Homework 5 Solution

Question 1

(1)

$$Z^\top Z = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 26 & 0 \\ 0 & 0 & 12 \end{bmatrix}$$

$$Z^\top Y = \begin{bmatrix} 0 \\ 8 \\ 0 \end{bmatrix}$$

Therefore,

$$\begin{aligned} \hat{\beta} &= (Z^\top Z)^{-1} Z^\top Y \\ &= \begin{bmatrix} \frac{1}{7} & 0 & 0 \\ 0 & \frac{1}{26} & 0 \\ 0 & 0 & \frac{1}{12} \end{bmatrix} \begin{bmatrix} 0 \\ 8 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ \frac{4}{13} \\ 0 \end{bmatrix} \end{aligned}$$

(2)

$$\begin{aligned} R^2 &= 1 - \frac{\|\hat{\epsilon}\|^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\ &= 1 - \frac{\|Y - Z\hat{\beta}\|^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\ &= 0.6154 \end{aligned}$$

(3)

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{n - r - 1} \|\hat{\epsilon}\|^2 \\ &= 0.3846 \end{aligned}$$

$$\begin{aligned} \hat{Cov}(\hat{\beta}) &= \hat{\sigma}^2 (Z^\top Z)^{-1} \\ &= \begin{bmatrix} 0.0549 & 0 & 0 \\ 0 & 0.0148 & 0 \\ 0 & 0 & 0.0321 \end{bmatrix} \end{aligned}$$

(4) The 95% confidence interval for β_1 is

$$\begin{aligned}
&[\hat{\beta}_1 - \hat{\sigma}\sqrt{\omega_{11}}t_{n-r-1}(\frac{0.05}{2}), \hat{\beta}_1 + \hat{\sigma}\sqrt{\omega_{11}}t_{n-r-1}(\frac{0.05}{2})] \\
&= [\frac{4}{13} - \sqrt{0.3846}\sqrt{\frac{1}{26}}t_{7-2-1}(\frac{0.05}{2}), \frac{4}{13} + \sqrt{0.3846}\sqrt{\frac{1}{26}}t_{7-2-1}(\frac{0.05}{2})] \\
&= [-0.0300, 0.6454]
\end{aligned}$$

(5) The 95% confidence intervals for $\beta_j,\ j = 0, 1, 2$ based on confidence region are

$$[\hat{\beta}_j - \hat{\sigma}\sqrt{\omega_{jj}}\sqrt{(r+1)F_{r+1,n-r-1}(0.05)}, \hat{\beta}_j + \hat{\sigma}\sqrt{\omega_{jj}}\sqrt{(r+1)F_{r+1,n-r-1}(0.05)}]$$

The results are

$$\begin{aligned}
\beta_0 &\in [-1.0423, 1.0423] \\
\beta_1 &\in [-0.2332, 0.8485] \\
\beta_2 &\in [-0.7961, 0.7961]
\end{aligned}$$

(6) The 95% confidence intervals for $\beta_j,\ j = 0, 1, 2$ based on Bonferroni correction are

$$[\hat{\beta}_j - \hat{\sigma}\sqrt{\omega_{jj}}t_{n-r-1}(\frac{0.05}{2(r+1)}), \hat{\beta}_j + \hat{\sigma}\sqrt{\omega_{jj}}t_{n-r-1}(\frac{0.05}{2(r+1)})]$$

The results are

$$\begin{aligned}
\beta_0 &\in [-0.9284, 0.9284] \\
\beta_1 &\in [-0.1740, 0.7894] \\
\beta_2 &\in [-0.7091, 0.7091]
\end{aligned}$$

(7) Let $C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. The F-test statistic is

$$\begin{aligned}
&\frac{1}{\hat{\sigma}^2}(C\hat{\beta})^\top(C(Z^\top Z)^{-1}C^\top)^{-1}(C\hat{\beta}) \\
&= \frac{1}{\hat{\sigma}^2}\hat{\beta}_{(2)}^T\Omega_{22}^{-1}\hat{\beta}_{(2)} \\
&= \frac{1}{0.3846}\begin{bmatrix} \frac{4}{13} & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{26} & 0 \\ 0 & \frac{1}{12} \end{bmatrix}^{-1}\begin{bmatrix} \frac{4}{13} \\ 0 \end{bmatrix} \\
&= 6.4
\end{aligned}$$

The critical value is

$$\begin{aligned}
(r-q)F_{r-q,n-r-1}(\alpha) &= (2-0)F_{2-0,7-2-1}(0.05) \\
&= 13.8885
\end{aligned}$$

As $6.4 < 13.8885$, we do not reject the H_0.

(8)

$$\bar{z}_1 = 0$$
$$\bar{z}_2 = 0$$

So

$$\vec{z}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

The confidence interval for the mean response $E(Y_0)$ given $\vec{z}_0$ is given by

$$[\vec{z}_0^\top \hat{\beta} - \hat{\sigma} t_{n-r-1}(\frac{\alpha}{2})\sqrt{\vec{z}_0^\top (Z^\top Z)^{-1}\vec{z}_0}, \vec{z}_0^\top \hat{\beta} + \hat{\sigma} t_{n-r-1}(\frac{\alpha}{2})\sqrt{\vec{z}_0^\top (Z^\top Z)^{-1}\vec{z}_0}]$$

The result is

$$[-0.6508, 0.6508]$$

(9) The prediction interval for the Y_0 given $\vec{z}_0$ is given by

$$[\vec{z}_0^\top \hat{\beta} - \hat{\sigma} t_{n-r-1}(\frac{\alpha}{2})\sqrt{1 + \vec{z}_0^\top (Z^\top Z)^{-1}\vec{z}_0}, \vec{z}_0^\top \hat{\beta} + \hat{\sigma} t_{n-r-1}(\frac{\alpha}{2})\sqrt{1 + \vec{z}_0^\top (Z^\top Z)^{-1}\vec{z}_0}]$$

The result is

$$[-1.8408, 1.8408]$$

Question 2

(1)

$$\begin{aligned} H^\top &= (Z(Z^\top Z)^{-1}Z^\top)^\top \\ &= (Z^\top)^\top((Z^\top Z)^{-1})^\top Z^\top \\ &= Z((Z^\top Z)^\top)^{-1}Z^\top \\ &= Z(Z^\top Z)^{-1}Z^\top \\ &= H \end{aligned}$$

$$\begin{aligned} (I-H)^\top &= I^\top - H^\top \\ &= I - H \end{aligned}$$

(2)

$$\begin{aligned} H^2 &= (Z(Z^\top Z)^{-1}Z^\top)(Z(Z^\top Z)^{-1}Z^\top) \\ &= Z(Z^\top Z)^{-1}(Z^\top Z)(Z^\top Z)^{-1}Z^\top \\ &= Z(Z^\top Z)^{-1}Z^\top \\ &= H \end{aligned}$$

$$
\begin{aligned}
(I-H)^2 &= (I-H)(I-H) \\
&= I - H - H + H^2 \\
&= I - 2H + H \\
&= I - H
\end{aligned}
$$

$$
\begin{aligned}
H(I-H) &= H - H^2 \\
&= H - H \\
&= \mathbf{0}
\end{aligned}
$$

(3) Suppose the spectral decomposition of H is

$$
H = P\Lambda P^\top
$$

where P is an orthogonal matrix and Λ is a diagonal matrix.
By $H^2 = H$, we have

$$
\begin{aligned}
P\Lambda P^\top P\Lambda P^\top &= P\Lambda P^\top \\
P\Lambda I \Lambda P^\top &= P\Lambda P^\top \\
P\Lambda^2 P^\top &= P\Lambda P^\top \\
P^\top (P\Lambda^2 P^\top) P &= P^\top (P\Lambda P^\top) P \\
\Lambda^2 &= \Lambda
\end{aligned}
$$

Because Λ is a diagonal matrix, by $\Lambda^2 = \Lambda$, all the eigenvalues of H must be 0 or 1.

$$
\begin{aligned}
I - H &= PP^\top - P\Lambda P^\top \\
&= P(I-\Lambda)P^\top
\end{aligned}
$$

So all the eigenvalues of $(I-H)$ must be 0 or 1.

(4) As all the eigenvalues of H are nonnegative, by the definition of positive semi-definite, H is a positive semi-definite matrix. Similarly, $(I-H)$ is also a positive semi-definite matrix.

(5)

$$
\begin{aligned}
HZ &= (Z(Z^\top Z)^{-1}Z^\top)Z \\
&= Z(Z^\top Z)^{-1}(Z^\top Z) \\
&= Z
\end{aligned}
$$

$$
\begin{aligned}
(I-H)Z &= IZ - HZ \\
&= Z - Z \\
&= \mathbf{0}
\end{aligned}
$$

Question 3 From Question 2 (5), we know $HZ = Z$. Then

$$\begin{aligned} H\begin{bmatrix} Z_{(1)} & Z_{(2)} \end{bmatrix} &= \begin{bmatrix} Z_{(1)} & Z_{(2)} \end{bmatrix} \\ \begin{bmatrix} HZ_{(1)} & HZ_{(2)} \end{bmatrix} &= \begin{bmatrix} Z_{(1)} & Z_{(2)} \end{bmatrix} \end{aligned}$$

Therefore, $HZ_{(1)} = Z_{(1)}$.

$$\begin{aligned} HH_{(red)} &= HZ_{(1)}(Z_{(1)}^\top Z_{(1)})^{-1}Z_{(1)}^\top \\ &= (HZ_{(1)})(Z_{(1)}^\top Z_{(1)})^{-1}Z_{(1)}^\top \\ &= Z_{(1)}(Z_{(1)}^\top Z_{(1)})^{-1}Z_{(1)}^\top \\ &= H_{(red)} \end{aligned}$$

$$\begin{aligned} H_{(red)}H &= Z_{(1)}(Z_{(1)}^\top Z_{(1)})^{-1}Z_{(1)}^\top H \\ &= Z_{(1)}(Z_{(1)}^\top Z_{(1)})^{-1}(H^\top Z_{(1)})^\top \\ &= Z_{(1)}(Z_{(1)}^\top Z_{(1)})^{-1}(HZ_{(1)})^\top \\ &= Z_{(1)}(Z_{(1)}^\top Z_{(1)})^{-1}Z_{(1)}^\top \\ &= H_{(red)} \end{aligned}$$

Here we also use $H^\top = H$ from Question 2 (1).

Question 4

(1)

$$\begin{aligned} W &= \begin{bmatrix} 1 & \vec{w}_1^\top \\ 1 & \vec{w}_2^\top \\ \cdot & \cdot \\ 1 & \vec{w}_n^\top \end{bmatrix} \\ &= \begin{bmatrix} 1 & (C\vec{z}_1)^\top \\ 1 & (C\vec{z}_2)^\top \\ \cdot & \cdot \\ 1 & (C\vec{z}_n)^\top \end{bmatrix} \\ &= \begin{bmatrix} 1 & \vec{z}_1^\top \\ 1 & \vec{z}_2^\top \\ \cdot & \cdot \\ 1 & \vec{z}_n^\top \end{bmatrix} \begin{bmatrix} 1 & \vec{0}^\top \\ \vec{0} & C^\top \end{bmatrix} \\ &= Z\tilde{C}^\top \end{aligned}$$

(2)

$$\begin{aligned} \hat{\gamma} &= (W^\top W)^{-1}W^\top Y \\ &= (\tilde{C}Z^\top Z\tilde{C}^\top)^{-1}\tilde{C}Z^\top Y \\ &= (\tilde{C}^\top)^{-1}(Z^\top Z)^{-1}\tilde{C}^{-1}\tilde{C}Z^\top Y \\ &= (\tilde{C}^\top)^{-1}(Z^\top Z)^{-1}Z^\top Y \\ &= (\tilde{C}^\top)^{-1}\hat{\beta} \end{aligned}$$

$$
\begin{aligned}
\hat{\epsilon}_w &= Y - W\hat{\gamma} \\
&= Y - Z\tilde{C}^\top (\tilde{C}^\top)^{-1}\hat{\beta} \\
&= Y - Z\hat{\beta} \\
&= \hat{\epsilon}_z
\end{aligned}
$$

(3)

$$
\begin{aligned}
R_w^2 &= 1 - \frac{\|\hat{\epsilon}_w\|^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\
&= 1 - \frac{\|\hat{\epsilon}_z\|^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\
&= R_z^2
\end{aligned}
$$

(4)

$$
\begin{aligned}
\hat{\sigma}_w^2 &= \frac{\|\hat{\epsilon}_w\|^2}{n-r-1} \\
&= \frac{\|\hat{\epsilon}_z\|^2}{n-r-1} \\
&= \hat{\sigma}_z^2
\end{aligned}
$$

Method 1:

Let $R = \begin{bmatrix} \vec{0} & I_r \end{bmatrix}$, then it is equivalent to test $H_0 : R\beta = \vec{0}$ and $H_0 : R\gamma = \vec{0}$.

$$
\begin{aligned}
R(\tilde{C}^\top)^{-1}\hat{\beta} &= \begin{bmatrix} \vec{0} & I_r \end{bmatrix} \begin{bmatrix} 1 & \vec{0}^\top \\ \vec{0} & (C^{-1})^\top \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_{(2)} \end{bmatrix} \\
&= (C^{-1})^\top \hat{\beta}_{(2)}
\end{aligned}
$$

where $\hat{\beta}_{(2)} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ . \\ \hat{\beta}_r \end{bmatrix}$

And

$$
\begin{aligned}
&(R((Z\tilde{C}^\top)^\top (Z\tilde{C}^\top))^{-1} R^\top)^{-1} \\
&= (R(\tilde{C}Z^\top Z\tilde{C}^\top)^{-1} R^\top)^{-1} \\
&= (R(\tilde{C}^\top)^{-1}(Z^\top Z)^{-1}\tilde{C}^{-1}R^T)^{-1} \\
&= (\begin{bmatrix} \vec{0} & I_r \end{bmatrix} \begin{bmatrix} 1 & \vec{0}^\top \\ \vec{0} & (C^{-1})^\top \end{bmatrix} \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \begin{bmatrix} 1 & \vec{0}^\top \\ \vec{0} & C^{-1} \end{bmatrix} \begin{bmatrix} \vec{0}^\top \\ I_r \end{bmatrix})^{-1} \\
&= ((C^{-1})^\top \Omega_{22} C^{-1})^{-1} \\
&= C\Omega_{22}^{-1}C^\top
\end{aligned}
$$

Therefore, the test statistic for $H_0 : R\gamma = \vec{0}$ is

$$
\begin{aligned}
&\frac{1}{\hat{\sigma}_w^2}(R\hat{\gamma})^\top (R(W^\top W)^{-1}R^\top)^{-1}(R\hat{\gamma}) \\
&= \frac{1}{\hat{\sigma}_z^2}(R(\tilde{C}^\top)^{-1}\hat{\beta})^\top (R((Z\tilde{C}^\top)^\top (Z\tilde{C}^\top))^{-1}R^\top)^{-1}(R(\tilde{C}^\top)^{-1}\hat{\beta}) \\
&= \frac{1}{\hat{\sigma}_z^2}((C^{-1})^\top \hat{\beta}_{(2)})^\top C\Omega_{22}^{-1}C^\top ((C^{-1})^\top \hat{\beta}_{(2)}) \\
&= \frac{1}{\hat{\sigma}_z^2}\hat{\beta}_{(2)}^\top \Omega_{22}^{-1}\hat{\beta}_{(2)}
\end{aligned}
$$

which is the same as the test statistic for $H_0 : R\beta = \vec{0}$.

Method 2:

Here we use the residuals to derive the F-test statistic. We can easily find that

$$\hat{\epsilon}_{w,(red)} = \hat{\epsilon}_{z,(red)}$$

because the reduced models are the same.

Then the test statistic for $H_0 : \gamma_1 = \gamma_2 = ... = \gamma_r = 0$ is

$$
\begin{aligned}
&\frac{1}{\hat{\sigma}_w^2}(\|\hat{\epsilon}_{w,(red)}\|^2 - \|\hat{\epsilon}_w\|^2) \\
&= \frac{1}{\hat{\sigma}_z^2}(\|\hat{\epsilon}_{z,(red)}\|^2 - \|\hat{\epsilon}_z\|^2)
\end{aligned}
$$

which is the same as the test statistic for $H_0 : \beta_1 = \beta_2 = ... = \beta_r = 0$.

(5)

$$
\begin{aligned}
\vec{w}_0 &= \begin{bmatrix} 1 \\ w_{01} \\ . \\ w_{0r} \end{bmatrix} \\
&= \begin{bmatrix} 1 & \vec{0}^\top \\ \vec{0} & C \end{bmatrix} \begin{bmatrix} 1 \\ z_{01} \\ . \\ z_{0r} \end{bmatrix} \\
&= \tilde{C}\vec{z}_0
\end{aligned}
$$

The center for the prediction interval for Y_0 based on $\hat{\gamma}$ is

$$
\begin{aligned}
\vec{w}_0^\top \hat{\gamma} &= (\tilde{C}\vec{z}_0)^\top (\tilde{C}^\top)^{-1}\hat{\beta} \\
&= \vec{z}_0^\top \tilde{C}^\top (\tilde{C}^\top)^{-1}\hat{\beta} \\
&= \vec{z}_0^\top \hat{\beta}
\end{aligned}
$$

which is the same as the center for the prediction interval for Y_0 based on $\hat{\beta}$.

The half length of the prediction interval for Y_0 based on $\hat{\gamma}$ is

$$\begin{aligned}
&\hat{\sigma}_w t_{n-r-1}(\frac{\alpha}{2})\sqrt{1+\vec{w}_0^\top (W^\top W)^{-1}\vec{w}_0} \\
&= \hat{\sigma}_z t_{n-r-1}(\frac{\alpha}{2})\sqrt{1+(\tilde{C}\vec{z}_0)^\top((Z\tilde{C}^\top)^\top(Z\tilde{C}^\top))^{-1}(\tilde{C}\vec{z}_0)} \\
&= \hat{\sigma}_z t_{n-r-1}(\frac{\alpha}{2})\sqrt{1+\vec{z}_0^\top\tilde{C}^\top(\tilde{C}Z^\top Z\tilde{C}^\top)^{-1}\tilde{C}\vec{z}_0} \\
&= \hat{\sigma}_z t_{n-r-1}(\frac{\alpha}{2})\sqrt{1+\vec{z}_0^\top\tilde{C}^\top(\tilde{C}^\top)^{-1}(Z^\top Z)^{-1}\tilde{C}^{-1}\tilde{C}\vec{z}_0} \\
&= \hat{\sigma}_z t_{n-r-1}(\frac{\alpha}{2})\sqrt{1+\vec{z}_0^\top(Z^\top Z)^{-1}\vec{z}_0}
\end{aligned}$$

which is the same as the half length for the prediction interval for Y_0 based on $\hat{\beta}$.

In all, the prediction interval for Y_0 based on $\hat{\beta}$ and the one based on $\hat{\gamma}$ are the same.

Question 5

(1)

$$\hat{\beta} = \begin{bmatrix} 30.967 \\ 2.634 \\ 0.045 \end{bmatrix}$$

(2)

$$R^2 = 0.834$$

(3)

$$\hat{\sigma}^2 = 12.059$$

$$\begin{aligned}
\hat{Cov}(\hat{\beta}) &= \hat{\sigma}^2(Z^\top Z)^{-1} \\
&= \begin{bmatrix} 62.129 & 3.068 & -1.765 \\ 3.068 & 0.617 & -0.207 \\ -1.765 & -0.207 & 0.081 \end{bmatrix}
\end{aligned}$$

(4) The 95% confidence interval for β_1 is

$$[0.977, 4.292]$$

(5) The 95% confidence intervals for $\beta_j,\ j = 0, 1, 2$ based on confidence region are

$$\begin{aligned}
\beta_0 &\in [6.557, 55.376] \\
\beta_1 &\in [0.202, 5.067] \\
\beta_2 &\in [-0.838, 0.928]
\end{aligned}$$

(6) The 95% confidence intervals for $\beta_j,\ j = 0, 1, 2$ based on Bonferroni correction are

$$\begin{aligned}
\beta_0 &\in [10.039, 51.894] \\
\beta_1 &\in [0.549, 4.720] \\
\beta_2 &\in [-0.712, 0.802]
\end{aligned}$$

(7) The F-test statistic is

$$\begin{aligned}
&\frac{1}{\hat{\sigma}^2}(C\hat{\beta})^\top (C(Z^\top Z)^{-1}C^\top)^{-1}(C\hat{\beta}) \\
&= 85.655
\end{aligned}$$

The critical value is

$$\begin{aligned}
(r-q)F_{r-q,n-r-1}(\alpha) &= (2-0)F_{2-0,20-2-1}(0.05) \\
&= 7.183
\end{aligned}$$

As $85.655 > 7.183$, we reject the H_0.

(8)

$$\begin{aligned}
\bar{z}_1 &= 16.222 \\
\bar{z}_2 &= 63.065
\end{aligned}$$

So

$$\vec{z}_0 = \begin{bmatrix} 1 \\ 16.222 \\ 63.065 \end{bmatrix}$$

The confidence interval for the mean response $E(Y_0)$ given $\vec{z}_0$ is given by

$$[\vec{z}_0^\top \hat{\beta} - \hat{\sigma} t_{n-r-1}(\frac{\alpha}{2})\sqrt{\vec{z}_0^\top (Z^\top Z)^{-1}\vec{z}_0}, \vec{z}_0^\top \hat{\beta} + \hat{\sigma} t_{n-r-1}(\frac{\alpha}{2})\sqrt{\vec{z}_0^\top (Z^\top Z)^{-1}\vec{z}_0}]$$

The result is

$$[74.9118, 78.1882]$$

(9) The prediction interval for the Y_0 given $\vec{z}_0$ is given by

$$[\vec{z}_0^\top \hat{\beta} - \hat{\sigma} t_{n-r-1}(\frac{\alpha}{2})\sqrt{1 + \vec{z}_0^\top (Z^\top Z)^{-1}\vec{z}_0}, \vec{z}_0^\top \hat{\beta} + \hat{\sigma} t_{n-r-1}(\frac{\alpha}{2})\sqrt{1 + \vec{z}_0^\top (Z^\top Z)^{-1}\vec{z}_0}]$$

The result is

$$[69.0427, 84.0573]$$