Sample Final Solution

Question 1

(1) We have the pooled sample covariance matrix

$$\boldsymbol{S}_{\text{pooled}} = \frac{n_1 - 1}{n_1 + n_2 - 2} S_1 + \frac{n_2 - 1}{n_1 + n_2 - 2} S_2 = \frac{1}{2} \begin{bmatrix} 16 & 8 \\ 8 & 16 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 16 & 8 \\ 8 & 16 \end{bmatrix} = \begin{bmatrix} 16 & 8 \\ 8 & 16 \end{bmatrix}$$

Moreover, we have $\bar{\vec{x}}_1 - \bar{\vec{x}}_2 = \begin{bmatrix} -3 \\ 3 \end{bmatrix}$. Then

$$T^2 = \left(\bar{\vec{x}}_1 - \bar{\vec{x}}_2\right)^\top \left(\left(\frac{1}{n_1} + \frac{1}{n_2} \right) S_{\text{pooled}} \right)^{-1} \left(\bar{\vec{x}}_1 - \bar{\vec{x}}_2\right) = \frac{81}{4}$$

In contrast, the critical value is

$$\frac{(n_1 + n_2 - 2)\, p}{n_1 + n_2 - 1 - p} F_{p, n_1 + n_2 - 1 - p}(\alpha) = \frac{68}{33} F_{2,33}(0.05) = 6.7689$$

Here $T^2 > \frac{(n_1+n_2-2)p}{n_1+n_2-1-p} F_{p,n_1+n_2-1-p}(\alpha)$. Therefore, we reject H_0 at the level of $\alpha = 0.05$.

(2) The $(1 - \alpha)$ confidence ellipse of $\vec{\mu}_1 - \vec{\mu}_2$ is

$$\begin{aligned} &((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - (\vec{\mu}_1 - \vec{\mu}_2))^\top ((\frac{1}{n_1} + \frac{1}{n_2}) S_{\text{pooled}})^{-1} ((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - (\vec{\mu}_1 - \vec{\mu}_2)) \\ &\leq \frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - 1 - p} F_{p, n_1+n_2-1-p}(\alpha) \end{aligned}$$

The spectral decomposition of $((\frac{1}{n_1} + \frac{1}{n_2}) S_{\text{pooled}})$ is

$$((\frac{1}{n_1} + \frac{1}{n_2}) S_{\text{pooled}}) = \frac{8}{3} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}^\top + \frac{8}{9} \begin{bmatrix} \frac{-\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{-\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}^\top$$

Also, $\bar{\vec{x}}_1 - \bar{\vec{x}}_2 = \begin{bmatrix} -3 \\ 3 \end{bmatrix}$, $c = \sqrt{\frac{(n_1+n_2-2)p}{n_1+n_2-1-p} F_{p,n_1+n_2-1-p}(\alpha)} = 2.6017$.

Therefore, the ellipse has center $\begin{bmatrix} -3 \\ 3 \end{bmatrix}$, with axes of directions $\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$ and $\begin{bmatrix} \frac{-\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$, and half axis lengths $c\sqrt{\lambda_1} = 4.2486$ and $c\sqrt{\lambda_2} = 2.4529$.

(3) The formulas for simultaneous $1 - \alpha$ confidence intervals for $\mu_{1j} - \mu_{2j}$, $j = 1, \ldots, p$ with Bonferroni correction are

$$\begin{aligned} &(\bar{x}_{1j} - \bar{x}_{2j}) - s_{\text{pooled},j} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2}(\frac{\alpha}{2p}) \\ &\leq \mu_{1j} - \mu_{2j} \\ &\leq (\bar{x}_{1j} - \bar{x}_{2j}) + s_{\text{pooled},j} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2}(\frac{\alpha}{2p}) \end{aligned}$$

We have $\bar{x}_{11} - \bar{x}_{21} = -3$, $s_{\text{pooled},1} = \sqrt{(\boldsymbol{S}_{\text{pooled}})_{11}} = 4$, $\bar{x}_{12} - \bar{x}_{22} = 3$,$s_{\text{pooled},2} = \sqrt{(\boldsymbol{S}_{\text{pooled}})_{22}} = 4$ and $t_{n_1+n_2-2}(\frac{\alpha}{2p}) = t_{34}(0.0125) = 2.345$. The resulting simultaneous confidence intervals are

$$-6.1267 \le \mu_{11} - \mu_{21} \le 0.1267, \quad -0.1267 \le \mu_{12} - \mu_{22} \le 6.1267$$

Question 2

(1)

$$Z^\top Z = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 26 & -24 \\ 0 & -24 & 26 \end{bmatrix}$$

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

And

$$\begin{aligned}(Z^\top Z)^{-1} &= \begin{bmatrix} \frac{1}{7} & 0 & 0 \\ 0 & \frac{26}{26^2-24^2} & \frac{-24}{26^2-24^2} \\ 0 & \frac{-24}{26^2-24^2} & \frac{26}{26^2-24^2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{7} & 0 & 0 \\ 0 & \frac{26}{100} & \frac{-24}{100} \\ 0 & \frac{-24}{100} & \frac{26}{100} \end{bmatrix}\end{aligned}$$

Therefore,

$$\begin{aligned}\hat{\beta} &= (Z^\top Z)^{-1} Z^\top Y \\ &= \begin{bmatrix} \frac{1}{7} & 0 & 0 \\ 0 & \frac{26}{100} & \frac{-24}{100} \\ 0 & \frac{-24}{100} & \frac{26}{100} \end{bmatrix} \begin{bmatrix} 0 \\ 8 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ \frac{4}{25} \\ -\frac{4}{25} \end{bmatrix}\end{aligned}$$

(2)

$$\begin{aligned}R^2 &= 1 - \frac{\|\hat{\epsilon}\|^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\ &= 1 - \frac{\|Y - Z\hat{\beta}\|^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \\ &= 0.64\end{aligned}$$

(3)

$$\begin{aligned}\hat{\sigma}^2 &= \frac{1}{n-r-1} \|\hat{\epsilon}\|^2 \\ &= 0.36\end{aligned}$$

$$
\begin{aligned}
\hat{Cov}(\hat{\beta}) &= \hat{\sigma}^2 (Z^\top Z)^{-1} \\
&= \begin{bmatrix} 0.0514 & 0 & 0 \\ 0 & 0.0936 & 0.0864 \\ 0 & 00864 & 0.0936 \end{bmatrix}
\end{aligned}
$$

(4) The 95% confidence interval for β_1 is

$$
\begin{aligned}
&[\hat{\beta}_1 - \hat{\sigma}\sqrt{\omega_{11}} t_{n-r-1}(\frac{0.05}{2}), \hat{\beta}_1 + \hat{\sigma}\sqrt{\omega_{11}} t_{n-r-1}(\frac{0.05}{2})] \\
&= [\frac{4}{25} - \sqrt{0.36}\sqrt{\frac{26}{100}} t_{7-2-1}(\frac{0.05}{2}), \frac{4}{25} + \sqrt{0.36}\sqrt{\frac{26}{100}} t_{7-2-1}(\frac{0.05}{2})] \\
&= [-0.6894, 1.0094]
\end{aligned}
$$

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

$$
\begin{aligned}
&\frac{1}{\hat{\sigma}^2} (C\hat{\beta})^\top (C(Z^\top Z)^{-1} C^\top)^{-1} (C\hat{\beta}) \\
&= \frac{1}{\hat{\sigma}^2} \hat{\beta}_{(2)}^T \Omega_{22}^{-1} \hat{\beta}_{(2)} \\
&= \frac{1}{0.36} \begin{bmatrix} 0.16 & -0.16 \end{bmatrix} \begin{bmatrix} \frac{26}{100} & \frac{-24}{100} \\ \frac{-24}{100} & \frac{26}{100} \end{bmatrix}^{-1} \begin{bmatrix} 0.16 \\ -0.16 \end{bmatrix} \\
&= 2.56
\end{aligned}
$$

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 $2.56 < 13.8885$, we do not reject the H_0.

(6)

$$
\begin{aligned}
\bar{z}_1 &= 0 \\
\bar{z}_2 &= 0
\end{aligned}
$$

So

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

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

$$
[-1.7809, 1.7809]
$$

Question 3

(1)

$$S = \frac{1}{6}\begin{bmatrix} 26 & -24 \\ -24 & 26 \end{bmatrix}$$

The spectral decomposition of S is

$$S = \frac{25}{3}\begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}\begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}^\top + \frac{1}{3}\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}^\top$$

Based on $\vec{u}_1 = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$ and $\vec{u}_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$, the principal components are

$$\hat{y}_1 = -\frac{\sqrt{2}}{2}x_1 + \frac{\sqrt{2}}{2}x_2$$
$$\hat{y}_2 = \frac{\sqrt{2}}{2}x_1 + \frac{\sqrt{2}}{2}x_2$$

(2) The proportion of total sample variance due to the first sample principal component is

$$\frac{\hat{\lambda}_1}{\hat{\lambda}_1 + \hat{\lambda}_2} = \frac{\frac{25}{3}}{\frac{25}{3} + \frac{1}{3}} = \frac{25}{26} = 96.15\%$$

(3) Based on loadings, as $|v_{11}| = \frac{\sqrt{2}}{2} = |v_{12}| = \frac{\sqrt{2}}{2}$, these two variates contribute the same to the determination of the first sample principal component.

(4) Based on sample correlations,

$$|v_{11}\sqrt{\frac{\hat{\lambda}_1}{s_{11}}}| = |-\frac{\sqrt{2}}{2}\sqrt{\frac{\frac{25}{3}}{\frac{26}{6}}}| = 0.981$$
$$|v_{12}\sqrt{\frac{\hat{\lambda}_1}{s_{22}}}| = |\frac{\sqrt{2}}{2}\sqrt{\frac{\frac{25}{3}}{\frac{26}{6}}}| = 0.981$$

these two variates contribute the same to the determination of the first sample principal component.

Question 4 For any $\vec{x}_p$ in P, as $\vec{x}_p \in L_{12}$, we have

$$(\vec{x}_p - \bar{\vec{x}}_1)^\top S_{pooled}^{-1}(\vec{x}_p - \bar{\vec{x}}_1) = (\vec{x}_p - \bar{\vec{x}}_2)^\top S_{pooled}^{-1}(\vec{x}_p - \bar{\vec{x}}_2)$$

which means for $\vec{x}_p$, the Mahalanobis distance to $\bar{\vec{x}}_1$ is equal to that to $\bar{\vec{x}}_2$.
Similarly, $\vec{x}_p \in L_{23}$, we have

$$(\vec{x}_p - \bar{\vec{x}}_2)^\top S_{pooled}^{-1}(\vec{x}_p - \bar{\vec{x}}_2) = (\vec{x}_p - \bar{\vec{x}}_3)^\top S_{pooled}^{-1}(\vec{x}_p - \bar{\vec{x}}_3)$$

which means for $\vec{x}_p$, the Mahalanobis distance to $\bar{\vec{x}}_2$ is equal to that to $\bar{\vec{x}}_3$. Therefore, when we compare π_1 and π_3, as

$$(\vec{x}_p - \bar{\vec{x}}_1)^\top S_{pooled}^{-1}(\vec{x}_p - \bar{\vec{x}}_1) = (\vec{x}_p - \bar{\vec{x}}_3)^\top S_{pooled}^{-1}(\vec{x}_p - \bar{\vec{x}}_3)$$

we know that $\vec{x}_p \in L_{31}$. Then then L_{31} also goes through P.

Question 5 For a regression model, we have

$$\hat{e} = (I - H)Y$$

So SSE can be written as

$$\begin{aligned}
\text{SSE} &= \sum_{i=1}^{n} \hat{e}_i^2 \\
&= \hat{e}^\top \hat{e} \\
&= Y^\top (I - H)^\top (I_H) Y \\
&= Y^\top (I - H)(I - H) Y \\
&= Y^\top (I - H) Y
\end{aligned}$$

here we use the results in homework 5.2.
And SSTO is

$$\text{SSTO} = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

In this problem as $\bar{y} = 0$, we have

$$\begin{aligned}
\text{SSTO} &= \sum_{i=1}^{n} y_i^2 \\
&= Y^\top Y
\end{aligned}$$

Therefore, R^2 can be written as

$$\begin{aligned}
R2 &= 1 - \frac{\text{SSE}}{\text{SSTO}} \\
&= 1 - \frac{Y^\top (I - H) Y}{Y^\top Y} \\
&= 1 - \frac{Y^\top Y - Y^\top H Y}{Y^\top Y} \\
&= \frac{Y^\top H Y}{Y^\top Y}
\end{aligned}$$

Then let's look at the properties for $\hat{Z}_1$ and $\hat{Z}_2$.

As $\hat{Z}_1$ and $\hat{Z}_2$ are sample principal components, we know

$$\hat{z}_{i1} = v_{11}x_{i1} + v_{12}x_{i2}$$

then we have

$$\begin{aligned}\bar{z}_1 &= v_{11}\bar{x}_1 + v_{12}\bar{x}_2 \\ &= 0\end{aligned}$$

Similarly,

$$\begin{aligned}\bar{z}_2 &= v_{21}\bar{x}_1 + v_{22}\bar{x}_2 \\ &= 0\end{aligned}$$

Also, we know that their sample covariance must be 0, that is to say,

$$\begin{aligned}0 &= \frac{1}{n-1}\sum_{i=1}^{n}(\hat{z}_{i1} - \bar{z}_1)(\hat{z}_{i2} - \bar{z}_2) \\ &= \frac{1}{n-1}\sum_{i=1}^{n}\hat{z}_{i1}\hat{z}_{i2} \\ &= \frac{1}{n-1}\hat{Z}_1^\top \hat{Z}_2\end{aligned}$$

So we have

$$\hat{Z}_1^\top \hat{Z}_2 = 0$$

Now let's look at the relationship for these models.

(a) For model: $y_i = \beta_1\hat{z}_{i1} + \epsilon_i$, the design matrix is $\hat{Z}_1$. Then,

$$H_1 = \hat{Z}_1(\hat{Z}_1^\top \hat{Z}_1)^{-1}\hat{Z}_1^\top$$

(b) Similarly, for model: $y_i = \beta_2\hat{z}_{i2} + \epsilon_i$, we have $H_2 = \hat{Z}_2(\hat{Z}_2^\top \hat{Z}_2)^{-1}\hat{Z}_2^\top$.

(c) For the full model: $y_i = \beta_1\hat{z}_{i1} + \beta_2\hat{z}_{i2} + \epsilon_i$, the design matrix

$$Z = \begin{bmatrix}\hat{Z}_1 & \hat{Z}_2\end{bmatrix}$$

So the hat matrix is

$$\begin{aligned}H_{full} &= Z(Z^\top Z)^{-1}Z^\top \\ &= \begin{bmatrix}\hat{Z}_1 & \hat{Z}_2\end{bmatrix}\left(\begin{bmatrix}\hat{Z}_1^\top \\ \hat{Z}_2^\top\end{bmatrix}\begin{bmatrix}\hat{Z}_1 & \hat{Z}_2\end{bmatrix}\right)^{-1}\begin{bmatrix}\hat{Z}_1^\top \\ \hat{Z}_2^\top\end{bmatrix} \\ &= \begin{bmatrix}\hat{Z}_1 & \hat{Z}_2\end{bmatrix}\begin{bmatrix}\hat{Z}_1^\top\hat{Z}_1 & \hat{Z}_1^\top\hat{Z}_2 \\ \hat{Z}_2^\top\hat{Z}_1 & \hat{Z}_2^\top\hat{Z}_2\end{bmatrix}^{-1}\begin{bmatrix}\hat{Z}_1^\top \\ \hat{Z}_2^\top\end{bmatrix} \\ &= \begin{bmatrix}\hat{Z}_1 & \hat{Z}_2\end{bmatrix}\begin{bmatrix}\hat{Z}_1^\top\hat{Z}_1 & 0 \\ 0 & \hat{Z}_2^\top\hat{Z}_2\end{bmatrix}^{-1}\begin{bmatrix}\hat{Z}_1^\top \\ \hat{Z}_2^\top\end{bmatrix} \\ &= \begin{bmatrix}\hat{Z}_1 & \hat{Z}_2\end{bmatrix}\begin{bmatrix}(\hat{Z}_1^\top\hat{Z}_1)^{-1} & 0 \\ 0 & (\hat{Z}_2^\top\hat{Z}_2)^{-1}\end{bmatrix}\begin{bmatrix}\hat{Z}_1^\top \\ \hat{Z}_2^\top\end{bmatrix} \\ &= \hat{Z}_1(\hat{Z}_1^\top\hat{Z}_1)^{-1}\hat{Z}_1^\top + \hat{Z}_2(\hat{Z}_2^\top\hat{Z}_2)^{-1}\hat{Z}_2^\top \\ &= H_1 + H_2\end{aligned}$$

Therefore, for R^2's, we have

$$\begin{aligned} R^2_{full} &= \frac{Y^\top H_{full} Y}{Y^\top Y} \\ &= \frac{Y^\top (H_1 + H_2) Y}{Y^\top Y} \\ &= \frac{Y^\top H_1 Y}{Y^\top Y} + \frac{Y^\top H_2 Y}{Y^\top Y} \\ &= R^2_1 + R^2_2 \end{aligned}$$