Econometrics 710 - Problem Set # 4

February 19, 2015

TA: Cristian Hernandez
cahernandez3@wisc.edu

Exercise 3.14

$$
\begin{aligned}
\hat{\beta}_{n+1} &= \left(\begin{bmatrix} X_n \\ x'_{n+1} \end{bmatrix}' \begin{bmatrix} X_n \\ x'_{n+1} \end{bmatrix} \right)^{-1} \begin{bmatrix} X_n \\ x'_{n+1} \end{bmatrix}' \begin{bmatrix} y_n \\ y_{n+1} \end{bmatrix} \\
&= (X'_n X_n + x'_{n+1} x_{n+1})^{-1} (X_n{}' y_n + x_{n+1} y_{n+1}) \\
&= [(X_n{}'X_n)^{-1} - (X_n{}'X_n)^{-1} x_{n+1} (1 + x'_{n+1} (X_n{}'X_n)^{-1} x_{n+1})^{-1} x'_{n+1} (X_n{}'X_n)^{-1}] \cdot \\
&\qquad (X_n{}' y_n + x_{n+1} y_{n+1}) \\
&\quad (\because \text{By using Woodbury matrix identity (A.2)}) \\
&= (X_n{}'X_n)^{-1} X_n{}' y_n + (X_n{}'X_n)^{-1} x_{n+1} y_{n+1} - \frac{(X_n{}'X_n)^{-1} x_{n+1} x'_{n+1} (X_n{}'X_n)^{-1}}{(1 + x'_{n+1} (X_n{}'X_n)^{-1} x_{n+1})} \cdot \\
&\qquad (X_n{}' y_n + x_{n+1} y_{n+1}) \\
&= \hat{\beta}_n + \frac{1}{(1 + x'_{n+1} (X_n{}'X_n)^{-1} x_{n+1})} (X_n{}'X_n)^{-1} x_{n+1} (y_{n+1} - x'_{n+1} \hat{\beta}_n)
\end{aligned}
$$

Exercise 3.18

From (3.37) $\hat{\beta}_{(-i)} = \hat{\beta} - (1 - h_{ii})^{-1} (X'X)^{-1} x_i \hat{e}_i$. Thus $\hat{\beta}_{-i} = \hat{\beta}$ when $x_i = 0$ or $\hat{e}_i = 0$. In words, if the observation y_i is exactly on the regression line (so that the predicted value $(x'_i \hat{\beta})$ and the actual value(y_i) are same), then excluding the observation does not change the coefficient. Excluding an observation which has $x_i = 0$ also does not change the OLS coefficient.

Exercise 3.21

Table 1 summarizes the regression results for the sub-sample of white-male-hispanics with regional and marital status dummies.

Exercise 4.1

$\mathbb{E}(\mathbf{x}_i \mathbf{x}'_i)$ is the population mean of the matrix $\mathbf{x}_i \mathbf{x}'_i$, and $\frac{1}{n} \sum_i^n \mathbf{x}_i \mathbf{x}'_i$ is its sample analogs of the unknown population mean. As sample size n increase, sample mean $\frac{1}{n} \sum_i^n \mathbf{x}_i \mathbf{x}'_i$ converges in probability to $\mathbb{E}(\mathbf{x}_i \mathbf{x}'_i)$ by the WLLN under $\mathbb{E}||\mathbf{x}_i \mathbf{x}'_i|| < \infty$

Exercise 4.2

False. Let $\mathbf{x} = (x_1, x_2, \cdots, x_n)', \hat{e} = (\hat{e}_1, \cdots, \hat{e}_n)', \mathbf{x}^2 = (x_1^2, \cdots, x_n^2)'$. For OLS residual vector $\hat{e}$, $\mathbf{x}$ and $\hat{e}$ are orthogonal, i.e. $\mathbf{x}'\hat{e} = 0$. In general, this does not imply $\mathbf{x}^{2'} \hat{e} \neq 0$ unless $\mathbf{x}$ and $\mathbf{x}^2$ are linearly dependent, i.e., $\mathbf{x}$ are constant vectors $(1, \cdots, 1)'$

Table 1: OLS Estimates Result for wage regression: All white male hispanics ($n = 4230$)

Variables	ln(Wage per hours)
Education	0.088
Experience	0.028
Experience squared/100	-0.037
Regional dummy	
Northeast	0.062
South	-0.068
West	0.020
Marital Status dummy	
Married	0.178
Widowed	0.243
Divorced	0.079
Separated	0.017
Constant	1.192
R^2	0.249
observations	4230

Counter example: $n = 3, \{(y_i, x_i)\} = \{(7,2),(4,3),(8,4)\}$. $\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2} = 2, \hat{e}_1 = 3, \hat{e}_2 = -2, \hat{e}_3 = 0$. Here $\sum x_i \hat{e}_i = 0$, but $\sum \mathbf{x}_i^2 \hat{e}_i = -6$

Exercise 4.3

Proof is similar to the proof of theorem 4.6.1.1 in the lecture notes. Clearly, $\tilde{\beta}$ is linear in $\mathbf{y}$, and unbiased estimator of β since $\mathbb{E}(\tilde{\beta}|X) = (X'D^{-1}X)^{-1}X'D^{-1}\mathbb{E}(\mathbf{y}|X) = \beta$. The variance of $\tilde{\beta}$ is

$$\begin{aligned}\mathrm{var}(\tilde{\beta}|X) &= \mathrm{var}((X'D^{-1}X)^{-1}X'D^{-1}y|X) \\ &= (X'D^{-1}X)^{-1}X'D^{-1}DD^{-1}X(X'D^{-1}X)^{-1} = (X'D^{-1}X)^{-1}\end{aligned}$$

For any linear estimator $\bar{\beta} = A'\mathbf{y}$, let A be any $n \times k$ function of X such that $A'X = I_k$. $\bar{\beta}$ unbiased if (and only if) $A'X = I_k$, since $\mathbb{E}(\bar{\beta}|X) = A'\mathbb{E}(y|X) = A'X\beta$. The variance of $\bar{\beta}$ is

$$\mathrm{var}(\bar{\beta}|X) = A'\mathrm{var}(\mathbf{y}|X)A = A'DA,$$

where $D = \mathbb{E}(\mathbf{ee}'|X) = \mathrm{diag}(\sigma_1^2, \cdots, \sigma_n^2)$.

To show that $\tilde{\beta}$ is best linear unbiased estimator, it is sufficient to show that the difference $\mathrm{var}(\bar{\beta}|X) - \mathrm{var}(\tilde{\beta}|X)$ is positive semi-definite. Set $C = A - D^{-1}X(X'D^{-1}X)^{-1}$. Note that $X'C = \mathbf{0}$. Then

$$\begin{aligned}\mathrm{var}(\bar{\beta}|X) - \mathrm{var}(\tilde{\beta}|X) &= A'DA - (X'D^{-1}X)^{-1} \\ &= (C + D^{-1}X(X'D^{-1}X)^{-1})'D(C + D^{-1}X(X'D^{-1}X)^{-1}) - (X'D^{-1}X)^{-1} \\ &= C'DC + C'DD^{-1}X(X'D^{-1}X)^{-1} + (X'D^{-1}X)^{-1}X'D^{-1}DC \\ &+ (X'D^{-1}X)^{-1}X'D^{-1}DD^{-1}X(X'D^{-1}X)^{-1} - (X'D^{-1}X)^{-1} \\ &= C'DC \\ &= (D^{1/2}C)'(D^{1/2}C)\end{aligned}$$

where $D^{1/2} = \text{diag}(\sigma_1, \cdots, \sigma_n)$. The last term is positive semi-definite (see appendix (A.8)).

Exercise 4.4

(a) $\mathbb{E}(\tilde{\beta}|X) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\mathbb{E}(\mathbf{y}|X) = \beta$

(b) $\text{var}(\tilde{\beta}|X) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\text{var}(\mathbf{y}|X)\Omega^{-1}X(X'\Omega^{-1}X)^{-1} = \sigma^2(X'\Omega^{-1}X)^{-1}$. The last equality holds since $\text{var}(\mathbf{y}|X) = \text{var}(\mathbf{e}|X) = \sigma^2\Omega$.

(c)

$$\begin{aligned}
\widehat{\mathbf{e}} &= \mathbf{y} - X\widetilde{\beta} \\
&= y - X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\mathbf{y} \\
&= M_1\mathbf{y} \\
&= M_1(X\beta + \mathbf{e}) \\
&= M_1\mathbf{e}
\end{aligned}$$

The last equality holds since $M_1X = X - X(X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}X) = \mathbf{0}$

(d)

$$\begin{aligned}
M_1'\Omega^{-1}M_1 &= (I - X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1})'\Omega^{-1}(I - X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}) \\
&= \Omega^{-1} - 2\Omega^{-1}X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} + \Omega^{-1}X(X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}X)X'\Omega^{-1} \\
&= \Omega^{-1} - \Omega^{-1}X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}
\end{aligned}$$

(e) Note that $\mathbb{E}(s^2|X) = \frac{1}{n-k}\mathbb{E}(\hat{\mathbf{e}}'\Omega^{-1}\hat{\mathbf{e}}|X) = \frac{1}{n-k}\mathbb{E}(e'M_1\Omega^{-1}M_1e|X)$ by using (c).

$$\begin{aligned}
\mathbb{E}(e'M_1\Omega^{-1}M_1e|X) &= \mathbb{E}\ (\text{tr}(e'M_1\Omega^{-1}M_1e)|X) \quad (tr(c) = c \text{ for scalar c}) \\
&= \mathbb{E}\ (\text{tr}(M_1\Omega^{-1}M_1ee')|X) \quad \text{(using properties of trace)} \\
&= \text{tr}\ (\mathbb{E}(M_1\Omega^{-1}M_1ee'|X)) \quad \text{(trace and expectation are linear operator)} \\
&= \text{tr}\ (M_1\Omega^{-1}M_1\mathbb{E}(ee'|X)) \quad (M_1, \Omega \text{ are function of } X) \\
&= \text{tr}\ (M_1\Omega^{-1}M_1\sigma^2\Omega) \\
&= \sigma^2\text{tr}(I_n - \Omega^{-1}X(X'\Omega^{-1}X)^{-1}X') \quad (\text{using}(d)) \\
&= \sigma^2(n - \text{tr}((X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}X))) = \sigma^2(n-k)
\end{aligned}$$

Thus, $\mathbb{E}(s^2|X) = \frac{1}{n-k}\mathbb{E}(e'M_1\Omega^{-1}M_1e|X) = \sigma^2$.

(f) In this (heteroskedastic) linear regression model with known Ω, s^2 is unbiased estimator, so it is fairly reasonable estimator for σ^2.

Exercise 4.5

(a) Note that $\tilde{\beta}$ is a unbiased linear estimator. Under homoskedastic regression model, OLS is the best linear unbiased estimator by Gauss-Markov Theorem (Theorem 4.6.1.1). $\tilde{\beta}$ would be a good estimator if model is not homoskedastic. If we consider following transformation, $\tilde{y} = W^{1/2}y, \tilde{X} = W^{1/2}X$ where $W^{1/2} = \text{diag}(w_1^{1/2}, \cdots, w_n^{1/2})$, then $\tilde{\beta} = (\tilde{X}'\tilde{X})^{-1}(\tilde{X}\tilde{y})$ is OLS estimator of $\tilde{y}$ on $\tilde{X}$.

From this intuition we guess $\tilde{\beta}$ would be a good estimator when transformed linear regression model is homoskedastic.

(b) By Theorem 4.6.1.2 and exercise 4.3, $\tilde{\beta}$ is the best unbiased linear estimator under $\mathbb{E}(e'e|X) = W^{-1} = (\mathrm{diag}(w_1, \cdots, w_n))^{-1}$, i.e., $\mathbb{E}(e_i^2|x_i) = w_i^{-1} = x_{ji}^2$

Exercise 4.6

$$
\begin{aligned}
\mathbb{E}(\bar{\sigma}^2|X) &= \frac{1}{n}\sum_{i=1}^{n}(1-h_{ii})^{-1}\mathbb{E}(\hat{e}_i^2|X) \\
&= \frac{1}{n}\sum_{i=1}^{n}(1-h_{ii})^{-1}(1-h_{ii})\sigma^2 \quad \text{(use (4.16) in the homoskedastic regression model)} \\
&= \sigma^2
\end{aligned}
$$