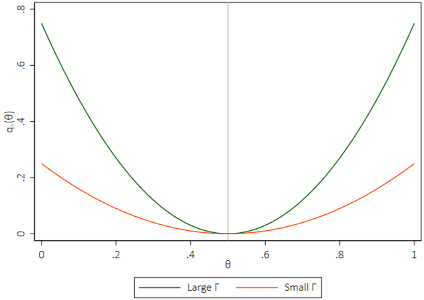
If Φ is small, there is little variation of this specific sample moment around zero and the moment condition is very informative about θ_0. So it is best to assign a high weight to it.



Asymptotic efficiency

$$V_{GMM,optimal} = 1/n[\Gamma(\theta_0)'\Phi^{-1}\Gamma(\theta_0)]^{-1}$$

If Γ is large, there is a large penalty from violating the moment condition by evaluating at $\theta \neq \theta_0$. Then the moment condition is very informative about θ_0. V is inversely related to Γ.



Estimate the variance in practice

$$\hat{V}_{GMM,optimal} = 1/n[\bar{G}_n(\hat{\theta})'\Phi_n^{-1}\bar{G}_n(\hat{\theta})]^{-1}$$

Consistent estimator

$$\Phi_n = NV(\bar{m}_n(\hat{\theta}))$$

$$\bar{G}_n(\hat{\theta}) = \frac{\partial m(X_i, Z_i, \hat{\theta})}{\partial \hat{\theta}'}$$

References I

Greene, W. H. (2011): *Econometric Analysis*. Prentice Hall, 5 edn.

Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50(4), 1029–1054.

——— (2012): "Proofs for large sample properties of generalized method of moments estimators," *Journal of Econometrics*, 170(2), 325–330, Thirtieth Anniversary of Generalized Method of Moments.