Spring 2022

Quantitative Analysis
Chung-Ming Kuan
Jui-Chung Yang

Problem Set 2

Due: 3/7

Part One: Hand-Written Exercise

1. Verify the statement in slide 18, Lecture 2. That is, let $\hat{r}_{i,1}$ be the OLS residual of regressing x_1 on the constant one and $x_2, ..., x_k$. Show that $\sum_{i=1}^n \hat{r}_{i,1} x_{i,1} = \sum_{i=1}^n \hat{r}_{i,1}^2$.

2. You have collected the following data on the response of a variable, y, to two other variables, x_1 and x_2:

y	1	0	2	1	3	5
x_1	0	0	0	1	1	1
x_2	0	1	2	0	1	2

   You want to fit the data to a multiple linear regression $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i$ and you have determined that $(\mathbf{X'X})^{-\mathbf{1}} = \frac{1}{12}\begin{bmatrix} 7 & -4 & -3 \\ -4 & 8 & 0 \\ -3 & 0 & 3 \end{bmatrix}$

   (a) Determine $\hat{\beta}_1$ and $\hat{\beta}_2$.

   (b) Let $\tilde{\beta}_1$ be the estimate of the coefficient of x_1 in the simple linear regression $y = \tilde{\beta}_0 + \tilde{\beta}_1 x_1 + u$ and $\check{\beta}_2$ be the estimate of the coefficient of x_2 in the simple linear regression $y = \check{\beta}_0 + \check{\beta}_2 x_2 + v$. Determine $\hat{\beta}_1 - \tilde{\beta}_1$ and $\hat{\beta}_2 - \check{\beta}_2$.

   (c) What's the implication behind your answer in (b)?

3. Consider the model $y_i = \beta_0 + \beta_1 x_i + u_i$ with $\text{Var}(y_i) = \sigma^2$. Under the Classical Assumptions, the OLS estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased. Let $\tilde{\beta}_1$ be the OLS estimator of β_1 by assuming the intercept is zero. That is, $\tilde{\beta}_1$ is obtained under the assumption $\beta_0 = 0$.

   (a) Calculate $\mathbb{E}(\tilde{\beta}_1)$ in terms of x_i, β_0, and β_1.

   (b) If $\beta_0 \neq 0$, is $\tilde{\beta}_1$ unbiased?

   (c) Calculate the variance of $\tilde{\beta}_1$.

   (d) Compare between $\text{Var}(\tilde{\beta}_1)$ and $\text{Var}(\hat{\beta}_1)$. Is it true that $\text{Var}(\tilde{\beta}_1) \leq \text{Var}(\hat{\beta}_1)$ in general?

   (e) Does the result in (d) violate the Gauss-Markov Theorem, which states that $\hat{\beta}_1$ should have the smallest variance? Explain.

Part Two: Computer Exercise

1. Let $\mathbf{X} = \begin{bmatrix} 7 & 2 & 3 \\ 4 & 6 & 7 \\ 9 & 2 & 0 \\ 0 & 9 & 0 \\ 5 & 3 & 5 \end{bmatrix}$ and $\mathbf{Y} = \begin{bmatrix} 6 \\ 2 \\ 4 \\ 2 \\ 1 \end{bmatrix}$.

   (a) Please construct the OLS estimator $\hat{\beta}$. (Reminder: Don't forget the intercept term.)

   (b) Given a new observation $x^* = (0, 4, 3)'$, please calculate $\hat{y}$.

2. Please load the data set "mtcars" from R using the code `data(mtcars)`. The data was extracted from the 1974 Motor Trend US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973-74 models).

   (a) Please show the data for the automobile "Camaro Z28".

   (b) Please show the `wt` (Weight (1000 lbs)) for all the automobile.

   (c) Please show the data with `gear` (Number of forward gears) = 3.

   (d) Please list the automobiles with `mpg` (miles/gallon) > 10, `cyl` (Number of cylinders) = 6, and `hp` (horsepower) between 90 and 110.

   (e) Suppose we have the following model:

   $$\texttt{drat}_i = \beta_0 + \beta_1 \texttt{wt}_i + \beta_2 \texttt{hp}_i + \beta_3 \texttt{qsec}_i + \beta_4 \texttt{vs}_i + u_i.$$

   Please find $\beta_0, \beta_1, \beta_2, \beta_3$ and β_4 without the function `lm()`.

   (f) Following (e), find those estimators with the function `lm()`.