MATH 423/533 Fall 2018

November 21, 2018

Homework 3

This homework is due on Dec 6 at 11:59pm.

Please submit your solutions with relevant R code snippets included as a pdf file via myCourses. Please also upload a separate file containing your entire code as an .rmd script.

Question 1

Suppose that

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i, \qquad i = 1, \ldots, n$$

where $\epsilon_i \sim N(0, \sigma^2)$. Notice that there is no intercept. Suppose that

$$\sum_{i=1}^{n} X_{1i} X_{2i} = 0.$$

Show that the least squares estimators $\widehat{\beta}_1$ and $\widehat{\beta}_2$ from the multiple regression are the same as if we were to fit separate, simple regressions on X_1 and X_2.

Question 2

Consider these data

X_1	4	3	10	7
X_2	5	4	9	10
Y	25	20	57	50

1. Fit the multiple regression in `R` and summarize the results.
2. Construct $\mathbf{X}^T\mathbf{X}$ and $(\mathbf{X}^T\mathbf{X})^{-1}$
3. Construct $\widehat{\beta}$ directly (show your work) and confirm that you get the same answer as you got from `R`.
4. Construct the hat matrix $\mathbf{H}$.
5. Compute $\text{Var}(\widehat{\beta})$ using your calculations.

Question 3

Recall that two vectors v and w or orthogonal if $v^T w = 0$. Let $\mathbf{e}$ be the vector of residuals and let $\widehat{\mathbf{Y}}$ be the vector of fitted values. Use the properties of the hat matrix to show that $\mathbf{e}$ and $\widehat{\mathbf{Y}}$ are orthogonal.

Question 4

Load the stackloss data:

```
data(stackloss)
names(stackloss)

## [1] "Air.Flow"   "Water.Temp" "Acid.Conc." "stack.loss"

help(stackloss)
```

1. Plot the data.
2. Fit a multiple regression model to predict stackloss from the three other variables. The model is

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon$$

   where Y is stackloss, X_1 is airflow, X_2 is water temperature and X_3 is acid. Summarize the results.

3. Construct 90 percent confidence intervals for the coefficients of the linear regresion model.

4. Construct a 99 percent prediction interval for a new observation when Airflow = 58, Water temperature = 20 and Acid = 86.

5. Test the null hypothesis $H_0 : \beta_3 = 0$. What is the p-value? What is the conclusion (at $\alpha = 0.10$)?