Advanced Econometrics

Chapter 11: Generalized Method of Moments

In Choi

Sogang University

References

- Newey, W. and D. McFadden (1994). "Large sample estimation and hypothesis testing," Chapter 36 in R. Engle and D. McFadden (eds.) Handbook of Econometrics, Vol. IV, North-Holland.
- Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.

Method of moments estimator

- Consider the linear model

$$y_i = \beta' x_i + u_i, \ (i = 1, ..., n),$$

where x_i is endogenous.

- Let $w_i = (y_i, x_i', z_i')'$ and $g(w_i, \beta) = z_i(y_i - \beta' x_i)$. Suppose that

$$E\left(g(w_i, \beta)\right) = 0$$

for all i.

Method of moments estimator

- The sample analogue is

$$g_n(\tilde{\beta}) = \frac{1}{n}\sum_{i=1}^{n} g(w_i, \tilde{\beta}) = \frac{1}{n}\sum_{i=1}^{n} z_i(y_i - \tilde{\beta}' x_i) = 0.$$

Solving this equation with respect to β, we obtain

$$\tilde{\beta} = \left(\sum_{i=1}^{n} x_i z_i'\right)^{-1} \left(\sum_{i=1}^{n} z_i y_i\right),$$

where it is assumed that $\left(\sum_{i=1}^{n} x_i z_i'\right)^{-1}$ is well defined.

- This is the method of moments estimator.

Generalized method of moments estimator

Definition

- Let

$$J(\beta, \hat{W}) = n g_n(\beta)' \hat{W} g_n(\beta),$$

where $\hat{W}$ is a $K \times K$ symmetric positive definite matrix such that $\hat{W} \xrightarrow{p} W$.

- The generalized method of moment estimator is defined as

$$\tilde{\beta} = \arg\min_{\beta} J(\beta, \hat{W})$$

- For the linear IV estimation,

$$g_n(\beta) = \frac{1}{n} \sum_{i=1}^{n} z_i (y_i - \tilde{\beta}' x_i).$$

- The first order condition for the linear IV estimation is

$$S_{zx}'\hat{W}S_{zy} = S_{zx}'\hat{W}S_{zx}\tilde{\beta},$$

  where $S_{zx} = \frac{1}{n}\sum_{i=1}^{n} z_i x_i'$ and $S_{zy} = \frac{1}{n}\sum_{i=1}^{n} z_i y_i$.
- If $S_{zx}'\hat{W}S_{zx}$ is invertible,
  $\tilde{\beta} = \left(S_{zx}'\hat{W}S_{zx}\right)^{-1} S_{zx}'\hat{W}S_{zy} (= \left(X'Z\hat{W}Z'X\right)^{-1} X'Z\hat{W}Z'y)$.
- If $\hat{W} = \left(\frac{1}{n}\sum_{i=1}^{n} z_i z_i'\right)^{-1}$, $\tilde{\beta}$ is the 2SLS estimator.

- Write
$$\tilde{\beta} - \beta = \left(S'_{zx}\hat{W}S_{zx}\right)^{-1} S'_{zx}\hat{W}\bar{g}$$
where $\bar{g} = \frac{1}{n}\sum_{i=1}^{n} z_i u_i$.
- Assume
(i) $S_{zx} = \frac{1}{n}\sum_{i=1}^{n} z_i x_i' \xrightarrow{p} M_{zx}$
(ii) $\sqrt{n}\bar{g} \xrightarrow{d} N(0, S)$, where $S = p\lim \frac{1}{n}\sum_{i=1}^{n} z_i z_i' u_i^2$.

- Then,

$$\sqrt{n}\left(\tilde{\beta}-\beta\right) \xrightarrow{d} N\left(0, \left(M_{zx}' W M_{zx}\right)^{-1} M_{zx}' W S W M_{zx} \left(M_{zx}' W M_{zx}\right)^{-1}\right).$$

- Since

$$\left(M_{zx}' W M_{zx}\right)^{-1} M_{zx}' W S W M_{zx} \left(M_{zx}' W M_{zx}\right)^{-1} \geq \left(M_{zx}' S^{-1} M_{zx}\right)^{-1},$$

  the optimal choice of $\hat{W}$ is $\hat{S}^{-1}$.

- A consistent estimator of S is $\hat{S} = \frac{1}{n}\sum_{i=1}^{n} z_i z_i' \hat{u}_i^2$, where $\hat{u}_i$ denotes the GMM residual.

- The optimal GMM estimator is

$$\tilde{\beta}_{opt} = \left(S'_{zx}\hat{S}^{-1}S_{zx}\right)^{-1} S'_{zx}\hat{S}^{-1}S_{zy}.$$

- $\sqrt{n}\left(\tilde{\beta}_{opt} - \beta\right) \xrightarrow{d} N\left(0, \left(M'_{zx}S^{-1}M_{zx}\right)^{-1}\right).$
- We can construct Wald test statistics and t-ratios based on this result.

- When all the conditions for the asymptotic result for $\sqrt{n}\left(\tilde{\beta}_{opt} - \beta\right)$ hold,

$$J(\tilde{\beta}_{opt}, \hat{S}^{-1}) = ng_n(\tilde{\beta}_{opt})'\hat{S}^{-1}g_n(\tilde{\beta}_{opt}) \xrightarrow{d} \chi^2(L-K),$$

  where L is the dimension of z_i.
- For this test to be valid, we should have L (# of instruments) $> K$ (# of regressors).
- When $J(\tilde{\beta}_{opt}, \hat{S}^{-1})$ takes a large value, the condition between z and x and/or some other conditions are false.
- Only when $J(\tilde{\beta}_{opt}, \hat{S}^{-1})$ takes a small value, we are confident that the GMM estimator is appropriate to use.

- For some function $g(w_i, \theta^o) \in R^L$, Assume that

$$E\left(g(w_i, \theta^o)\right) = 0,$$

where $\theta^o \in \Theta \subset R^p$ is the true parameter value (θ^o was β in previous discussions).

- If $\hat{W} \xrightarrow{p} W$,

$$\left(\frac{1}{n}\sum_{i=1}^{n} g(w_i, \theta)\right)' \hat{W} \left(\frac{1}{n}\sum_{i=1}^{n} g(w_i, \theta)\right) \xrightarrow{p} \left(Eg(w, \theta)\right)' WEg(w, \theta).$$

- Since W is positive definite, $\left(Eg(w_i, \theta)\right)' W \left(Eg(w_i, \theta)\right)$ is uniquely minimized at $\theta = \theta^o$.

Theorem 1 Assume
(i) Θ is compact.
(ii) For each $w \in$ W, $g(w, \cdot)$ is continuous on Θ.
(iii) $|g_j(w, \theta^o)| \leq b(w)$ for all $\theta \in \Theta$ and $j = 1, ..., L$ where $b(\cdot)$ is a nonnegative function on W such that $E(b(w)) < \infty$.
(iv) $\hat{W} \xrightarrow{p} W$, a positive definite matrix.
(v) θ^o is unique.
Then, the GMM estimator of θ is consistent.

- Let $G_o = E\left(\frac{\partial g(w_i,\theta)}{\partial \theta'}\right)$, $A_o = G_o' W G_o$, $\Lambda_o = E\left(g(w_i,\theta^o) g(w_i,\theta^o)'\right)$, $B_o = G_o' W \Lambda_o W G_o$. Then,

$$\sqrt{n}(\hat{\theta}_{GMM} - \theta) \xrightarrow{d} N(0, A_o^{-1} B_o A_o^{-1}).$$

- A consistent estimator of Λ_o^{-1} is the optimal choice for $\hat{W}$.