GMM Basics

$$y_i = z_i'\delta + \varepsilon_i$$

$$E(g_i) \equiv E(x_i\varepsilon_i) = E(x_i(y_i - z_i'\delta)) = 0$$

$$g_n(\tilde{\delta}) = \frac{1}{n}\sum_{i=1}^{n} x_i(y_i - z_i'\tilde{\delta}) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i z_i'\tilde{\delta} = s_{xy} - S_{xz}\tilde{\delta}$$

GMM estimator:

$$\tilde{\delta}(\hat{W}) = \arg\min_{\tilde{\delta}} n\, g_n(\tilde{\delta})'\hat{W} g_n(\tilde{\delta}) = \arg\min_{\tilde{\delta}} n(s_{xy} - S_{xz}\tilde{\delta})'\hat{W}(s_{xy} - S_{xz}\tilde{\delta})$$

FOC:

$$\tilde{\delta}(\hat{W}) = \left(S_{xz}'\hat{W}S_{xz}\right)^{-1} S_{xz}'\hat{W}s_{xy}$$

sample error:

$$\tilde{\delta}(\hat{W}) - \delta = \left(S_{xz}'\hat{W}S_{xz}\right)^{-1} S_{xz}'\hat{W}s_{xy} - \left(S_{xz}'\hat{W}S_{xz}\right)^{-1}(S_{xz}'\hat{W}S_{xz})\delta = (S_{xz}'\hat{W}S_{xz})^{-1}S_{xz}'\hat{W}\bar{g}$$

base on

$$s_{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i(z_i'\delta + \varepsilon_i) = \frac{1}{n}\sum_{i=1}^{n} x_i z_i'\delta + \frac{1}{n}\sum_{i=1}^{n} x_i\varepsilon_i = S_{xz}\delta + \bar{g}$$

Assumptions

the same as chap4

Proposition 3.1 Asymptotic distribution of GMM estimator

Consistency

Under Ass 3.1-3.4:

$$p\lim_{n\to\infty} \tilde{\delta}(\hat{W}) = \delta$$

Proof:

Ass 3.2, 3.4 ⇒ $S_{xz} \xrightarrow{p} \Sigma_{xz}$, Σ_{xz} is column full rank$(K \geq L)$

Ass 3.2-3.3 ⇒ $\bar{g} = \frac{1}{n}\sum_{i=1}^{n} g_i \xrightarrow{p} 0$

$\hat{W} \xrightarrow{p} W$ by the definition of W(**symmetric** and **positive definite** weight matrix)

By Slutsky theorem:

$$S'_{xz}\hat{W}S_{xz} \xrightarrow{p} \Sigma'_{xz}W\Sigma_{xz}$$

$$S'_{xz}\hat{W}\bar{g} \xrightarrow{p} \Sigma'_{xz}W \cdot 0 = 0$$

hence:

$$\tilde{\delta}(\hat{W}) - \delta \xrightarrow{p} \left(\Sigma'_{xz}W\Sigma_{xz}\right)^{-1} \cdot 0 = 0$$

$$p\lim \tilde{\delta}(\hat{W}) = \delta$$

Asymptotic normality

Under Ass 3.5:

$$\sqrt{n}\left(\tilde{\delta}(\hat{W}) - \delta\right) \xrightarrow{d} N\left(0,\ \mathrm{Avar}(\tilde{\delta}(\hat{W}))\right)$$

$$\mathrm{Avar}(\tilde{\delta}(\hat{W})) = \left(\Sigma'_{xz}W\Sigma_{xz}\right)^{-1}\Sigma'_{xz}WSW\Sigma_{xz}\left(\Sigma'_{xz}W\Sigma_{xz}\right)^{-1}$$

$$S = E[g_i g_i'] = E[x_i x_i' \varepsilon_i^2]$$

proof:

$$\sqrt{n}\left(\tilde{\delta}(\hat{W}) - \delta\right) = \left(S'_{xz}\hat{W}S_{xz}\right)^{-1}S'_{xz}\hat{W} \cdot \sqrt{n}\,\bar{g}$$

By CLT:

$$\sqrt{n}\,\bar{g} \xrightarrow{d} N(0, S)$$

By Slutsky theorem:

$$S'_{xz}\hat{W}S_{xz} \xrightarrow{p} \Sigma'_{xz}W\Sigma_{xz}$$

$$S'_{xz}\hat{W} \xrightarrow{p} \Sigma'_{xz}W$$

hence:

$$\sqrt{n}\left(\tilde{\delta}(\hat{W}) - \delta\right) \xrightarrow{d} \left(\Sigma'_{xz}W\Sigma_{xz}\right)^{-1}\Sigma'_{xz}W \cdot N(0, S)$$

$$\mathrm{Avar}(\tilde{\delta}(\hat{W})) = \left(\Sigma'_{xz}W\Sigma_{xz}\right)^{-1}\Sigma'_{xz}WSW\Sigma_{xz}\left(\Sigma'_{xz}W\Sigma_{xz}\right)^{-1}$$

Consistent estimate of $\mathrm{Avar}(\tilde{\delta}(\hat{W}))$

$$\widehat{\mathrm{Avar}}(\tilde{\delta}(\hat{W})) \xrightarrow{p} \mathrm{Avar}(\tilde{\delta}(\hat{W}))$$

Proof:

$$\widehat{\text{Avar}}(\tilde{\delta}(\hat{W})) = \left(S'_{xz}\hat{W}S_{xz}\right)^{-1}S'_{xz}\hat{W}\hat{S}\hat{W}S_{xz}\left(S'_{xz}\hat{W}S_{xz}\right)^{-1}$$

$$S_{xz} \xrightarrow{p} \Sigma_{xz}, \quad \hat{W} \xrightarrow{p} W, \quad \hat{S} \xrightarrow{p} S$$

$$\widehat{\text{Avar}}(\tilde{\delta}(\hat{W})) \xrightarrow{p} \left(\Sigma'_{xz}W\Sigma_{xz}\right)^{-1}\Sigma'_{xz}WSW\Sigma_{xz}\left(\Sigma'_{xz}W\Sigma_{xz}\right)^{-1}$$

Proposition 3.2 consistent estimation of error variance

Under Ass 3.1-3.2 + $E(z_i z_i')$ exists and is finite:

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^2 \xrightarrow{p} E(\varepsilon_i^2)$$

Proof:

$$\hat{\varepsilon}_i = \varepsilon_i - z_i'(\hat{\delta}-\delta)$$

$$\hat{\varepsilon}_i^2 = \varepsilon_i^2 - 2(\hat{\delta}-\delta)'z_i\varepsilon_i + (\hat{\delta}-\delta)'z_iz_i'(\hat{\delta}-\delta)$$

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^2 = \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 - 2(\hat{\delta}-\delta)'\left(\frac{1}{n}\sum_{i=1}^{n}z_i\varepsilon_i\right) + (\hat{\delta}-\delta)'\left(\frac{1}{n}\sum_{i=1}^{n}z_iz_i'\right)(\hat{\delta}-\delta)$$

- For the first term in RHS, using Ass 3.2 + WLLN: $\frac{1}{n}\sum\varepsilon_i^2 \xrightarrow{p} E(\varepsilon_i^2)$
- For the second term, $\hat{\delta}-\delta \to 0$ (consistency), $\frac{1}{n}\sum z_i\varepsilon_i \xrightarrow{p} E(z_i\varepsilon_i)$ exists and is finite (by Cauchy-Schwartz inequality)
- For the third term, $\hat{\delta}-\delta \to 0, \ \frac{1}{n}\sum z_iz_i' \xrightarrow{p} E(z_iz_i')$ exists and is finite

hence:

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^2 \xrightarrow{p} E(\varepsilon_i^2)$$

The Efficient GMM Estimator

To minimize $\text{Avar}(\tilde{\delta}(\hat{W}))$, choose $\hat{S}^{-1}$ as the weighting matrix.

$$Avar(\tilde{\delta}(\hat{W})) = (\Sigma'_{xz}W\Sigma_{xz})^{-1}\Sigma'_{xz}WSW\Sigma_{xz}(\Sigma'_{xz}W\Sigma_{xz})^{-1}$$

$$Avar(\tilde{\delta}(\hat{S}^{-1})) = (\Sigma'_{xz}S^{-1}\Sigma_{xz})^{-1}$$

$$\widehat{Avar}(\tilde{\delta}(\hat{S}^{-1})) = (S'_{xz}\hat{S}^{-1}S_{xz})^{-1}$$

From GMM to 2SLS

Under Ass 3.7(conditional homoskedasticity):

$$S = E(x_i x_i' \varepsilon_i^2) = E[x_i x_i' E(\varepsilon_i^2 | x_i)] = \sigma^2 E(x_i x_i')$$

as is proofed in chap2:
$s^2 \xrightarrow{p} \sigma^2$, $S_{xx} \xrightarrow{p} E(x_i x_i') = \Sigma_{xx}$, $\hat{S} \xrightarrow{p} S$

$$\begin{aligned}
\hat{\beta}_{GMM} &= (S_{xz}' \hat{S}^{-1} S_{xz})^{-1} S_{xz}' \hat{S}^{-1} s_{xy} \\
&= \left(S_{xz}' (s^2 S_{xx})^{-1} S_{xz}\right)^{-1} S_{xz}' (s^2 S_{xx})^{-1} s_{xy} \\
&= \left(S_{xz}' S_{xx}^{-1} S_{xz}\right)^{-1} S_{xz}' S_{xx}^{-1} s_{xy} \\
&= \left(\frac{Z'X}{n} (\frac{X'X}{n})^{-1} \frac{X'Z}{n}\right)^{-1} \frac{Z'X}{n} (\frac{X'X}{n})^{-1} \frac{X'y}{n} \\
&= \left(Z'X(X'X)^{-1}X'Z\right)^{-1} Z'X(X'X)^{-1}X'y \\
&= \left(Z'P_X P_X Z\right)^{-1} Z'P_X y \\
&= (\hat{Z}'\hat{Z})^{-1} \hat{Z}'y \\
&= \hat{\beta}_{2SLS}
\end{aligned}$$

don't forget the projection matrix P is symmetric and idempotent.

Proposition 3.3 Robust t ratio and Wald statistics

Under Ass 3.1-3.5:

(a) Under the null $H_0 : \delta = \bar{\delta}$

$$t \equiv \frac{\sqrt{n}(\tilde{\delta}(\hat{W}) - \bar{\delta})}{\sqrt{\widehat{Avar}(\tilde{\delta}(\hat{W}))}} \xrightarrow{d} N(0,1)$$

(b) Under the null $H_0 : R\delta = r$

$$W \equiv n(R\tilde{\delta}(\hat{W}) - r)' \{R[\widehat{Avar}(\tilde{\delta}(\hat{W}))]R'\}^{-1} (R\tilde{\delta}(\hat{W}) - r) \xrightarrow{d} \chi^2(\#r)$$

(c) Under the null $H_0 : a(\delta) = 0$

$$W \equiv n\, a(\tilde{\delta}(\hat{W}))' \{A(\tilde{\delta}(\hat{W}))[\widehat{Avar}(\tilde{\delta}(\hat{W}))] A(\tilde{\delta}(\hat{W}))'\}^{-1} a(\tilde{\delta}(\hat{W})) \xrightarrow{d} \chi^2(\#a)$$

(d) $\bar{\delta}$ is from restricted model, $\tilde{\delta}$ is from unrestricted model:

$$LR \equiv J(\bar{\delta}(\hat{S}^{-1}), \hat{S}^{-1}) - J(\tilde{\delta}(\hat{S}^{-1}), \hat{S}^{-1}) \xrightarrow{d} \chi^2$$

Proposition 3.6 Hansen's test of overidentifying restrictions

Under Ass 3.1-3.5:

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

$\sqrt{n}\bar{g} \xrightarrow{d} N(0, \mathrm{S}), \hat{S} \xrightarrow{p} S$

Under conditional homoskedasticity, J becomes **Sargan Statistic**:

$$n(s_{xy} - S_{xz}\tilde{\delta}_{2SLS})'(s^2 S_{xx})^{-1}(s_{xy} - S_{xz}\tilde{\delta}_{2SLS})$$

Proposition 3.7 Testing a subset of orthogonality conditions

Under Ass 3.1-3.5:

$$C \equiv J - J_1 \rightarrow \chi^2(K - K_1)$$

need $K_1 \geq L$