STAT 3008
Exercise 4

Problems refer to the problem sets in the textbook: Applied Linear Regression, 3rd edition by Weisberg.

1. Problem 3.2.
   3.2. Added-variable plots
   This problem uses the United Nations example in Section 3.1 to demonstrate many of the properties of added-variable plots. This problem is based on the mean function

   $$E(log(Fertility)|log(PPgdp) = x_1, Purban = x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

   There is nothing special about the two-predictor regression mean function, but we are using this case for simplicity.
   3.2.1. Show that the estimated coefficient for log(PPgdp) is the same as the estimated slope in the added-variable plot for log(PPgdp) after Purban. This correctly suggests that all the estimates in a multiple linear regression model are adjusted for all the other terms in the mean function. Also, show that the residuals in the added-variable plot are identical to the residuals from the mean function with both predictors.

   3.2.2. Show that the t -test for the coefficient for log(PPgdp) is not quite the same from the added-variable plot and from the regression with both terms, and explain why they are slightly different.

2. Problem 3.3.
   3.3. The following questions all refer to the mean function

   $$E(Y|X_1 = x_1, X_2 = x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

   (3.26)

3.3.1. Suppose we fit (3.26) to data for which $x_1 = 2.2x_2$, with no error. For example, x_1 could be a weight in pounds, and x_2 the weight of the same object in kg. Describe the appearance of the added-variable plot for X_2 after X_1.
3.3.2. Again referring to (3.26), suppose now that Y and X_1 are perfectly correlated, so Y $= 3X_1$, without any error. Describe the appearance of the added-variable plot for X_2 after X_1.
3.3.3. Under what conditions will the added-variable plot for X_2 after X_1 have exactly the same shape as the scatterplot of Y versus X_2?
3.3.4. True or false: The vertical variation in an added-variable plot for X_2 after X_1 is always less than or equal to the vertical variation in a plot of Y versus X_2. Explain.

3. Using matrix calculation, reproduce the result of simple linear regression:

$$\begin{aligned}\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1\bar{x} \\ \hat{\beta}_1 &= SXY/SXX\end{aligned}$$

4. Let

$$M = \begin{pmatrix} 1 & 3 & 4 \\ 1 & 2 & 5 \\ 1 & 1 & 2 \end{pmatrix}. \qquad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}.$$

Let $f(\beta) = \beta' M \beta$

(i) Compute $f(\beta)$ explicitly in terms of β_1, β_2 and β_3, then find $\frac{\partial f(\beta)}{\partial \beta}$.

(ii) Use matrix differentiation to find $\frac{\partial f(\beta)}{\partial \beta}$.

5. Let

$$A = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}. \quad X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}. \quad E(X) = \begin{pmatrix} 2 \\ -3 \end{pmatrix}. \quad Var(X) = \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}.$$

(i) Find $E(AX)$, $E(X'A)$, $E(X'A')$ and $E(A'X)$.

(ii) Find $Var(AX)$ and $Var(A'X)$.

(iii) Compute $X'AX$ explicitly in terms of X_1, X_2, then find $E(X'AX)$.

(iv) Using the fact that $E(X'AX) = E[tr(X'AX)] = tr[AE(XX')]$, find $E(X'AX)$.