ECON 5350 Solutions to the Midterm Exam – Fall 2009

1. (40 pts) Consider the following bivariate regression model:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i;\ i = 1, ..., n.$$

Show the equivalence of the matrix and summation forms for the OLS slope estimator, b_1. You do not need to derive the summation form. Hint: The inverse of A is

$$A^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Solution. The estimates in summation form are

$$\begin{aligned} b_0 &= \bar{y} - b_1\bar{x} \\ b_1 &= \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}. \end{aligned}$$

The matrix estimates are

$$\begin{aligned} b &= \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & \sum_{i=1}^n x_i^2 \end{bmatrix}^{-1} \begin{bmatrix} n\bar{y} \\ \sum_{i=1}^n x_i y_i \end{bmatrix} \\ &= \frac{1}{n\sum_{i=1}^n x_i^2 - n^2\bar{x}^2} \begin{bmatrix} \sum_{i=1}^n x_i^2 & -n\bar{x} \\ -n\bar{x} & n \end{bmatrix} \begin{bmatrix} n\bar{y} \\ \sum_{i=1}^n x_i y_i \end{bmatrix} \\ &= \frac{1}{n\sum_{i=1}^n x_i^2 - n^2\bar{x}^2} \begin{bmatrix} n\bar{y}\sum_{i=1}^n x_i^2 - n\bar{x}\sum_{i=1}^n x_i y_i \\ -n^2\bar{x}\bar{y} + n\sum_{i=1}^n x_i y_i \end{bmatrix}. \end{aligned}$$

The slope estimator from the matrix form is

$$b_1 = \frac{-n^2\bar{x}\bar{y} + n\sum_{i=1}^n x_i y_i}{n\sum_{i=1}^n x_i^2 - n^2\bar{x}^2} = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}.$$

Expanding the summation form gives

$$\begin{aligned} b_1 &= \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n x_i y_i - \bar{y}\sum_{i=1}^n x_i - \bar{x}\sum_{i=1}^n y_i + \sum_{i=1}^n \bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - 2\bar{x}\sum_{i=1}^n x_i + \sum_{i=1}^n \bar{x}^2} \\ &= \frac{\sum_{i=1}^n x_i y_i - n\bar{y}\bar{x} - n\bar{x}\bar{y} + n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - 2n\bar{x}^2 + n\bar{x}^2} = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}. \end{aligned}$$

2. (30 pts) Consider the Classical multivariate linear regression model:

$$Y = X\beta + \epsilon$$

with two error structures: (i) $\epsilon \sim N(0, \sigma^2 I)$ and (ii) $\epsilon \sim N(0, \sigma^2 \Omega)$, where Ω is a symmetric positive-definite matrix.

(a) Is the ordinary least squares, $b = (X'X)^{-1}(X'Y)$, unbiased for each model? Defend your answer.

Solution. Yes. The errors are mean zero in each case and uncorrelated with X.

(b) Derive the variance of b for each model.

Solution. The respective variances are

$$\begin{aligned}
\text{Case(i). } var(b) &= E[(b-\beta)(b-\beta)'] \\
&= E[(X'X)^{-1}X'\epsilon\epsilon'X(X'X)^{-1}] = \sigma^2(X'X)^{-1}. \\
\text{Case(ii). } var(b) &= E[(b-\beta)(b-\beta)'] \\
&= E[(X'X)^{-1}X'\epsilon\epsilon'X(X'X)^{-1}] = \sigma^2(X'X)^{-1}X'\Omega X(X'X)^{-1}.
\end{aligned}$$

(c) Find the efficient estimator for the partitioned model

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix},$$

where ϵ_1 and ϵ_2 are mutually independent and have variance-covariance matrices, $\sigma_1^2 I$ and $\sigma_2^2 I$.

Solution. The efficient estimator is OLS on each partition.

3. (20 pts) True or False? If false, explain why.

(a) High multicollinearity leads to biased coefficient estimates and low t statistics.

Solution. False. Multicollinearity does not bias the coefficient estimates.

(b) The null hypothesis, $H_0 : \beta_1 + \beta_2 = 1$, can be tested with the following t statistic

$$t = \frac{b_1 + b_2 - 1}{se(b_1) + se(b_2)}.$$

Solution. False. The correct standard error is

$$se(b_1 + b_2) = \sqrt{var(b_1) + var(b_2) + 2cov(b_1, b_2)}.$$

(c) Consider the following model $y_i = \beta_0 + \beta_1 x_i^* + \epsilon_i$, where $x_i = x_i^* + \nu_i$ is measured with error. The least squares estimate of β_1 is unbiased if $E(\epsilon_i \cdot \nu_i) = 0$.

Solution. False. The LS estimate is biased (towards zero) even if $E(\epsilon_i \cdot \nu_i) = 0$.

(d) The variance of the estimated marginal effect, $\delta = \partial\hat{y}/\partial x$, from the quadratic model $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i$ is

$$var(\delta) = var(b_1) + 2var(b_2).$$

Solution. False. The correct variance is

$$var(\delta) = var(b_1 + 2b_2 x_i) = var(b_1) + 4x_i^2 var(b_2) + 2x_i cov(b_1, b_2).$$

4. (10 pts) What is the output for the following gauss code?

```
x=zeros(3,3);
for i (1,3,1);
for j (1,i,1);
x[i,j]=(i+j)^2;
endfor;
endfor;
print x;
```

Solution. The output is

$$x = \begin{bmatrix} 4 & 0 & 0 \\ 9 & 16 & 0 \\ 16 & 25 & 36 \end{bmatrix}.$$