Homework 5

Richard L. McCormick

Northern Arizona University

STA471: Statistical Regression

Dr. Jin Wang

October 31st, 2023

1. Show that (*I*–*H*) is a symmetric and idempotent matrix, where ***H* = *X*(*X*^T^*X*)^-1^*X*^T^**.

I) Symmetric

$$\left(X(X^TX)^{-1}X^T\right)^T = (X^T)^T \circ \left[(X^TX)^{-1}\right]^T \circ X^T$$

$$= X \cdot \left[\underbrace{(X^TX)^{-1}}\right]^T \circ X^T$$

(X^TX) is symmetric, so $(X^TX)^{-1}$ must also be $\Rightarrow (X^TX)^{-1^T} = (X^TX)^{-1}$

$$= X(X^TX)^{-1}X^T$$

Since I only has diagonal elements, $I - H$, where H is symmetric, results in a symmetric matrix. Thus, $(I-H) = (I-H)^T$.

II) Idempotent

$$H^2 = \left(X(X^TX)^{-1}X^T\right)\left(X(X^TX)^{-1}X^T\right)$$

$$= X\left(\cancel{(X^TX)^{-1}}\,\cancel{(X^TX)}(X^TX)^{-1}\right)X^T$$

$$= X(X^TX)^{-1}X^T$$

$$= H$$

So H is idempotent...

$$(I-H)^2 = (I-H)(I-H)$$

$$= I^2 - HI - IH + H^2$$

$$= I - H - H + H$$

$$= I - H \quad \blacksquare$$

So $(I-H)$ is idempotent.

2. Suppose that we are given the data below:

X	1	3	6	4
Y	4.2	2.3	2.7	2.0

a. Write out the matrix form of the linear regression model for the data.

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$$

$$\begin{bmatrix} 4.2 \\ 2.3 \\ 2.7 \\ 2.0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 6 \\ 1 & 4 \end{bmatrix} \circ \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}$$

b. Compute X^TX, X^TY, and Y^TY.

I. $X^T \times X$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 6 & 4 \end{bmatrix} \times \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 6 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 14 \\ 14 & 62 \end{bmatrix}$$

II. X^TY

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 6 & 4 \end{bmatrix} \times \begin{bmatrix} 4.2 \\ 2.3 \\ 2.7 \\ 2.0 \end{bmatrix} = \begin{matrix} 4.2 + 2.3 + 2.7 + 2 \\ 4.2 + (3 \times 2.3) + (6 \times 2.7) + 8 \end{matrix} = \begin{bmatrix} 11.2 \\ 35.3 \end{bmatrix}$$

III. Y^TY

$$\begin{bmatrix} 4.2 & 2.3 & 2.7 & 2.0 \end{bmatrix} \times \begin{bmatrix} 4.2 \\ 2.3 \\ 2.7 \\ 2.0 \end{bmatrix} = 4.2^2 + 2.3^2 + 2.7^2 + 2.0^2 = \begin{bmatrix} 34.22 \end{bmatrix}$$

c. Find $(X^TX)^{-1}$.

$$X^T \times X = \begin{bmatrix} 4 & 14 \\ 14 & 62 \end{bmatrix} \quad \text{(Problem 2b.)}$$

$$|X^T \times X| = (4 \times 62) - (14 \times 14) = 52$$

$$\boxed{(X^TX)^{-1} = \frac{1}{52}\begin{bmatrix} 62 & -14 \\ -14 & 4 \end{bmatrix}}$$

d. Find $H = X(X^TX)^{-1}X^T$.

$$X \times (X^TX)^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 6 \\ 1 & 4 \end{bmatrix} \cdot \begin{bmatrix} 62 & -14 \\ -14 & 4 \end{bmatrix} \cdot \frac{1}{52}$$

$$= \begin{bmatrix} 62-14 & -14+1(4) \\ 62-3(14) & -14+3(4) \\ 62-6(14) & -14+6(4) \\ 62-4(14) & -14+4(4) \end{bmatrix} = \begin{bmatrix} 48 & -10 \\ 20 & -2 \\ -22 & 10 \\ 6 & 2 \end{bmatrix} \cdot \frac{1}{52}$$

$$X \times (X^TX)^{-1}X^T = \begin{bmatrix} 48 & -10 \\ 20 & -2 \\ -22 & 10 \\ 6 & 2 \end{bmatrix} \times \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 6 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 48-10 & 48+3(-10) & 48+6(-10) & 48+4(-10) \\ 20-2 & 20+3(-2) & 20+6(-2) & 20+4(-2) \\ -22+10 & -22+3(10) & -22+6(10) & -22+4(10) \\ 6+2 & 6+3(2) & 6+6(2) & 6+4(2) \end{bmatrix}$$

$$\boxed{H = \begin{bmatrix} 38 & 18 & -12 & 8 \\ 18 & 14 & 8 & 12 \\ -12 & 8 & 38 & 18 \\ 8 & 12 & 18 & 14 \end{bmatrix} \times \frac{1}{52}}$$

e. Find ***b***.

$$\underline{b} = (X \cdot X^T)^{-1} \cdot (X^T \cdot \underline{y})$$

$$\underline{b} = \begin{bmatrix} 62 & -14 \\ -14 & 4 \end{bmatrix} \times \begin{bmatrix} 11.2 \\ 35.3 \end{bmatrix} \times \frac{1}{52}$$

$$= \begin{bmatrix} 62 \times 11.2 + -14 \times 35.3 \\ -14 \times 11.2 + 4 \times 35.3 \end{bmatrix} \times \frac{1}{52}$$

$$= \begin{bmatrix} 200.2 \\ -15.6 \end{bmatrix} \times \frac{1}{52}$$

$$\boxed{\underline{b} = \begin{bmatrix} 3.85 \\ -0.3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_0 \end{bmatrix}}$$

f. Find $\hat{Y}$.

$$\underline{\hat{y}} = X\underline{b}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 6 \\ 1 & 4 \end{bmatrix} \times \begin{bmatrix} 3.85 \\ -0.3 \end{bmatrix} = \begin{bmatrix} 3.85 + (-0.3)(1) \\ 3.85 + (-0.3)(3) \\ 3.85 + (-0.3)6 \\ 3.85 + (-0.3)4 \end{bmatrix}$$

$$\boxed{\underline{\hat{y}} = \begin{bmatrix} 3.55 \\ 2.95 \\ 2.05 \\ 2.65 \end{bmatrix}}$$

g. Find ***e***.

$$\underline{e} = \underline{y} - \underline{\hat{y}}$$

$$= \begin{bmatrix} 4.2 \\ 2.3 \\ 2.7 \\ 2.0 \end{bmatrix} - \begin{bmatrix} 3.55 \\ 2.95 \\ 2.05 \\ 2.65 \end{bmatrix}$$

$$\underline{e} = \begin{bmatrix} 0.65 \\ -0.65 \\ 0.65 \\ -0.65 \end{bmatrix}$$

h. Find $s^2 = MS_{resid}$.

$$RSS = \underline{e}' \times \underline{e} = \sum_{i=1}^{4} e_i^2$$

$$= (0.65)^2 + (-0.65)^2 + (0.65)^2 + (-0.65)^2$$

$$RSS = 1.69$$

$$\boxed{s^2 = \frac{RSS}{n-2} = \frac{1.69}{4-2} = 0.845}$$

i. Verify that $S_{yy} = \Sigma (Y_i - \bar{Y})^2 = Y^T(I_n - 1/n\ 1\ 1^T)Y$ using the sample information.

$$TSS = S_{yy} = \sum_{i=1}^{4} (Y_i - \bar{Y})^2 \quad \text{where } \bar{Y} = 2.8$$

$$= (4.2 - 2.8)^2 + (2.7 - 2.8)^2 + (2.3 - 2.8)^2 + (2.0 - 2.8)^2$$

$$\boxed{TSS = 2.86}$$

j. Find R^2 and r_{xy}.

$$SS_{reg} = TSS - RSS = 2.86 - 1.69 = 1.17$$

$$R^2 = \frac{SS_{reg}}{TSS} = \frac{1.17}{2.86} = 0.40\overline{9}$$

$$r_{xy} = \sqrt{R^2} = \sqrt{0.40\overline{9}} = 0.639$$

$$\boxed{R^2 = (0.40\overline{9}) \quad ; \quad r_{xy} = 0.639}$$