Stat 354 Fall 2018
Solutions to Assignment #3

(4.19) (a) Observe that we can reformulate the linear regression model in matrix notation as

$$\mathbf{y} = \mathbf{X}\beta + \boldsymbol{\epsilon}$$

where β is a (one-dimensional) parameter,

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_{12} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_{12} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{12} \end{bmatrix},$$

and $\boldsymbol{\epsilon}$ satsifies $\mathbb{E}(\boldsymbol{\epsilon}) = \mathbf{0}$ and $V(\boldsymbol{\epsilon}) = \sigma^2 V$ with

$$V = \text{diag}(x_1^2, \ldots, x_{12}^2) = \begin{bmatrix} x_1^2 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & x_{12}^2 \end{bmatrix}.$$

We compute

$$\mathbf{X}'V^{-1} = \begin{bmatrix} x_1 & \cdots & x_{12} \end{bmatrix} \begin{bmatrix} x_1^{-2} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & x_{12}^{-2} \end{bmatrix} = \begin{bmatrix} x_1^{-1} & \cdots & x_{12}^{-1} \end{bmatrix}$$

and

$$\mathbf{X}'V^{-1}\mathbf{X} = \begin{bmatrix} x_1^{-1} & \cdots & x_{12}^{-1} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{12} \end{bmatrix} = 12$$

so that from equations (4.58) and (4.59) we conclude that the generalized least squares estimator is

$$\hat{\beta}^{\text{GLS}} = (\mathbf{X}'V^{-1}\mathbf{X})^{-1}\mathbf{X}'V^{-1}\mathbf{y} = \frac{1}{12} \begin{bmatrix} x_1^{-1} & \cdots & x_{12}^{-1} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_{12} \end{bmatrix} = \frac{1}{12} \sum_{i=1}^{12} \frac{y_i}{x_i}$$

and has variance

$$V(\hat{\beta}^{\text{GLS}}) = \sigma^2 (\mathbf{X}'V^{-1}\mathbf{X})^{-1} = \frac{\sigma^2}{12}.$$

(b) From the given data, we obtain

$$\hat{\beta}^{\text{GLS}} = \frac{1}{12} \sum_{i=1}^{12} \frac{y_i}{x_i} = \frac{1}{12} \sum_{i=1}^{12} z_i = \frac{30}{12}$$

and

$$V(\hat{\beta}^{\text{GLS}}) = \frac{\sigma^2}{12}.$$

(4.20) **(a)** Observe that we can reformulate the linear regression model in matrix notation as

$$\mathbf{y} = \mathbf{X}\beta + \boldsymbol{\epsilon}$$

where β is a (one-dimensional) parameter,

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_{10} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_{10} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{10} \end{bmatrix},$$

and $\boldsymbol{\epsilon}$ satsifies $\mathbb{E}(\boldsymbol{\epsilon}) = \mathbf{0}$ and $V(\boldsymbol{\epsilon}) = \sigma^2 V$ with

$$V = \operatorname{diag}(x_1, \ldots, x_{10}) = \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & x_{10} \end{bmatrix}.$$

We compute

$$\mathbf{X}'V^{-1} = \begin{bmatrix} x_1 & \cdots & x_{10} \end{bmatrix} \begin{bmatrix} x_1^{-1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & x_{10}^{-1} \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}$$

and

$$\mathbf{X}'V^{-1}\mathbf{X} = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{10} \end{bmatrix} = \sum_{i=1}^{10} x_i$$

so that from equations (4.58) and (4.59) we conclude that the generalized least squares estimator is

$$\hat{\beta}^{\text{GLS}} = (\mathbf{X}'V^{-1}\mathbf{X})^{-1}\mathbf{X}'V^{-1}\mathbf{y} = \left(\sum_{i=1}^{10} x_i \right)^{-1} \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_{10} \end{bmatrix} = \frac{\sum y_i}{\sum x_i}$$

and has variance

$$V(\hat{\beta}^{\text{GLS}}) = \sigma^2 (\mathbf{X}'V^{-1}\mathbf{X})^{-1} = \frac{\sigma^2}{\sum x_i}.$$

(b) From the given data, we obtain

$$\hat{\beta}^{\text{GLS}} = \frac{\sum y_i}{\sum x_i} = \frac{10 \cdot \bar{y}}{10 \cdot \bar{x}} = \frac{30}{15} = 2$$

and

$$V(\hat{\beta}^{\text{GLS}}) = \frac{\sigma^2}{\sum x_i} = \frac{\sigma^2}{10 \cdot \bar{x}} = \frac{\sigma^2}{150}.$$

(Note that the textbook as an error in the solution.)

(5.4) From equation (5.29), we know

$$\text{VIF}_j = \frac{1}{1 - R_j^2}$$

where R_j^2 is the coefficient of determination from the regression of x_j on all other regressors. Hence,

$$\text{VIF}_1 = \frac{1}{1-R_1^2} = \frac{1}{1-0.6} = \frac{1}{0.4} = \frac{5}{2} = 2.5,$$

$$\text{VIF}_2 = \frac{1}{1-R_2^2} = \frac{1}{1-0.8} = \frac{1}{0.2} = \frac{10}{2} = 5,$$

$$\text{VIF}_3 = \frac{1}{1-R_3^2} = \frac{1}{1-0.9} = \frac{1}{0.1} = \frac{10}{1} = 10.$$

(5.5) The correct answer is (e), namely high R^2 and mostly insignificant t ratios suggest the presence of a multicollinearity problem.

(5.14) (a) We know the distribution of the least squares estimator $\hat{\boldsymbol{\beta}}$ is

$$\hat{\boldsymbol{\beta}} \sim \mathcal{N}(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}).$$

We showed in class that the fact that $\mathbf{X}$ is orthogonal implies $\mathbf{X}'\mathbf{X}$ is diagonal. (And therefore $(\mathbf{X}'\mathbf{X})^{-1}$ is diagonal as well.) This implies that the components of $\hat{\boldsymbol{\beta}}$ are uncorrelated, and since $\hat{\boldsymbol{\beta}}$ has a multivariate normal distribution, we deduce that its components must be independent. Hence, $\hat{\beta}_1$ and $\hat{\beta}_j$ are independent as required.

(b) Form the augmented matrix $\mathbf{Z} = \begin{bmatrix}\mathbf{X} & \mathbf{z}\end{bmatrix}$ and the augmented parameter $\boldsymbol{\gamma} = \begin{bmatrix}\boldsymbol{\beta} & \gamma\end{bmatrix}'$ so that we can express the expanded model in matrix notation as $\mathbf{y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$. The least squares estimate for the expanded model is $\hat{\boldsymbol{\gamma}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}$. Now the form of $\mathbf{Z}$, along with the fact that $\mathbf{z}$ is orthogonal to the columns of $\mathbf{X}$, implies that

$$\mathbf{Z}'\mathbf{Z} = \begin{bmatrix}\mathbf{X} & \mathbf{z}\end{bmatrix}' \begin{bmatrix}\mathbf{X} & \mathbf{z}\end{bmatrix} = \begin{bmatrix}\mathbf{X}' \\ \mathbf{z}'\end{bmatrix} \begin{bmatrix}\mathbf{X} & \mathbf{z}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{z}'\mathbf{z}\end{bmatrix}$$

so that

$$(\mathbf{Z}'\mathbf{Z})^{-1} = \begin{bmatrix}(\mathbf{X}'\mathbf{X})^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{z}'\mathbf{z})^{-1}\end{bmatrix}.$$

Therefore,

$$\begin{aligned}\hat{\boldsymbol{\gamma}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y} = \begin{bmatrix}(\mathbf{X}'\mathbf{X})^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{z}'\mathbf{z})^{-1}\end{bmatrix} \begin{bmatrix}\mathbf{X}' \\ \mathbf{z}'\end{bmatrix} \mathbf{y} &= \begin{bmatrix}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' & \mathbf{0} \\ \mathbf{0} & (\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}\end{bmatrix} \mathbf{y} \\ &= \begin{bmatrix}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ (\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}\mathbf{y}\end{bmatrix}.\end{aligned}$$

In other words,

$$\begin{bmatrix}\hat{\boldsymbol{\beta}} \\ \hat{\gamma}\end{bmatrix} = \begin{bmatrix}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ (\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}\mathbf{y}\end{bmatrix}$$

implying that $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ as required.

(c) Write the design matrix $\mathbf{X}$ as an augmented matrix $\mathbf{X} = \begin{bmatrix}\mathbf{1} & \mathbf{X}_c\end{bmatrix}$ where $\mathbf{X}_c$ is an $(n \times p)$ matrix with the property that each column has mean 0. We then compute

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix}\mathbf{1} & \mathbf{X}_c\end{bmatrix}' \begin{bmatrix}\mathbf{1} & \mathbf{X}_c\end{bmatrix} = \begin{bmatrix}\mathbf{1}' \\ \mathbf{X}_c'\end{bmatrix} \begin{bmatrix}\mathbf{1} & \mathbf{X}_c\end{bmatrix} = \begin{bmatrix}\mathbf{1}'\mathbf{1} & \mathbf{1}'\mathbf{X}_c \\ \mathbf{X}_c'\mathbf{1}' & \mathbf{X}_c'\mathbf{X}_c\end{bmatrix} = \begin{bmatrix}n & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_c'\mathbf{X}_c\end{bmatrix}$$

using the fact that each column of $\mathbf{X}_c$ has mean 0. Therefore,

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 1/n & \mathbf{0} \\ \mathbf{0} & (\mathbf{X}_c'\mathbf{X}_c)^{-1} \end{bmatrix}$$

and so

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{bmatrix} 1/n & \mathbf{0} \\ \mathbf{0} & (\mathbf{X}_c'\mathbf{X}_c)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{1}' \\ \mathbf{X}_c' \end{bmatrix} \mathbf{y} = \begin{bmatrix} \sum y_i/n \\ (\mathbf{X}_c'\mathbf{X}_c)^{-1}\mathbf{X}_c'\mathbf{y} \end{bmatrix}$$

implying that

$$\hat{\beta}_0 = \frac{1}{n}\sum y_i = \bar{y}$$

as required.

(5.16) (a) Since the possible values of z are 0 or 1, the parameter β_3 represents the change in the yield of the chemical reaction (y) due to the second catalyst at a fixed temperature level (x). Note that if $\beta_3 > 0$, then this would imply that the yield increases due to the second catalyst, while $\beta_3 < 0$ implies that the yield decreases due to the second catalyst.

(b) From the data given, we find $\hat{\beta}_2 = 0.41$ and $\mathrm{SE}(\hat{\beta}_2) = 0.11$. Therefore, a 95% confidence interval for β_2 is

$$\hat{\beta}_2 \pm t(0.025; 26)\mathrm{SE}(\hat{\beta}_2) = 0.41 \pm (2.056)(0.11) = [0.184, 0.636].$$

Note that the degrees of freedom are $df = n - p - 1 = 30 - 3 - 1 = 26$.

(c) (i) Since the vector of errors $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2 I)$, we conclude that the vector of least squares estimates $[\hat{\beta}_0,\ \hat{\beta}_1,\ \hat{\beta}_2,\ \hat{\beta}_3]'$ has a multivariate normal distribution. Hence, any uncorrelated components are necessarily independent. Thus, since $\mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_3) = 0$, we conclude that $\hat{\beta}_1$ and $\hat{\beta}_3$ are independent.

(c) (ii) When the standard temperature ($x = 0$) and catalyst 2 ($z = 1$) are used, the expected yield is

$$\hat{\mu} = \mathbb{E}(y) = \hat{\beta}_0 + \hat{\beta}_3 = 29.83 - 0.32 = 29.51.$$

Since the residual sum of squares is SSE = 25.05, we conclude that

$$s^2 = \frac{\mathrm{SSE}}{n - p - 1} = \frac{25.05}{26} \doteq 0.96.$$

Thus, a 95% confidence interval for $\hat{\mu}$ is

$$\hat{\mu} \pm t(0.025; 26)s = 29.51 \pm (2.056)\sqrt{0.96} = [27.495,\ 31.525].$$

Note that the degrees of freedom are $df = n - p - 1 = 30 - 3 - 1 = 26$.