Westminster Business School
School of Finance and Accounting
7FNCE044 Predictive Analysis for Decision Making
Semesters 1 and 2
Weeks 1 and 2
Dr. Issam Malki

Seminar Week 1 and 2 Sketch Answers

Linear Regression: Extension I

Computer Based Questions:

You are expected to write R studio (Python) codes to the questions. See codes in Blackboard for more details.

Theoretical Questions

1. Some linear Algebra

1.

a) $\mathbf{X}'\mathbf{X} = \begin{bmatrix} \mathbf{1}' \\ \mathbf{x}' \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{x} \end{bmatrix}$.

$$= \begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\mathbf{x} \\ \mathbf{1}'\mathbf{x} & \mathbf{x}'\mathbf{x} \end{bmatrix},$$

Using summation notation, we know:

$$A = \mathbf{1}'\mathbf{1} = \sum_{i=1}^{n} 1^2 = n,$$

$$B = \mathbf{1}'\mathbf{x} = \sum_{i=1}^{n} 1 \cdot x_i = \sum_{i=1}^{n} x_i .$$

b) $\mathbf{X}'\mathbf{y} = \begin{bmatrix} \mathbf{1}' \\ \mathbf{x}' \end{bmatrix} \mathbf{y}$

$$= \begin{bmatrix} \mathbf{1}'\mathbf{y} \\ \mathbf{x}'\mathbf{y} \end{bmatrix}$$

Using summation notation, we know:

$$C = \mathbf{1}'\mathbf{y} = \sum_{i=1}^{n} y_i \ ,$$

$$D = \mathbf{x}'\mathbf{y} = \sum_{i=1}^{n} x_i y_i \ .$$

c) The rule says that if: $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $\mathbf{A} = \dfrac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

From (a), we know:

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix}$$

Therefore:

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}\begin{bmatrix} \sum_{i=1}^{n} x_i^2 & -\sum_{i=1}^{n} x_i \\ -\sum_{i=1}^{n} x_i & n \end{bmatrix}$$

d) The OLS coefficient vector is given by:

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

From (b) and (c), we know:

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \frac{1}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}\begin{bmatrix} \sum_{i=1}^{n} x_i^2 & -\sum_{i=1}^{n} x_i \\ -\sum_{i=1}^{n} x_i & n \end{bmatrix}\begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix}$$

Therefore,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2};$$

$$\hat{\beta}_2 = \frac{-\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i + n\sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}.$$

e) We know that **X** must have full rank in order for $\mathbf{X}'\mathbf{X}^{-1}$ to exist, which means that the rank must be 2 in this case.

In order for **X** to have full rank, one column cannot be a linear combination of the other columns. In this case, there are only two variables, therefore we require that **x** is not proportion to the vector of ones. This means that **x** cannot be a constant, *i.e*. the x_i cannot all be equal.

2. **OLS Properties**

(a) and (b)

$$\mathrm{E}(\hat{\beta}) = E\{(X'X)^{-1}X'Y\} = E\{(X'X)^{-1}X'X\beta + (X'X)^{-1}X'u\} = \beta$$

Find the variance:

$$\mathrm{var}(\hat{\beta}) = E\{(\hat{\beta} - E(\hat{\beta}))'(\hat{\beta} - E(\hat{\beta}))\} = E\{(X'X)^{-1}u'u\} = (X'X)^{-1}\sigma^2$$

The distribution, assuming normality, can be described as:

$$\hat{\beta} \sim N(\beta, (X'X)^{-1}\sigma^2)$$