(see page 224)

Digging into the OLS estimator, $\mathbf{b}$.

Recall that:

$$\mathbf{b} = (X'X)^{-1}X'\mathbf{y}$$
$$= (X'X)^{-1}X'(X\boldsymbol{\beta} + \mathbf{e})$$
$$= (X'X)^{-1}X'X\boldsymbol{\beta} + (X'X)^{-1}X'\mathbf{e}$$
$$= \boldsymbol{\beta} + (X'X)^{-1}X'\mathbf{e}$$

<u>Do this before proving $\mathbf{b}$ is BLUE.</u>

Gauss Markov Assumptions:

$$\mathbf{b} \sim (\boldsymbol{\beta}, \sigma^2(X'X)^{-1})$$
$$\mathbf{y} \sim (X\boldsymbol{\beta}, \sigma^2 I_n)$$
$$\mathbf{e} \sim (0, \sigma^2 I_n)$$

Provided that the Gauss Markov Assumptions hold, the Ordinary Least Squares (OLS) estimator, $\mathbf{b}$, will be the best linear unbiased estimator (BLUE).

Prove that, given the Gauss Markov Assumptions hold, the OLS estimator $\mathbf{b}$ is in fact BLUE.

To prove that the OLS estimator $\mathbf{b}$ is BLUE I must first show that it is a linear estimator, second that it is an unbiased linear estimator, and lastly that it is the best of unbiased linear estimators.

- Assume that $E[\mathbf{e}] = \mathbf{0}$ and $E[\mathbf{ee}'] = \sigma^2 I$.

Any linear estimator of $\boldsymbol{\beta}$ may be written as:

$\tilde{\mathbf{b}} = A\mathbf{y}$ where A is a $K \times n$ matrix made up of elements independent of the dependent variable and the unknown parameter to be estimated.

Now, recall that the OLS estimator is:

$\mathbf{b} = (X'X)^{-1}X'\mathbf{y}$ (where $\underset{K\times n\ n\times K}{(X'X)^{-1}}\underset{K\times n}{X'}$ is a $K \times n$ matrix that does not depend on $\boldsymbol{\beta}$ or $\mathbf{y}$)

$\therefore$ $\mathbf{b}$ is a linear estimator of $\boldsymbol{\beta}$.

To show that $\mathbf{b}$ is an unbiased estimator, I must show that $E[\mathbf{b}] = \boldsymbol{\beta}$

$E[\mathbf{b}] = E[(X'X)^{-1}X'\mathbf{y}]$ (by definition)

$= [(X'X)^{-1}X'] \cdot E[\mathbf{y}]$ (by rules regarding mathematical expectation)

$= [(X'X)^{-1}X'] \cdot E[X\boldsymbol{\beta} + \mathbf{e}]$ (by definition of $\mathbf{y}$)

$= [(X'X)^{-1}X'] \cdot (X\boldsymbol{\beta} + E[\mathbf{e}])$ (by rules of mathematical expectation)

$= [(X'X)^{-1}X'X\boldsymbol{\beta}] + (X'X)^{-1}X' \cdot E[\mathbf{e}]$ (by distributive property)

$= I\boldsymbol{\beta} + \mathbf{0}$ (by Gauss-Markov assumption that $E[\mathbf{e}] = \mathbf{0}$)

$= \boldsymbol{\beta}$ (by properties of identity matrices)

$\therefore \; E[\mathbf{b}] = \boldsymbol{\beta}$; that is, the OLS estimator is unbiased.

To show that the linear, unbiased estimator is best of the linear unbiased estimators of the unknown population parameter, I must show that it has the minimum variance of the other linear unbiased estimators.

The sampling variability of $\mathbf{b}$ is given by its variance-covariance matrix.

$\text{cov}(\mathbf{b}) = E[(\mathbf{b} - E[\mathbf{b}])(\mathbf{b} - E[\mathbf{b}])']$ (by definition of the covariance matrix)

$= E[(\mathbf{b} - \boldsymbol{\beta})(\mathbf{b} - \boldsymbol{\beta})']$ (since OLS estimator is unbiased)

$= E[(\boldsymbol{\beta} + (X'X)^{-1}X'\mathbf{e} - \boldsymbol{\beta})(\boldsymbol{\beta} + (X'X)^{-1}X'\mathbf{e} - \boldsymbol{\beta})']$

(from earlier equality* $\mathbf{b} = \boldsymbol{\beta} + (X'X)^{-1}X'\mathbf{e}$)

*see "Digging into the OLS estimator"

$= E[([X'X]^{-1}X'\mathbf{e})([X'X]^{-1}X'\mathbf{e})']$

$= E[([X'X]^{-1}X'\mathbf{e})(\mathbf{e}'X[X'X]^{-1})]$

(by linear algebra rules:

1] $(ABC)' = C'B'A'$ when A, B, and

$$\overset{n\times 1}{y} = \overset{n\times 1}{X}\overset{k\times 1}{\beta} + e$$

C are conformable matrices, and

2] the transpose of a symmetric matrix is itself ☆ pg. 225

$$= \underbrace{\underbrace{(X'X)^{-1}}_{(k\times n)(n\times k)}}_{(k\times k)} \underbrace{X'}_{(k\times n)} E[\mathbf{e}\mathbf{e}'] \underbrace{X}_{(n\times k)} \underbrace{(X'X)^{-1}}_{(k\times n)\,(n\times k)\;(k\times k)}$$

(since X is fixed from sample to sample)

$$= \underbrace{(X'X)^{-1}X'}_{k\times n}\sigma^2 \underbrace{I}_{(n\times n)} \underbrace{X(X'X)^{-1}}_{n\times k}$$

(by Gauss Markov assumption that $E(\mathbf{e}\mathbf{e}') = \sigma^2 I_n$)

$$= \sigma^2 \underbrace{(X'X)^{-1}}_{(k\times k)} \underbrace{X'X}_{(k\times k)} \underbrace{(X'X)^{-1}}_{(k\times k)}$$

(σ^2 is a scalar)

$$= \sigma^2 (X'X)^{-1} I_k$$

$$= \sigma^2 (X'X)^{-1}$$

$$\therefore \; \mathrm{cov}(\mathbf{b}) = \sigma^2 (X'X)^{-1}$$

and from above results that the estimator is unbiased, we can say that

$$\mathbf{b} \sim (\boldsymbol{\beta}, \sigma^2 (X'X)^{-1})$$

Consider $\tilde{\mathbf{b}}$, which is an alternative estimator of $\boldsymbol{\beta}$, where

$$\tilde{\mathbf{b}} = [(X'X)^{-1}X' + C]\mathbf{y}$$

(where C is a constant $k\times n$ matrix)

$\tilde{\mathbf{b}}$ is in the form $A\mathbf{y}$ where A is still independent of $\boldsymbol{\beta}$ & $\mathbf{y}$, and thus is a linear estimator.

To affirm that $\tilde{\mathbf{b}}$ is an unbiased linear estimator, I must show that $E[\tilde{\mathbf{b}}] = \boldsymbol{\beta}$.

$$E[\tilde{\mathbf{b}}] = E[([X'X]^{-1}X' + C)\mathbf{y}]$$ (by definition of $\tilde{\mathbf{b}}$)

$$= ([X'X]^{-1}X' + C)(E[\mathbf{y}])$$ (by nature of mathematical expectation)

$$= ([X'X]^{-1}X' + C)(E[X\boldsymbol{\beta} + \mathbf{e}])$$

$$= ([X'X]^{-1}X' + C)(X\boldsymbol{\beta} + E[\mathbf{e}])$$

$$= ([X'X]^{-1}X' + C)(X\boldsymbol{\beta})$$ (by Gauss Markov Assumption that $E[\mathbf{e}] = \mathbf{0}$)

$$= (X'X)^{-1}X'X\boldsymbol{\beta} + CX\boldsymbol{\beta}$$

$$= \boldsymbol{\beta} + CX\boldsymbol{\beta}$$

$$= \boldsymbol{\beta} \quad \text{iff} \quad CX\boldsymbol{\beta} = \mathbf{0}$$

$\therefore$ $\tilde{\mathbf{b}}$ is unbiased if $CX\boldsymbol{\beta} = \mathbf{0}$, so we will assume this restriction to check whether $\mathbf{b}$ is best.

To determine whether $\mathbf{b}$ is best (has minimum sampling variability) I must know that of $\tilde{\mathbf{b}}$).

$$\operatorname{cov}(\tilde{\mathbf{b}}) = E[(\tilde{\mathbf{b}} - E[\tilde{\mathbf{b}}])(\tilde{\mathbf{b}} - E[\tilde{\mathbf{b}}])']$$

(by definition of $\operatorname{cov}(\tilde{\mathbf{b}})$)

$$= E[(\tilde{\mathbf{b}} - \boldsymbol{\beta})(\tilde{\mathbf{b}} - \boldsymbol{\beta})']$$

$$= E[(([X'X]^{-1}X' + c)\mathbf{y} - \boldsymbol{\beta})(([X'X]^{-1}X' + c)\mathbf{y} - \boldsymbol{\beta})']$$

$$= E[(([X'X]^{-1}X' + C)(X\boldsymbol{\beta} + \mathbf{e}) - \boldsymbol{\beta})(([X'X]^{-1}X' + C)$$

$$* (X\boldsymbol{\beta} + \mathbf{e}) - \boldsymbol{\beta})']$$

$$= E[((X'X)^{-1}X'X\boldsymbol{\beta} + (X'X)^{-1}X'\mathbf{e} + CX\boldsymbol{\beta} + C\mathbf{e} - \boldsymbol{\beta})$$

$$* ((X'X)^{-1}X'X\boldsymbol{\beta} - (X'X)^{-1}X'\mathbf{e} + CX\boldsymbol{\beta} + C\mathbf{e} - \boldsymbol{\beta})']$$

$$= E[((X'X)^{-1}X'\mathbf{e} + C\mathbf{e})((X'X)^{-1}X'\mathbf{e} + C\mathbf{e})']$$

(because $CX\boldsymbol{\beta} = 0$ for unbiasedness)

$$= E[([X'X]^{-1}X'\mathbf{e} + C\mathbf{e})(\mathbf{e}'X[X'X]^{-1} + \mathbf{e}'C')]$$

(because transpose of a sum is the sum of a transpose, transpose of a symmetrix matrix is the matrix, and $(ABC)' = C'B'A'$ with conformable matrices)

$$= E[(X'X)^{-1}X'\mathbf{ee}'X(X'X)^{-1} + (X'X)^{-1}X'\mathbf{ee}'C' + C\mathbf{ee}'X + C\mathbf{ee}'C']$$

$$= (X'X)^{-1}X'E[\mathbf{ee}']X(X'X)^{-1} + (X'X)^{-1}X'E[\mathbf{ee}']C' + CE[\mathbf{ee}']X + CE[\mathbf{ee}']C'$$

(because X is fixed from sample to sample)

$$= (X'X)^{-1}X'\sigma^2 I_n X(X'X)^{-1} + (X'X)^{-1}X'\sigma^2 I_n C' + C\sigma^2 I_n X + C\sigma^2 I_n C'$$

(by Gauss Markov assumption that $E[\mathbf{ee}'] = \sigma^2 I_n$)

$$= \sigma^2([X'X]^{-1} + [X'X]^{-1}X'C' + CX + CC')$$

$$= \sigma^2([X'X]^{-1} + (CX[X'X]^{-1})' + CX + CC')$$

(because $(ABC)' = C'B'A'$)

$$= \sigma^2[(X'X)^{-1} + CC']$$

(because $CX = \mathbf{0}$ by unbiased assumption)

$\therefore \text{var}(\tilde{b}) = \sigma^2[(X'X)^{-1} + CC']$ where $CC' \geq 0$ because any matrix multiplied by its own transpose is positive semidefinite.

Therefore $\text{var}(\tilde{\mathbf{b}}) > \text{var}(\mathbf{b})$ in every case except when $C = \mathbf{0}$, at which point $\tilde{\mathbf{b}} = \mathbf{b}$ and I conclude that $\mathbf{b}$ is the best of all linear unbiased estimators.

☆ pg. 207

5C.1

$$A = \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$$

(a) Find A^{-1}, B^{-1}, and $B^{-1}A^{-1}$:

$$A^{-1} = \frac{1}{2} \cdot \begin{bmatrix} 2 & -1 \\ -2 & 2 \end{bmatrix} \qquad B^{-1} = 4 \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -\frac{1}{2} \\ -1 & 1 \end{bmatrix} \qquad = \begin{bmatrix} 4 & 4 \\ -4 & 12 \end{bmatrix}$$

$$B^{-1}A^{-1} = \begin{bmatrix} 0 & 2 \\ & \end{bmatrix}$$

complete later if there is time

The takeaway is that the transpose of a symmetric matrix is itself, but where I'm going to take for granted is that $(X'X)^{-1}$ is symmetric.