Homework 4

Haixiang Zhu

March 27, 2021

1. Consider GLS model

$$\begin{cases} \hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y \\ Var(\hat{\beta}_{GLS}) = (X'\Omega^{-1}X)^{-1} \end{cases}$$

Proof.

$$E(\hat{\beta}_{GLS}) = E[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}(X\beta + u)] = \beta$$

Let b be an alternative linear unbiased estimator such that

$$b = [(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} + A]y$$

Unbiasedness implies that $AX = 0$

$$\begin{aligned} Var(b) &= [(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} + A]\Omega[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} + A]' \\ &= (X'\Omega^{-1}X)^{-1} + A\Omega A' + (X'\Omega^{-1}X)^{-1}X'A' + AX(X'\Omega^{-1}X)^{-1} \\ &= (X'\Omega^{-1}X)^{-1} + A\Omega A' \end{aligned}$$

The second term is positive semi-definite, so $A = 0$ is best. $\square$

2. Statistical properties of double sampling

Let $y^* = \begin{pmatrix} y \\ y \end{pmatrix}, \beta^* = \begin{pmatrix} \beta \\ \beta \end{pmatrix}, X^* = \begin{pmatrix} X \\ X \end{pmatrix}, u^* = \begin{pmatrix} u \\ u \end{pmatrix}$. Then

$$\begin{aligned} y^* &= X^*\beta^* + u^* \\ \hat{\beta}^* &= \left(X^{*\prime}X^*\right)^{-1} X^{*\prime}y^* \\ \Omega &= Eu^*u^{*\prime} \\ &= \sigma^2 \begin{pmatrix} I_n & I_n \\ I_n & I_n \end{pmatrix} \end{aligned}$$

(a) Unbiasedness of $\hat{\beta}^*$

$$E(\hat{\beta}^*) = E[(X^{*\prime}X^*)^{-1}X^{*\prime}(X^*\beta^* + u^*)] = \beta^*$$

(b) Variance of $\hat{\beta}^*$ becomes larger

$$\begin{aligned}
Var(\hat{\beta}^*) &= E[(X^{*\prime}X^*)^{-1}X^{*\prime}u^*u^{*\prime}X^*(X^{*\prime}X^*)^{-1}] \\
&= \sigma^2(X^{*\prime}X^*)^{-1}X^{*\prime}\begin{pmatrix} I_n & I_n \\ I_n & I_n \end{pmatrix}X^*(X^{*\prime}X^*)^{-1} \\
&\geq \sigma^2(X^{*\prime}X^*)^{-1}
\end{aligned}$$

3. WLS

(a) Let $\alpha_i = p(X_i)$. Since $\sum_{i=1}^{N} p(X_i) = 1$

$$E(\hat{\mu}_y) = \sum_{i=1}^{N} p(X_i)E(y_i) = \mu_y$$

$$Var(\hat{\mu}_y) = \sum_{i=1}^{N} p^2(X_i)Var(y_i)$$

(b) Incorporating WLS in GLS framework to get a more efficient estimator. Because of the heteroskedasticity, we should want to have small weights where the noise variance is large, because there the data tends to be far from the true regression. Let $w_i = p(X_i) \propto \dfrac{1}{\sigma_i^2}$.

$$\hat{\beta}_{WLS} = (X'w^{-1}X)^{-1}X'w^{-1}y$$

4. Consider the causal ARMA(2,1) process

$$u_t = \sum_{i=1}^{2} \rho_i u_{t-1} + a_1\varepsilon_{t-1} + a_0\varepsilon_t$$

where $\varepsilon \sim iid(0, \sigma_\varepsilon^2)$. By the causality property, the process can be written as

$$u_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

where ψ_j denotes the ψ-weights.

$$\psi_j = a_j + \sum_{k=1}^{2} \rho_k \psi_{j-k}, \quad j = 0, 1, 2, \ldots$$

where $a_0 = 1, a_j = 0$ for $j > 1$ and $\psi_j = 0$ for $j < 0$

(a) Let $\gamma(k) = E(u_t u_{t-k})$. The autocovariance function is

$$\gamma(k) = \begin{cases} \rho_1\gamma(k-1) + \rho_2\gamma(k-2) & k \geq 2 \\ \rho_1\gamma(k-1) + \rho_2\gamma(k-2) + \sigma_\varepsilon^2 \sum_{j=k}^{1} a_j \psi_{j-k} & 0 \leq k < 2 \end{cases}$$

where

$$\begin{cases} \gamma(0) = \rho_1\gamma(1) + \rho_2\gamma(2) + (a_0^2 + a_1\rho_1 a_0 + a_1^2)\sigma_\varepsilon^2 \\ \gamma(1) = \rho_1\gamma(0) + \rho_2\gamma(1) + a_0 a_1 \sigma_\varepsilon^2 \end{cases}$$

(b) Let $z_t - \theta z_{t-1} = u_t$. Then

$$\begin{cases} z_{t-1} - \theta z_{t-2} = u_{t-1} \\ z_{t-2} - \theta z_{t-3} = u_{t-2} \end{cases}$$

Plugging back in

$$\begin{aligned} & z_t - \theta z_{t-1} - \rho_1(z_{t-1} - \theta z_{t-2}) - \rho_2(z_{t-2} - \theta z_{t-3}) = a_0\varepsilon_t + a_1\varepsilon_{t-1} \\ \Rightarrow & z_t - (\rho_1 + \theta)z_{t-1} - (\rho_2 - \theta\rho_1)z_{t-2} + \theta\rho_2 z_{t-3} = a_0\varepsilon_t + a_1\varepsilon_{t-1} \end{aligned}$$

Therefore, z_t follows a ARMA(3,1) process.