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

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:

$$\widehat{ heta}_{GMM}(W_n) = \operatorname*{argmin}_{ heta} q_n(heta),$$

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

$$q_n(heta) = ar{m}_n(heta)' W ar{m}_n(heta).$$

The sample moments are a function

$$ar{m}_n(heta) = 1/n \sum_{i=1}^N m(X_i, Z_i, heta_0)$$

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

What are the properties of the quadratic form

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

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

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

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:

$$ar{m}_n(\hat{ heta}_{GMM}) pprox ar{m}_n(heta_0) + ar{G}_n(ar{ heta})(\hat{ heta}_{GMM} - heta_0),$$

where $ar{G}_n(ar{ heta}) = rac{\partial ar{m}_n(ar{ heta})}{\partial ar{ heta}'}$ and $ar{ heta}$ is a point between $\hat{ heta}_{GMM}$ and $heta_0$.

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 $\bar{G}_{n}(\bar{\theta}) = \frac{\partial \bar{m}_{n}(\bar{\theta})}{\frac{L \times 1}{\partial |\bar{\theta}'|}}$ 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:

 $ar{G}_n(ar{ heta}) = rac{ar{m}_n(\hat{ heta}_{GMM}) - ar{m}_n(heta_0)}{\hat{ heta}_{GMM} - heta_0} ext{ for } heta_0 < ar{ heta} < \hat{ heta}_{GMM}.$



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

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

$$rac{\partial q_n(\hat{ heta}_{GMM})}{\partial \hat{ heta}_{GMM}} = 2 \, ar{G}_n(\hat{ heta}_{GMM})' W_n ar{m}_n(\hat{ heta}_{GMM}) = 0.$$

Express as much as possible asymptotically

$$rac{\partial q_n(\hat{ heta}_{GMM})}{\partial \hat{ heta}_{GMM}} = 2 ar{G}_n(\hat{ heta}_{GMM})' W_n ar{m}_n(\hat{ heta}_{GMM}) = 0,$$

Plug in the approximation from before

$$ar{m}_n(\hat{ heta}_{GMM})pproxar{m}_n(heta_0)+ar{G}_n(ar{ heta})(\hat{ heta}_{GMM}- heta_0)$$

to obtain

$$ar{G}_n(\hat{ heta}_{GMM})'W_nar{m}_n(heta_0)+ar{G}_n(\hat{ heta}_{GMM})'W_nar{G}_n(ar{ heta})(\hat{ heta}_{GMM}- heta_0)pprox 0$$

which we rearrange to get the very useful

$$\hat{ heta}_{GMM} pprox heta_0 - (ar{G}_n(\hat{ heta}_{GMM})'W_nar{G}_n(ar{ heta}))^{-1}ar{G}_n(\hat{ heta}_{GMM})'W_nar{m}_n(heta_0).$$

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

Useful approximation

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

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

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.

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

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

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

$$\bar{m}(heta_0)=0.$$

Then

$$ar{m}_n(heta_0) \stackrel{p}{
ightarrow} ar{m}(heta_0) = 0,$$

such that

$$\hat{ heta}_{GMM} \stackrel{p}{
ightarrow} heta_0$$
 for $N
ightarrow \infty$

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

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 $N \rightarrow \infty$. Under the central limit theorem (GMM3)

$$\sqrt{n}\bar{m}_n(heta_0) \stackrel{d}{
ightarrow} N[0,\Phi]$$

$$plimW_n = W$$
$$plim\bar{G}_n(\hat{\theta}_{GMM}) = plim\bar{G}_n(\bar{\theta}) = plim\frac{\partial m(X_i, Z_i, \theta_0)}{\partial \theta'_0} = E\left[\frac{\partial \bar{m}(\theta_0)}{\partial \theta'_0}\right] = \Gamma(\theta_0)$$

Asymptotic normality

With $plimW_n = W$ 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

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

with

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

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

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

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

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

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

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

Consistent estimator

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

$$ar{G}_n(\hat{ heta}) = rac{\partial m(X_i, Z_i, \hat{ heta})}{\partial \hat{ heta'}}$$

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.