Econometrics II TA Session #8

Royou Kiku

December 22nd, 2021

Contents

1.Hansen's J Test

- The J test is called a test for over-identifying restrictions (過剰識別制約).

$$H_0: E(z'u) = 0$$
$$H_1: E(z'u) \neq 0$$

- Test statistics:

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z'_i \hat{u}_i\right)' \left(V\left(\widehat{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i \hat{u}_i}\right)\right)^{-1} \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z'_i \hat{u}_i\right) \longrightarrow \chi^2(r-k)$$

Where $\hat{u}_i = y_i - x_i\beta_{GMM}$.

$V\left(\widehat{\frac{1}{n}\sum_{i=1}^{n} z'_i}\hat{u}_i\right)$ indicates the estimate of $V\left(\frac{1}{n}\sum_{i=1}^{n} z'_i u_i\right)$ for $u_i = y_i - x_i\beta$.

In the case of GMM

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z'_i u_i \longrightarrow N(0, \Sigma)$$

Where $\Sigma = V\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z'_i u_i\right)$.

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z'_i u_i\right)' \Sigma^{-1} \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z'_i u_i\right) \longrightarrow \chi^2(r)$$

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z'_i \widehat{u}_i\right)' \Sigma^{-1} \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z'_i \widehat{u}_i\right) \longrightarrow \chi^2(r-k)$$

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z'_i \widehat{u}_i\right)' \hat{\Sigma}^{-1} \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z'_i \widehat{u}_i\right) \longrightarrow \chi^2(r-k)$$

2.Generalized Method of Moments II (Nonlinear Case)

- Consider the general case

$$E\left(h(\theta; w)\right) = 0$$

Which is the orthogonality condition.

Where θ is a $k \times 1$ vector of parameter, $h(\theta; w)$ is a $r \times 1$ vector for $r \geq k$.

Let $w_i = (y_i, x_i)$ be the ith observed data.

$$g(\theta; W) = \frac{1}{n} \sum_{i=1}^{n} h(\theta; w_i)$$

Where $W = \{w_n, w_{n-1}, \dots, w_1\}$.

$g(\theta; W)$ is a $r \times 1$ vector for $r \geq k$.

- In the same way as the GMM estimator in linear case, we define the GMM estimator $\hat{\theta}$, which minimizes:

$$g(\theta; W)' S^{-1} g(\theta; W)$$

- The first-order condition of GMM is

$$\frac{\partial g(\theta; W)'}{\partial \theta} S^{-1} g(\theta; W) = 0$$

- The second derivative is omitted.

- Solving for the first-order condition, we can get

$$\hat{\theta}^{(i+1)} = \hat{\theta}^{(i)} - \left(\widehat{D}^{(i)\prime} S^{-1} \widehat{D}^{(i)}\right)^{-1} \widehat{D}^{(i)\prime} S^{-1} g(\hat{\theta}^{(i)}; W)$$

- Calculation procedure of the first-order condition
- To obtain $\hat{\theta}$, we linearize the first-order condition around $\theta = \hat{\theta}$,

$$0 = \frac{\partial g(\theta;W)'}{\partial \theta} S^{-1} g(\theta;W)$$

$$\approx \frac{\partial g(\hat{\theta};W)'}{\partial \theta} S^{-1} g(\hat{\theta};W) + \frac{\partial g(\hat{\theta};W)'}{\partial \theta} S^{-1} \frac{\partial g(\hat{\theta};W)}{\partial \theta'} (\theta - \hat{\theta})$$

$$= \hat{D}' S^{-1} \mathrm{g}(\hat{\theta};W) + \hat{D}' S^{-1} \hat{D} (\theta - \hat{\theta})$$

Where $\hat{D} = \frac{\partial g(\hat{\theta};W)}{\partial \theta'}$, which is a $r \times k$ matrix.

- Rewriting, we have the following equation:

$$\theta - \hat{\theta} = -(\hat{D}'S^{-1}\hat{D})^{-1}\hat{D}'S^{-1}g(\hat{\theta};W)$$

- Replacing θ and $\hat{\theta}$ by $\hat{\theta}^{(i+1)}$ and $\hat{\theta}^{(i)}$, respectively.

$$\hat{\theta}^{(i+1)} = \hat{\theta}^{(i)} - \left(\hat{D}^{(i)\prime}S^{-1}\hat{D}^{(i)}\right)^{-1}\hat{D}^{(i)\prime}S^{-1}g(\hat{\theta}^{(i)};W)$$

Where $\hat{D}^{(i)} = \frac{\partial g(\hat{\theta}^{(i)};W)}{\partial \theta\prime}$.

- How to calculate the weight matrix S ?
- If $h(\theta; w_i)$, $\mathrm{i} = 1, \dots, \mathrm{n}$, are mutually independent, S is

$$S = V(\sqrt{n}g(\theta; W)) = nE(g(\theta; W)g(\theta; W)')$$

$$= nE\left(\left(\frac{1}{n}\sum_{i=1}^{n} h(\theta; w_i)\right)\left(\frac{1}{n}\sum_{j=1}^{n} h(\theta; w_j)\right)'\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} E\left(h(\theta; w_i)h(\theta; w_j)'\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} E(h(\theta; w_i)h(\theta; w_i)')$$

- Note that

(i) $E(h(\theta; w_i)) = 0$ for all i and accordingly $E(g(\theta; W)) = 0$,

(ii) $g(\theta; W) = \frac{1}{n}\sum_{i=1}^{n} h(\theta; w_i) = \frac{1}{n}\sum_{j=1}^{n} h(\theta; w_j)$,

(iii)$E\left(h(\theta; w_i)h(\theta; w_j)'\right) = 0$ for $i \neq j$.

- The estimator of S, denote by $\hat{S}$ is given by

$$\hat{S} = \frac{1}{n}\sum_{i=1}^{n} h(\hat{\theta}; w_i)h(\hat{\theta}; w_i)' \longrightarrow S$$

- Suppose that $h(\theta; w_i)$ is stationary and $\Gamma_\tau = E(h(\theta; w_i)h(\theta; w_{i-\tau})') < \infty$

Stationarity:

(i) $E(h(\theta; w_i))$ does not depend on i.

(ii) $E(h(\theta; w_i)h(\theta; w_{i-\tau})')$ depends on time difference τ.

$$
\begin{aligned}
S &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} E\big(h(\theta; w_i)h(\theta; w_j)'\big) \\
&= \frac{1}{n}\left(n\Gamma_0 + (n-1)(\Gamma_1 + \Gamma'_1) + (n-2)(\Gamma_2 + \Gamma'_2) + \cdots + (\Gamma_{n-1} + \Gamma'_{n-1})\right) \\
&= \Gamma_0 + \textstyle\sum_{i=1}^{n-1} \frac{n-i}{n}(\Gamma_i + \Gamma'_i) = \Gamma_0 + \sum_{i=1}^{n-1}(1 - \frac{i}{n})(\Gamma_i + \Gamma'_i) \\
&= \Gamma_0 + \sum_{i=1}^{q}\left(1 - \frac{i}{q+1}\right)(\Gamma_i + \Gamma'_i)
\end{aligned}
$$

Where $\Gamma'_\tau = E(h(\theta; w_{i-\tau})h(\theta; w_i)') = \Gamma_{-\tau}$.

- S is estimated as

$$\hat{S} = \widehat{\Gamma_0} + \sum_{i=1}^{q} \left(1 - \frac{i}{q+1}\right) \left(\hat{\Gamma}_i + \hat{\Gamma}'_i\right)$$

This is the Newey-West estimator. Where n is replaced by $q + 1$, where $q \leq n$. We need to estimate $\hat{\Gamma}_\tau$ as

$$\hat{\Gamma}_\tau = \frac{1}{n} \sum_{i=\tau+1}^{n} h\left(\hat{\theta}; w_i\right) h\left(\hat{\theta}; w_{i-\tau}\right)'$$

Note that $\hat{S} \longrightarrow S$, because $\hat{\Gamma}_\tau \longrightarrow \Gamma_\tau$ as $n \longrightarrow +\infty$.

2.1 Asymptotic Distribution of GMM Estimator

- We assume that the GMM estimator has the following properties.
- Assumption

I. $\hat{\theta} \longrightarrow \theta$

II. $\sqrt{n}g(\theta;W) \longrightarrow N(0,S),\ S=\lim_{n\to\infty} V\left(\sqrt{n}g(\theta;W)\right).$

- Theorem

Asymptotic Normality of the GMM Estimator $\hat{\theta}_{GMM}$ satisfies

$$\sqrt{n}\left(\hat{\theta}-\theta\right) \longrightarrow N(0,(D'S^{-1}D)^{-1})$$

Where D is a $r\times k$ matrix, and $\widehat{D}$ is an estimator of D, defined as: $D=\frac{\partial g(\theta;W)}{\partial\theta\prime}$, $\widehat{D}=\frac{\partial g(\widehat{\theta};W)}{\partial\theta\prime}$.

- Proof of Asymptotic Normality:
- The first-order condition of GMM is:

$$\frac{\partial g(\hat{\theta};W)'}{\partial\theta}\hat{S}^{-1}g(\hat{\theta};W) = 0$$

where $\hat{\theta}$ is the estimator of GMM.

- Using the Theorem of Mean Value, linearize $g(\hat{\theta};W)$ around $\theta = \hat{\theta}$ can be written as fellows

$$g(\hat{\theta};W) = g(\theta;W) + \frac{\partial g(\bar{\theta};W)}{\partial\theta'}(\hat{\theta}-\theta) = \mathrm{g}(\theta;W) + \bar{D}(\hat{\theta}-\theta)$$

where $\bar{\theta} \in (\hat{\theta},\theta)$ and $\bar{D} = \frac{\partial g(\bar{\theta};W)}{\partial\theta\prime}$.

- Substituting the linear approximation at $\hat{\theta} = \theta$, we obtain:

$$0 = \widehat{D}'\hat{S}^{-1}\mathrm{g}(\hat{\theta}; W)$$

$$= \widehat{D}'\hat{S}^{-1}\left(g(\theta; W) + \overline{D}(\hat{\theta} - \theta)\right)$$

$$= \widehat{D}'\hat{S}^{-1}g(\theta; W) + \widehat{D}'\hat{S}^{-1}\overline{D}(\hat{\theta} - \theta)$$

Which can be written as

$$\hat{\theta} - \theta = -\left(\widehat{D}'\hat{S}^{-1}\overline{D}\right)^{-1}\widehat{D}\hat{S}^{-1}g(\theta; W)$$

- From Assumption 1, $\hat{\theta} \longrightarrow \theta$ implies $\bar{\theta} \longrightarrow \theta$. Therefore,

$$\sqrt{n}\left(\hat{\theta} - \theta\right) = -\left(\hat{D}'\hat{S}^{-1}\bar{D}\right)^{-1}\hat{D}\hat{S}^{-1}\times\sqrt{n}g(\theta;W)$$

- The GMM estimator $\hat{\theta}$ has the following asymptotic distribution:

$$\sqrt{n}\left(\hat{\theta} - \theta\right) \longrightarrow \mathrm{N}(0,(D'S^{-1}D)^{-1})$$

Note that $\hat{D} \longrightarrow D$, $\bar{D} \longrightarrow D$, $\hat{S} \longrightarrow S$ and Assumption 2 are utilized.

2.2 Testing Hypothesis:

$$H_0:\ \ R(\theta) = 0$$
$$H_1:\ \ R(\theta) \neq 0$$

where $R(\theta)$ is a $p{\times}1$ vector function for $p \leq k$.

p denotes the number of restrictions.

$R(\theta)$ is linearized as:

$$R\left(\hat{\theta}_{GMM}\right) = R(\theta) + R_{\overline{\theta}}\left(\hat{\theta}_{GMM} - \theta\right)$$

Where $R_{\overline{\theta}} = \frac{\partial R(\overline{\theta})}{\partial \theta\prime}$, which is a $p{\times}k$ matrix.

- Asymptotic distribution of $\sqrt{n}\left(R\left(\hat{\theta}_{GMM}\right) - R(\theta)\right)$ is

$$\sqrt{n}\left(R\left(\hat{\theta}_{GMM}\right) - R(\theta)\right) = R_{\overline{\theta}}\sqrt{n}\left(\hat{\theta}_{GMM} - \theta\right) \longrightarrow N(0, R_{\theta}(D'S^{-1}D)^{-1}R'_{\theta})$$

- Because $R_{\overline{\theta}} \longrightarrow R_{\theta}$ as $\hat{\theta}_{GMM} \longrightarrow \theta$. So we have following distribution.

$$n\left(R\left(\hat{\theta}_{GMM}\right) - R(\theta)\right)'(R_{\theta}(D'S^{-1}D)^{-1}R'_{\theta})^{-1}\left(R\left(\hat{\theta}_{GMM}\right) - R(\theta)\right) \longrightarrow \chi^2(p)$$

- Under $H_0: R(\theta) = 0$, the test statistic is

$$n\left(R\left(\hat{\theta}_{GMM}\right)\right)'\left(R_{\hat{\theta}_{GMM}}\left(\hat{D}'\hat{S}^{-1}\hat{D}\right)^{-1}R'_{\hat{\theta}_{GMM}}\right)^{-1}\left(R\left(\hat{\theta}_{GMM}\right)\right) \longrightarrow \chi^2(p)$$

3.3 Example of $h(\theta; w)$

1. OLS:

Regression Model:

$$y_i = x_i\beta + \epsilon_i, \qquad E(x'_i\epsilon_i) = 0$$
$$h(\theta; w_i) = x'_i(y_i - x_i\beta)$$

2. IV:

Regression Model:

$$y_i = x_i\beta + \epsilon_i, \qquad E(x'_i\epsilon_i) \neq 0, \qquad E(z'_i\epsilon_i) = 0,$$
$$h(\theta; w_i) = z'_i(y_i - x_i\beta)$$

2. NLS:

Regression Model:

$$f(y_i, x_i, \beta) = \epsilon_i, \qquad E(x'_i\epsilon_i) \neq 0, \qquad E(z'_i\epsilon_i) = 0$$
$$h(\theta; w_i) = z'_i f(y_i, x_i, \beta)$$