Agenda:

1. Estimating $\text{Avar}[\hat\beta_n]$
2. Testing linear restrictions on β_*

Recall: $\sqrt{n}(\hat\beta_n - \beta_*) \xrightarrow{d} N(0, \underbrace{\Sigma_x^{-1}\Sigma_g\Sigma_x^{-1}}_{\text{Avar}[\hat\beta_n]})$

Interpretation: $\hat\beta_n \overset{d}{\approx} N(\beta_*, \frac{1}{n}\text{Avar}[\hat\beta_n])$

$\Sigma_x = E[x_i x_i^T]$

$\Sigma_g = E[\underbrace{(x_i\epsilon_i)}_{g_i}(x_i\epsilon_i)^T] = E[\underset{1\times1}{\epsilon_i^2}\underset{p\times1}{x_i}\underset{1\times p}{x_i^T}] = E[\underset{p\times1}{x_i}\underset{1\times1}{\epsilon_i^2}\underset{1\times p}{x_i^T}]$

Assume consistent estimates of Σ_x & Σ_g: $\hat\Sigma_x \xrightarrow{p} \Sigma_x$

$\hat\Sigma_g \xrightarrow{p} \Sigma_g$

Define $\widehat{\text{Avar}}[\hat\beta_n] \doteq \hat\Sigma_x^{-1}\hat\Sigma_g\hat\Sigma_x^{-1}$

$\widehat{\text{Avar}}[\hat\beta_n]$ is a consistent estimate of $\text{Avar}[\hat\beta_n]$ (consequence of CMT)

Think of $\text{Avar}[\hat\beta_n]$ as $f(\Sigma_x, \Sigma_g)$, where $f(A,B) \doteq A^{-1}BA^{-1}$

Assume $\hat\Sigma_x, \hat\Sigma_g$ are consistent, then we wish to deduce

$f(\hat\Sigma_x, \hat\Sigma_g) \xrightarrow{p} f(\Sigma_x, \Sigma_g)$

consistent estimator of Σ_x is $\hat\Sigma_x \doteq \frac{1}{n}\sum_{i=1}^n x_i x_i^T$ (consequence of LLN)

consistent estimator of Σ_g is $\hat\Sigma_g \doteq \frac{1}{n}\sum_{i=1}^n \hat\epsilon_i^2 x_i x_i^T$, where

$\hat\epsilon_i$'s are the OLS residuals

$\hat\epsilon_i \doteq Y_i - \hat Y_i = Y_i - x_i^T\hat\beta_n$.

Intuition: $\hat\epsilon_i \approx \epsilon_i$ so $\hat\Sigma_g \approx \frac{1}{n}\sum_{i=1}^n \epsilon_i^2 x_i x_i^T \xrightarrow{p} \Sigma_g$ (consequence of LLN)

Recall: $\hat\epsilon_i = Y_i - x_i^T\hat\beta_n = x_i^T\beta_* + \epsilon_i - x_i^T\hat\beta_n = \epsilon_i + x_i^T(\beta_* - \hat\beta_n)$

$$\hat\Sigma_g = \frac{1}{n}\sum_{i=1}^n \hat\epsilon_i^2 x_i x_i^T$$

$$= \frac{1}{n}\sum_{i=1}^n (\epsilon_i + x_i^T(\beta_* - \hat\beta_n))^2 x_i x_i^T$$ (plugging in exp for $\hat\epsilon_i$)

$$= \frac{1}{n}\sum_{i=1}^n \epsilon_i^2 x_i x_i^T + \frac{2}{n}\sum_{i=1}^n x_i(\epsilon_i x_i^T(\beta_* - \hat\beta_n))x_i^T$$

$$+ \frac{1}{n}\sum_{i=1}^n x_i x_i^T(\beta_* - \hat\beta_n)(\beta_* - \hat\beta_n)^T x_i x_i^T$$

$$= \overbrace{\frac{1}{n}\sum_{i=1}^n \epsilon_i^2 x_i x_i^T}^{I} + \overbrace{2\left(\frac{1}{n}\sum_{i=1}^n x_i^T x_i \epsilon_i x_i^T\right)(\beta_* - \hat\beta_n)}^{II}$$ (trace trick)

$$\frac{1}{n}\sum_{i=1} \epsilon_i^2 x_i x_i^T + 2\left(\frac{1}{n}\sum_{i=1} x_i x_i^T \epsilon_i x_i^T\right)(\beta_* - \hat{\beta}_n)$$

$$\underbrace{+ (\beta_* - \hat{\beta}_n)^T \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T x_i x_i^T\right)(\beta_* - \hat{\beta}_n)}_{\text{III}}$$

$\text{I} \xrightarrow{p} \Sigma_g$ (by LLN)

All other terms will converge in probability to 0 because $\hat{\beta}_* - \hat{\beta}_n \xrightarrow{p} 0$

In summary, $\widehat{\text{Avar}}[\hat{\beta}_n] = \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T\right)}_{\hat{\Sigma}_x}^{-1} \underbrace{\left[\frac{1}{n}\sum_{i=1}^{n} \hat{\epsilon}_i^2 x_i x_i^T\right]}_{\hat{\Sigma}_g} \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T\right)^{-1}$

is a consistent estimator of $\text{Avar}[\hat{\beta}_n]$

Aside: $\widehat{\text{Avar}}[\hat{\beta}_n] = (\overbrace{\Sigma_x + E_x}^{\hat{\Sigma}_x})^{-1} (\Sigma_g + E_g)(\Sigma_x + E_x)^{-1}$, where $E_x \doteq \hat{\Sigma}_x - \Sigma_x$, $E_g \doteq \hat{\Sigma}_g - \Sigma_g$

CLT: $\sqrt{n}(\hat{\Sigma}_x - \Sigma_x) \xrightarrow{d} N(0, \text{some variance})$

intuitively: $\hat{\Sigma}_x - \Sigma_x = Z/\sqrt{n}$, where $Z \sim N(0, \text{some var})$

$$= \left(\Sigma_x + Z_x/\sqrt{n}\right)^{-1} \left(\Sigma_g + Z_g/\sqrt{n}\right)\left(\Sigma_x + Z_x/\sqrt{n}\right)^{-1}$$

(by Taylor's Thm)

$$\widehat{\text{Avar}}[\hat{\beta}_n] - \text{Avar}[\beta_*] = \left(\Sigma_x + Z_x/\sqrt{n}\right)^{-1}\left(\Sigma_g + Z_g/\sqrt{n}\right)\left(\Sigma_x + Z_x/\sqrt{n}\right)^{-1} - \Sigma_x^{-1}\Sigma_g\Sigma_x^{-1}$$

CMT

$$\approx \left(\Sigma_x^{-1} + Z_x'/\sqrt{n}\right)\left(\Sigma_g + Z_g/\sqrt{n}\right)\left(\Sigma_x^{-1} + Z_x'/\sqrt{n}\right) - \Sigma_x^{-1}\Sigma_g\Sigma_x^{-1}$$

$O\left(\frac{1}{\sqrt{n}}\right) \rightarrow$ $\approx Z_x'/\sqrt{n}\, \Sigma_g \Sigma_x^{-1} + \Sigma_x^{-1} Z_g/\sqrt{n}\, \Sigma_x^{-1} + \Sigma_x^{-1}\Sigma_g Z_x'/\sqrt{n}$

$O\left(\frac{1}{n}\right) \rightarrow$ $+ Z_x'/\sqrt{n} \cdot Z_g/\sqrt{n} \cdot \Sigma_x^{-1} + Z_x'/\sqrt{n}\, \Sigma_g\, Z_x'/\sqrt{n} + \Sigma_x^{-1} \frac{Z_g}{\sqrt{n}} \frac{Z_x'}{\sqrt{n}}$

$O\left(\frac{1}{n^{3/2}}\right) \rightarrow$ $+ Z_x'/\sqrt{n}\, Z_g/\sqrt{n}\, Z_x'/\sqrt{n}$

Condition (conditional homoskedasticity): $E[\epsilon_i^2 | x_i] = \sigma^2$ for all i

Under this condition, $\Sigma_g = E[\epsilon_i^2 x_i x_i^T]$

$$= E\left[E[\epsilon_i^2 | x_i]\, x_i x_i^T\right] \quad \text{(tower property)}$$

$$= E[\sigma^2 x_i x_i^T]$$

$$= \sigma^2 \Sigma_x$$

$\text{Avar}[\hat{\beta}_n] = \Sigma_x^{-1}\Sigma_g\Sigma_x^{-1} = \Sigma_x^{-1}(\sigma^2 \Sigma_x)\Sigma_x^{-1}$

$= \sigma^2 \Sigma_X^{-1}$ under conditional homoskedasticity

Recall we estimated Σ_X with $\hat{\Sigma}_X = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T$

〃 we estimate σ^2 with $s_n^2 := \frac{1}{n}\sum_{i=1}^{n} \hat{\epsilon}_i^2$

s_n^2 is consistent b/c $s_n^2 = \frac{1}{n}\sum_{i=1}^{n} \hat{\epsilon}_i^2 = \frac{1}{n}\sum_{i=1}^{n} \left(\epsilon_i + x_i^T(\beta_* - \hat{\beta}_n)\right)^2$

$$= \underbrace{\frac{1}{n}\sum_{i=1}^{n} \epsilon_i^2}_{\xrightarrow{p} \sigma^2} + \underbrace{\frac{2}{n}\sum_{i=1}^{n} \epsilon_i x_i^T(\beta_* - \hat{\beta}_n)}_{\xrightarrow{p} 0} + \underbrace{\frac{1}{n}\sum_{i=1}^{n} (\beta_* - \hat{\beta}_n) x_i x_i^T (\beta_* - \hat{\beta}_n)}_{\xrightarrow{p} 0}$$

Under conditional homoskedasticity: $\widehat{Avar}[\hat{\beta}_n] = s_n^2 \Big(\underbrace{\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T}_{\frac{1}{n}X^T X}\Big)^{-1}$

z-statistic: $$\frac{\sqrt{n}\,[\hat{\beta}_n]_j}{[\widehat{Avar}[\hat{\beta}_n]]_{j,j}^{\frac{1}{2}}} = \frac{\sqrt{n}\,[\hat{\beta}_n]_j}{[s_n^2(\frac{1}{n}X^T X)^{-1}]_{j,j}^{\frac{1}{2}}}$$

$$= \frac{\hat{\beta}_n}{s_n [(X^T X)^{-1}]_{j,j}}$$