6 Time Series Models (Ref: Chap 19 & 21)

Consider the following regression

$$y_t = bx_t + u_t, \quad t = 1, ..., T$$

where

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim iid\left(0, \sigma^2\right)$$

Let's assume that $E\left(x_t u_s\right) = E\left(x_t x_s\right) = 0$ **for all** t **and** s. Q1: Find the limiting distribution of $\hat{b}$.

We know

$$\hat{b} - b = \frac{\sum x_t u_t}{\sum x_t^2},$$

and

$$E\left(\sum x_t u_t\right) = 0.$$

Consider

$$E\left(\sum x_t u_t\right)^2.$$

Observe this

$$\left(\sum x_t u_t\right)^2 = \left(x_1 u_1 + ... + x_T u_T\right)^2 = \sum x_t^2 u_t^2 + 2\left(x_1 u_1 x_2 u_2 + ... + x_{T-1} u_{T-1} x_T u_T\right) \quad (16)$$

Now

$$E\sum x_t^2 u_t^2 = \sum E x_t^2 E u_t^2$$

To calculate Eu_t^2, consider the folllowings

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad u_{t-1} = \rho u_{t-2} + \varepsilon_{t-1}$$

so that

$$\begin{aligned} u_t &= \rho^2 u_{t-2} + \rho\varepsilon_{t-1} + \varepsilon_t \\ &= \rho^3 u_{t-3} + \rho^2 \varepsilon_{t-2} + \rho\varepsilon_{t-1} + \varepsilon_t \\ &\vdots \\ &= \varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + ... = \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j} \end{aligned}$$

Next,

$$u_t^2 = \left(\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + ...\right)^2 = \sum_{j=0}^{\infty} \rho^{2j}\varepsilon_{t-j}^2 + \text{cross product terms}$$

so that

$$Eu_t^2 = E\left(\varepsilon_t^2 + \rho^2\varepsilon_{t-1}^2 + \rho^4\varepsilon_{t-2}^2 + ...\right) + E\left(\text{cross}\right)$$

Since

$$E\varepsilon_t\varepsilon_s = 0 \text{ for } t \neq s,\ E\left(\text{cross}\right) = 0.$$

Hence we have

$$\begin{aligned} Eu_t^2 &= E\left(\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + ...\right)^2 \\ &= E\left(\varepsilon_t^2 + \rho^2\varepsilon_{t-1}^2 + \rho^4\varepsilon_{t-2}^2 + ...\right) \\ &= \sigma^2\left(1 + \rho^2 + \rho^4 + ...\right) \end{aligned}$$

Note that

$$\begin{aligned} 1 + \rho^2 + \rho^4 + ... &= \frac{1}{1-\rho^2}, \\ 1 + \rho + \rho^2 + ... &= \frac{1}{1-\rho}, \\ 1 + \rho + \rho^2 + ... + \rho^T &= \frac{1-\rho^{T+1}}{1-\rho}. \end{aligned}$$

Finally

$$Eu_t^2 = \frac{\sigma^2}{1-\rho^2} = \sigma_u^2.$$

Also note that

$$\begin{aligned} Eu_tu_{t-1} &= E\left(\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + ...\right)\left(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \rho^2\varepsilon_{t-3} + ...\right) \\ &= \sigma^2\left(\rho + \rho^3 + ...\right) = \sigma^2\rho\left(1 + \rho^2 + ...\right) = \frac{\sigma^2}{1-\rho^2}\rho \\ &= \rho\left(Eu_t^2\right), \end{aligned}$$

and

$$\begin{aligned} Eu_tu_{t-2} &= E\left(\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + ...\right)\left(\varepsilon_{t-2} + \rho\varepsilon_{t-3} + \rho^2\varepsilon_{t-4} + ...\right) \\ &= \sigma^2\left(\rho^2 + \rho^4 + ...\right) = \sigma^2\rho^2\left(1 + \rho^2 + ...\right) = \frac{\sigma^2}{1-\rho^2}\rho^2 \\ &= \rho^2\left(Eu_t^2\right). \end{aligned}$$

In general

$$Eu_t u_{t-k} = Eu_{t+k}u_t = \rho^k Eu_t^2 = \rho^k \sigma_u^2$$

Then we have

$$E\left(\sum x_t u_t\right)^2 = \sum Ex_t^2 Eu_t^2 + 2E\left(x_1u_1x_2u_2 + ... + x_{T-1}u_{T-1}x_Tu_T\right) = T\sigma_x^2\sigma_u^2$$

Hence there is no much difference.

Let's assume that $E(x_t u_s) = 0$ for all t and s but $E(x_t x_s) = \rho^{t-s}E(x_t^2)$. Then we have

$$\begin{aligned} Ex_sx_tu_su_t &= \rho^{t-s}E\left(x_t^2\right)\rho^{t-s}E\left(u_t^2\right) \\ &= \rho^{2(t-s)}\sigma_x^2\sigma_u^2 \end{aligned}$$

Consider the cross product term carefully

$$\begin{aligned} \text{Cross Term} &= x_1u_1\left(x_2u_2 + ... + x_Tu_T\right) \\ &\quad + x_2u_2\left(x_1u_1 + x_3u_3... + x_Tu_T\right) \\ &\quad + ... \\ &\quad + x_Tu_T\left(x_1u_1 + ... + x_{T-1}u_{T-1}\right) \end{aligned}$$

Hence

$$\begin{aligned} E\text{Cross Term} &= Ex_1u_1\left(x_2u_2 + ... + x_Tu_T\right) \\ &\quad + Ex_2u_2\left(x_1u_1 + x_3u_3... + x_Tu_T\right) \\ &\quad + ... \\ &\quad + Ex_Tu_T\left(x_1u_1 + ... + x_{T-1}u_{T-1}\right) \end{aligned}$$

where

$$\begin{aligned} Ex_1u_1\left(x_2u_2 + ... + x_Tu_T\right) &= \rho^2\sigma_x^2\sigma_u^2 + \rho^4\sigma_x^2\sigma_u^2 + ... + \rho^{2(T-1)}\sigma_x^2\sigma_u^2 \\ &= \sigma_x^2\sigma_u^2\rho^2\left(1 + \rho^2 + ... + \rho^{2(T-2)}\right) = \sigma_x^2\sigma_u^2\rho^2\frac{1-\rho^{2(T-1)}}{1-\rho^2}, \end{aligned}$$

$$\begin{aligned} Ex_2u_2\left(x_1u_1+x_3u_3...+x_Tu_T\right) &= \rho^2\sigma_x^2\sigma_u^2+\rho^2\sigma_x^2\sigma_u^2+...+\rho^{2(T-2)}\sigma_x^2\sigma_u^2 \\ &= \sigma_x^2\sigma_u^2\rho^2\left(2+\rho^2+...+\rho^{2(T-3)}\right) \\ &= \sigma_x^2\sigma_u^2\rho^2\frac{1-\rho^{2(T-2)}}{1-\rho^2}+\sigma_x^2\sigma_u^2\rho^2, \end{aligned}$$

$$\begin{aligned} & Ex_3u_3\left(x_1u_1+x_2u_2+x_4u_4+...+x_Tu_T\right) \\ = & \ \rho^4\sigma_x^2\sigma_u^2+\rho^2\sigma_x^2\sigma_u^2+\rho^2\sigma_x^2\sigma_u^2+...+\rho^{2(T-3)}\sigma_x^2\sigma_u^2 \\ = & \ \sigma_x^2\sigma_u^2\rho^2\left(\rho^2+2+\rho^2+...+\rho^{2(T-4)}\right) \\ = & \ \sigma_x^2\sigma_u^2\rho^2\frac{1-\rho^{2(T-3)}}{1-\rho^2}+\sigma_x^2\sigma_u^2\rho^2\left(1+\rho^2\right), \end{aligned}$$

$$Ex_Tu_T\left(x_1u_1+...+x_{T-1}u_{T-1}\right)=\rho^{2T-2}\sigma_x^2\sigma_u^2+...+\rho^2\sigma_x^2\sigma_u^2=\sigma_x^2\sigma_u^2\rho^2\frac{1-\rho^{2(T-1)}}{1-\rho^2}$$

Hence the total sum becomes

$$2\sigma_x^2\sigma_u^2\rho^2\sum_{i=1}^{T}\frac{1-\rho^{2(T-i)}}{1-\rho^2}=2\frac{\sigma_x^2\sigma_u^2\rho^2}{1-\rho^2}T+O\left(1\right)$$

Or

$$\begin{aligned} E\left(\sum x_tu_t\right)^2 &= \sum Ex_t^2Eu_t^2+2E\left(x_1u_1x_2u_2+...+x_{T-1}u_{T-1}x_Tu_T\right) \\ &= T\sigma_x^2\sigma_u^2+2\frac{\sigma_x^2\sigma_u^2\rho^2}{1-\rho^2}T+O\left(1\right) \\ &= T\sigma_x^2\sigma_u^2\left(\frac{1+\rho^2}{1-\rho^2}\right)+O\left(1\right) \end{aligned}$$

Finally we have

$$\sqrt{T}\left(\hat{b}-b\right)\to^d N\left(0,\omega_b^2\right)$$

where

$$\omega_b^2=\frac{\sigma_x^2\sigma_u^2\left(\frac{1+\rho^2}{1-\rho^2}\right)}{\sigma_x^2\sigma_x^2}=\left(\frac{1+\rho^2}{1-\rho^2}\right)\frac{\sigma_u^2}{\sigma_x^2}\geq\frac{\sigma_u^2}{\sigma_x^2}$$

In other words, the typical limiting distribution such as

$$\sqrt{T}\left(\hat{b}-b\right)\to^d N\left(0,\sigma_u^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}\right)$$

does not work here.

6.1 Definitions

1. Strong Stationary: A time-series process, $\{z_t\}_{t=-\infty}^{t=\infty}$ is stronlgy stationary if the joint probability distribution of any set of k obersvations in the sequence $\{z_t, ..., z_{t+k}\}$ is the same regardless of the origin t, in the time scale.

2. Weak Stationary: $\{z_t\}$ is weakly stationary if (i) $E(z_t)$ is finite, (ii) $Cov(z_t, z_{t-k})$ is a finite function only of k and model parameters. (In other words, it should not be time varying)

3. Ergodicity: A strongly stationary time series process is ergodic if

$$\begin{aligned} & \lim_{k\to\infty} \left|E\left[f(z_t, z_{t+1}, ..., z_{t+a})\, g(z_{t+k}, z_{t+k+1}, ..., z_{t+k+b})\right]\right| \\ = & \left|Ef(z_t, z_{t+1}, ..., z_{t+a})\right| \left|Eg(z_{t+k}, z_{t+k+1}, ..., z_{t+k+b})\right| \end{aligned}$$

   (a) Example: Let $z_t = \rho z_{t-1} + u_t,\ \ u_t \sim iid(0,1)$

$$\lim_{k\to\infty} \left|E(z_t z_{t+k})\right| = \lim_{k\to\infty} \left|\rho^k \sigma_z^2\right| = 0 = |Ez_t|\,|Ez_{t+k}|$$

4. The Ergodic Theorem: If z_t is strongly stationary and ergodic and $E|z_t|$ is a finite constant, then $\bar{z}_T = T^{-1}\sum z_t \to^{a.s.} \mu = E(z_t)$.

5. Martingale Sequence: z_t is a martingale sequence if

$$E(z_t|z_{t-1}, z_{t-2}, ...) = z_{t-1}$$

   (a) Example: $z_t = z_{t-1} + u_t,\ \ E(z_t|z_{t-1}, z_{t-2}, ...) = z_{t-1}$

6. Martingale Difference Sequence: z_t is a martingale difference sequence if

$$E(z_t|z_{t-1}, z_{t-2}, ...) = 0$$

7. White Noise process: stationary but not-autocorrelated process.

6.2 Long Run Variance

Q1: Consider $u_t = \rho u_{t-1} + e_t,\;\; e_t$ is a white noise process with a finite variance of σ^2. Find the limiting distribution of the sample mean of u_t.

$$\mu_T = \frac{1}{T}\sum_{t=1}^{T} u_t$$

Mean: $E\mu_T = 0$.

Variance:

$$\begin{aligned}
E\left(\frac{1}{T}\sum_{t=1}^{T} u_t\right)^2 &= \frac{1}{T^2}E\left(u_1+...+u_T\right)\left(u_1+...+u_T\right) = \frac{1}{T^2}\left[\sum_{t=1}^{T} Eu_t^2 + 2E\sum_{t=1}^{T-1}\sum_{s=t}^{T} u_t u_s\right] \\
&= \frac{1}{T^2}\left[T\sigma_u^2 + 2\sum_{t=1}^{T-1}\sum_{s=t}^{T}\rho^{t-s}\sigma_u^2\right] = \frac{1}{T^2}\sigma_u^2\left[T + 2\frac{\rho}{1-\rho}T + O\left(1\right)\right] \\
&= \frac{1}{T}\frac{\sigma^2}{1-\rho^2}\left(1 + 2\frac{\rho}{1-\rho} + O\left(T^{-1}\right)\right) = \frac{1}{T}\frac{\sigma^2}{1-\rho^2}\left(\frac{1+\rho}{1-\rho} + O\left(T^{-1}\right)\right) \\
&= \frac{1}{T}\frac{\sigma^2}{\left(1-\rho\right)^2} + O\left(T^{-2}\right) := \frac{1}{T}\omega^2 \quad \text{omega}
\end{aligned}$$

Hence we have

$$d\mu_T \to^d N\left(0, \frac{1}{T}\omega^2\right)$$

or

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t \to^d N\left(0, \frac{\sigma^2}{\left(1-\rho\right)^2}\right) \tag{17}$$

We call ω^2 long run variance of u_t.

6.3 Estimation of Long Run Variance (HAC Estimation)

How to estimate the long run variance of u_t in (17) then? The unknowns are σ^2 and ρ. How many observations do we have? T. So it is easy to estimate it.

Now what if the parametric structure is unknown. Let say u_t follows $AR(T)$ or $ARMA(p,q)$ where p and q are unknown? Is it possible to estimate ω^2? No. The total number of unknowns becomes $\frac{T(T-1)}{2}+1$. The first term is the sum of cross product terms and the last term, 1, is the unknown variance term (diagonal term). If variance is time varying, then it becomes $\frac{T(T-1)}{2}+T$. Simply impossible to estimate the long run variance in this case.

Therefore we are imposing regularity: Ergodic and stationary process. And then we assume that

$$E(u_t u_{t-k}) \simeq 0 \text{ for a large } k.$$

Alternatively let say

$$\begin{aligned} E\frac{1}{T}\sum_{t=1}^{T-1}\sum_{s=t}^{T} u_t u_s &= E\frac{1}{T}\sum_{t=1}^{T-1}\sum_{s=t}^{t+k} u_t u_s + E\frac{1}{T}\sum_{t=1}^{T-1}\sum_{s=t+k+1}^{T} u_t u_s \\ &= E\frac{1}{T}\sum_{t=1}^{T-1}\sum_{s=t}^{t+k} u_t u_s + o_p(1) \end{aligned} \quad (18)$$

In this case, we don't need to estimate the second term.

Newey and West Estimator Let

$$\omega^2 = \omega_0^2 + \sum_{j=1}^{\infty}\left(\omega_j^2 + \omega_{-j}^2\right)$$

where

$$\omega_j^2 = E u_t u_{t-j}$$

Then we can apply the above concept in (18), so we have

$$\hat{\omega}^2 = \hat{\omega}_0^2 + \sum_{j=1}^{k}\left(\hat{\omega}_j^2 + \hat{\omega}_{-j}^2\right)$$

According to Andrews (1991), we can modify the estimator further in an elegant way

$$\hat{\omega}^2 = \hat{\omega}_0^2 + \sum_{j=1}^{k} w_j \left(\hat{\omega}_j^2 + \hat{\omega}_{-j}^2\right)$$

where w_j is some optimal weight. Newey and West (1992) suggest

$$w_j = 1 - \frac{j}{k+1}, \quad k = int\left(T^{1/3}\right).$$

We call such weight Bartlett kernel weight. They show that this type of estimator becomes consistent.

Parametric Version: Andrews and Monahan's Prewhitening HAC estimator Let e_t is a stationary and ergodic process. Then we may have

$$u_t = \rho u_{t-1} + e_t$$

and

$$E\left(\frac{1}{\sqrt{T}}\sum u_t\right)^2 = \frac{\omega_e^2}{(1-\rho)^2}$$

where ω_e^2 is the long run variance of e_t. Now we estimate $\hat{\rho}$ and replace this. That is,

$$\hat{\omega}_u^2 = \frac{\hat{\omega}_e^2}{(1-\hat{\rho})^2}.$$

Conversion to Matrix Form Consider

$$y_t = \mathbf{X}_t'\mathbf{b} + u_t$$

$$\sqrt{T}\left(\hat{\mathbf{b}} - \mathbf{b}\right) \to^d N\left(0, \mathbf{V}_{\hat{b}}\right)$$

where

$$\mathbf{V}_{\hat{b}} = \left(\frac{1}{T}\sum \mathbf{X}_t\mathbf{X}_t'\right)^{-1} \frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T} Eu_s\mathbf{X}_s\left(u_t\mathbf{X}_t\right)' \left(\frac{1}{T}\sum \mathbf{X}_t\mathbf{X}_t'\right)^{-1}.$$

Now let

$$\boldsymbol{\xi}_t = u_t \cdot \mathbf{X}_t = \left(u_t x_{1t}, u_t x_{2t}, ..., u_t x_{kt}\right)$$

Then

$$\Omega^2 = \Omega_0 + \Omega_j + \Omega_{-j}$$

$$\hat{\Omega}^2 = \hat{\Omega}_0 + \sum_{j=1}^{k} w_j \left(\hat{\Omega}_j + \hat{\Omega}_{-j}\right)$$

Then we have

$$\hat{\mathbf{V}}_{\hat{b}} = \left(\frac{1}{T}\sum \mathbf{X}_t\mathbf{X}_t'\right)^{-1} \hat{\Omega}^2 \left(\frac{1}{T}\sum \mathbf{X}_t\mathbf{X}_t'\right)^{-1}.$$

Alternative Approach Let assume

$$y_t = \mathbf{X}_t'\mathbf{b} + u_t, \quad u_t = \sum_{j=1}^{p} \rho_j u_{t-j} + e_t$$

Then

$$\begin{aligned} \rho_1 y_{t-1} &= \rho_1 \mathbf{X}_{t-1}'\mathbf{b} + \rho_1 u_{t-1} \\ &\vdots \\ \rho_p y_{t-p} &= \rho_p \mathbf{X}_{t-p}'\mathbf{b} + \rho_p u_{t-p} \end{aligned}$$

Now subtract $\rho y_{t-1}, ...\rho_p y_{t-p}$ from y_t.

$$\begin{aligned} y_t &= \mathbf{X}_t'\mathbf{b} - \sum_{j=1}^{p} \rho_j \mathbf{X}_{t-j}'\mathbf{b} + \sum_{j=1}^{p} \rho_j y_{t-j} + \mathbf{b}u_t - \sum_{j=1}^{p} \rho_j u_{t-j} \\ &= \mathbf{X}_t'\mathbf{b} - \sum_{j=1}^{p} \rho_j \mathbf{X}_{t-j}'\mathbf{b} + \sum_{j=1}^{p} \rho_j y_{t-j} + e_t = \mathbf{Z}_t\boldsymbol{\gamma} + e_t \end{aligned}$$

where $Z_t = \left(\mathbf{X}_t, \mathbf{X}_{t-1}, ..., \mathbf{X}_{t-p}, y_{t-1}, ..., y_{t-p}\right)$. Let rewrite it as

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\gamma} + \mathbf{e},$$

and then we have

$$\sqrt{T}\left(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}\right) \to^d N\left(0, \sigma_e^2 Q_Z^{-1}\right).$$

where

$$Q_Z = \text{plim}_{T\to\infty} \frac{\mathbf{Z}'\mathbf{Z}}{T}$$

Conventional Approach (Generalized Least Squares GLS: Chapter 8) Suppose that we know the AR order. Let say AR(1). Then we have

$$u_t = \rho u_{t-1} + e_t$$

so that

$$\begin{aligned} E\mathbf{uu}' &= \Omega_{T\times T} = \sigma_e^2 \begin{bmatrix} \frac{1}{1-\rho^2} & \frac{\rho}{1-\rho^2} & \cdots & \frac{\rho^{T-1}}{1-\rho^2} \\ \frac{\rho}{1-\rho^2} & \frac{1}{1-\rho^2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \frac{\rho}{1-\rho^2} \\ \frac{\rho^{T-1}}{1-\rho^2} & \cdots & \frac{\rho}{1-\rho^2} & \frac{1}{1-\rho^2} \end{bmatrix} \\ &= \frac{\sigma_e^2}{1-\rho^2} \begin{bmatrix} 1 & \rho & \cdots & \rho^{T-1} \\ \rho & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \rho \\ \rho^{T-1} & \cdots & \rho & 1 \end{bmatrix} \end{aligned}$$

Now we know

$$\Omega = \mathbf{C\Lambda C}'$$

where $\mathbf{C}'\mathbf{C} = \mathbf{I}$, and

$$\begin{aligned} \Omega^{-1} &= \mathbf{C\Lambda}^{-1}\mathbf{C}' \\ &= \mathbf{P}'\mathbf{P} \end{aligned}$$

where

$$\mathbf{P} = \mathbf{\Lambda}^{-1/2}\mathbf{C}'$$

Next consider the following transformation

$$\mathbf{Py} = \mathbf{PXb} + \mathbf{Pu}$$

or

$$\mathbf{y}^* = \mathbf{X}^*\mathbf{b} + \mathbf{u}^* \tag{19}$$

Now define the GLS estimator

$$\hat{\mathbf{b}}_{gls} = \left(\mathbf{X}^{*\prime}\mathbf{X}^*\right)^{-1}\mathbf{X}^{*\prime}\mathbf{y}^*$$

or alternatively we can say

$$\mathbf{X}^{*\prime}\mathbf{X}^{*} = \mathbf{X}'\mathbf{P}'\mathbf{P}\mathbf{X} = \mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}$$

and

$$\mathbf{X}^{*\prime}\mathbf{y}^{*} = \mathbf{X}'\mathbf{P}'\mathbf{P}\mathbf{y} = \mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{y}$$

so that we have

$$\hat{\mathbf{b}}_{gls} = \left(\mathbf{X}'\Omega^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\Omega^{-1}\mathbf{y}$$

and find its limiting distribution.

First note that

$$E\mathbf{u}^{*}\mathbf{u}^{*\prime} = \mathbf{P}E\mathbf{u}\mathbf{u}'\mathbf{P}' = \mathbf{P}\boldsymbol{\Omega}\mathbf{P}' = \boldsymbol{\Lambda}^{-1/2}\mathbf{C}'\mathbf{C}\boldsymbol{\Lambda}\mathbf{C}'\mathbf{C}\boldsymbol{\Lambda}^{-1/2} = \mathbf{I}$$

Hence the limiting distribution of $\hat{\mathbf{b}}_{gls}$ is given by

$$\sqrt{n}\left(\hat{\mathbf{b}}_{gls} - \mathbf{b}\right) \to^{d} N\left(0, \left(\frac{\mathbf{X}^{*\prime}\mathbf{X}^{*}}{n}\right)^{-1}\right)$$

or

$$\sqrt{n}\left(\hat{\mathbf{b}}_{gls} - \mathbf{b}\right) \to^{d} N\left(0, \left(\frac{\mathbf{X}'\Omega^{-1}\mathbf{X}}{n}\right)^{-1}\right)$$

Feasible GLS Replace Ω by $\hat{\Omega}$.

$$\hat{\mathbf{b}}_{fgls} = \left(\mathbf{X}'\hat{\Omega}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\hat{\Omega}^{-1}\mathbf{y}$$

7 Heteroskedasticity (Chapter 8 Continue)

Now we allow heterogenous variance for each i or t. That is,

$$Eu_i^2 = \sigma_i^2 \neq \sigma_j^2 = Eu_j^2$$

However we assume that

$$Eu_i u_j = 0.$$

Then we have

$$E\mathbf{uu}' = \Omega_{T\times T} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n^2 \end{bmatrix}$$

Note that

$$\mathbf{X}'E\mathbf{uu}'\mathbf{X} = \mathbf{X}'\boldsymbol{\Omega}\mathbf{X} \neq \mathbf{X}'\mathbf{X}.$$

But in this case, we have

$$\begin{aligned} [\mathbf{X}_1, ..., \mathbf{X}_n] \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n^2 \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \\ \mathbf{X}_n \end{bmatrix} &= \sigma_1^2 \mathbf{X}_1'\mathbf{X}_1 + \sigma_2^2 \mathbf{X}_2'\mathbf{X}_2 + ... + \sigma_n^2 \mathbf{X}_n'\mathbf{X}_n \\ &= \sum_{i=1}^n \sigma_i^2 \mathbf{X}_i'\mathbf{X}_i \end{aligned}$$

Therefore we have

$$\sqrt{n}\left(\hat{\mathbf{b}} - \mathbf{b}\right) \to^d N(0, \mathbf{V}_b)$$

where

$$\begin{aligned} \mathbf{V}_b &= (\mathbf{X}'\mathbf{X})^{-1} (\mathbf{X}'\Omega\mathbf{X})^{-1} (\mathbf{X}'\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{X})^{-1} \left(\sum_{i=1}^n \sigma_i^2 \mathbf{X}_i'\mathbf{X}_i\right) (\mathbf{X}'\mathbf{X})^{-1} \end{aligned}$$

If we replace σ_i^2 but $\hat{\sigma}_i^2 = \hat{u}_i^2$, then we call this estimator 'White' heteroskedasticity consistent estimator.

8 Instrumental Variables

Consider the following data generating process

$$y_i = ax_i + u_i$$

where

$$u_i = \beta x_i + e_i$$

we assume that $E(x_i e_j) = 0$ for all i and j.

Now we have

$$\hat{a} = a + (x'x)^{-1} x'u = a + \beta + (x'x)^{-1} x'e$$

so that

$$E(\hat{a} - a | x) = \beta \neq 0.$$

We say that x is endogeneous in this case. Note that the concept of endogeneity is in general somewhat different. We will explain it later in Chapter 20.

8.1 Can we know if $\beta = 0$ or not?

1. Hausman Test: testing for exogeneity. We will study it later.

2. u is unknown. How do we know if x is correlated with u ?

3. Known case: Lagged dependent variable.

$$y_t = a + y_t^o, \ y_t^o = \rho y_{t-1}^o + u_t$$

so that

$$y_t = a(1-\rho) + \rho y_{t-1} + u_t$$

Then we can rewrite it as

$$\tilde{y}_t = \rho \tilde{y}_{t-1} + \tilde{u}_t.$$

Note that $E(\tilde{y}_{t-1}\tilde{u}_t) \neq 0$. However as $t \to \infty$, this bias goes away at the $O_p(T^{-1})$ rate.

4. Measurement error: True model

$$y_i = \alpha x_i + u_i \tag{20}$$

But we observe $x_i^* = x_i + e_i$. So you run

$$y_i = \alpha x_i^* + v_i.$$

From (20), we haev

$$y_i = \alpha (x_i + e_i) - \alpha e_i + u_i = \alpha x_i^* + v_i$$

Now $E(v_i x_i^*) = E(u_i - \alpha e_i)(x_i + e_i) \neq 0.$

8.2 Solution I

Including control variables.

$$y_i = \alpha x_i + \mathbf{w}_i' \boldsymbol{\gamma} + v_i$$

where $\mathbf{w}_i = (w_{1i}, ..., w_{ki})'$. Now $\mathbf{w}_i$ becomes a proxy variable for u_i.

Problem: We don't know how many control variables should be included.

8.3 Solution II

Construct instrumental variable, z_i such that

$$E(x_i z_i) \neq 0$$

but

$$E(z_i u_i) = 0.$$

Then construct IV estimator

$$\begin{aligned} \hat{\alpha}_{IV} &= (\mathbf{z}'\mathbf{x})^{-1} \mathbf{z}'\mathbf{y} \\ &= (\mathbf{z}'\mathbf{x})^{-1} \mathbf{z}' (\mathbf{x}\boldsymbol{\alpha} + \mathbf{u}) \\ &= \alpha + (\mathbf{z}'\mathbf{x})^{-1} \mathbf{z}'\mathbf{u} \end{aligned}$$

Next,

$$\hat{\alpha}_{IV} - \alpha = (\mathbf{z}'\mathbf{x})^{-1} \mathbf{z}'\mathbf{u}$$

and

$$\begin{aligned} \text{plim}\,(\hat{\alpha}_{IV} - \alpha) &= \text{plim}\left(\frac{\mathbf{z}'\mathbf{x}}{n}\right)^{-1} \text{plim}\frac{\mathbf{z}'\mathbf{u}}{n} \\ &= Q_{zx} \cdot 0 = 0 \end{aligned}$$

so that $\hat{\alpha}_{IV}$ is a consistent estimator of α.

Asymptotic variance:

$$E\,(\hat{\alpha}_{IV} - \alpha)\,(\hat{\alpha}_{IV} - \alpha)' = E\left[(\mathbf{z}'\mathbf{x})^{-1}\,\mathbf{z}'\mathbf{u}\mathbf{u}'\mathbf{z}\,(\mathbf{z}'\mathbf{x})^{-1}\right]$$

If z and x are non-stochastic, we have

$$E\,(\hat{\alpha}_{IV} - \alpha)\,(\hat{\alpha}_{IV} - \alpha)' = (z'x)^{-1}\,z'\Omega_u z\,(z'x)^{-1}$$

8.3.1 Getting into details: Measurement error

$$y_i = \alpha x_i^* + v_i, \quad x_i^* = x_i + e_i, \quad v_i = -\alpha e_i + u_i$$

Find a variable such that

$$z_i = \beta x_i + m_i$$

but

$$Em_i e_i = 0 \text{ and } Em_i u_i = 0.$$

Then z_i is the right instrumental variable.

How to find such a good IV then? Ask GOD.