STAT151A Quiz 2 (Feb 13th)

Please write your full name and email address:

For this quiz, we'll consider the linear models

$$y_n = \boldsymbol{\beta}^\intercal \boldsymbol{x}_n + \varepsilon_n \quad \text{and} \quad y_n = \boldsymbol{\gamma}^\intercal \boldsymbol{z}_n + \eta_n$$

with

$$\boldsymbol{x}_n = (1, x_n)^\intercal \quad \text{and} \quad \boldsymbol{z}_n = (1, z_n)^\intercal \text{ where}$$
$$\overline{x} := \frac{1}{N}\sum_{n=1}^{N} x_n \quad \text{and} \quad z_n := x_n - \overline{x}.$$

Assume that x_n is not a constant (i.e., for at least one pair n and m, $x_n \neq x_m$.).

Let $\boldsymbol{X}$ denote the $N \times 2$ matrix whose n–th row is $\boldsymbol{x}_n^\intercal$, and $\boldsymbol{Z}$ denote the $N \times 2$ matrix whose n–th row is $\boldsymbol{z}_n^\intercal$.

Recall that the inverse of a 2x2 matrix is given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

You have 20 minutes for this quiz.

There are three parts, (a), (b), and (c), each weighted equally..

(a)

Find a 2×2 matrix $\boldsymbol{A}$ such that $\boldsymbol{Z} = \boldsymbol{X}\boldsymbol{A}$.

(b)

Suppose I tell you that the OLS estimate of β is given by $\hat{\beta} = (2, 3)$, and that $\overline{x} = 4$. What is the value of $\hat{\gamma}$, the OLS estimate of γ?

(c)

In general, can you say whether one regression will provide a better fit than the other? That is, can you say which of $\frac{1}{N}\sum_{n=1}^{N}(y_n - \boldsymbol{z}_n^\intercal\hat{\gamma})^2$ and $\frac{1}{N}\sum_{n=1}^{N}(y_n - \boldsymbol{x}_n^\intercal\hat{\beta})^2$ is smaller? Argue why or why not.