MATH 4630 / 6632 3.0 - Fall 2022
Solution for Assignment 3

Question 1:

a. Model $Y = X\beta + \epsilon$ where

$$Y = \begin{pmatrix} Y_{11} & Y_{12} \\ \vdots & \vdots \\ Y_{n1} & Y_{n2} \end{pmatrix},\ X = \begin{pmatrix} 1 & X_{11} & X_{12} \\ \vdots & \vdots & \vdots \\ 1 & X_{n1} & X_{n2} \end{pmatrix},\ \beta = \begin{pmatrix} \beta_{01} & \beta_{02} \\ \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix},\ \epsilon = \begin{pmatrix} \epsilon_{11} & \epsilon_{12} \\ \vdots & \vdots \\ \epsilon_{n1} & \epsilon_{n2} \end{pmatrix},\ n = 25$$

Assumptions:

– $\underset{\sim}{\epsilon}_i$ are iid with mean $\underset{\sim}{0}$ and variance $\sigma_{ii}I$ for $i = 1, 2$

– $cov(\underset{\sim}{\epsilon}_i, \underset{\sim}{\epsilon}_j) = \sigma_{ij}I$

b. Using SAS or R, we have

$$\hat{Y} = (\hat{Y}_1, \hat{Y}_2) = (1, X_1, X_2)\begin{pmatrix} 28.0021 & 25.8024 \\ 0.3876 & 0.2447 \\ 0.5766 & 0.4713 \end{pmatrix}$$

or we can write it as

$$\begin{aligned} \hat{Y}_1 &= 28.0021 + 0.3876X_1 + 0.5766X_2 \\ \hat{Y}_2 &= 25.8024 + 0.2447X_1 + 0.4713X_2 \end{aligned}$$

c. $H_0 : \beta_{ij} = 0$ (for all $i, j = 1, 2$ vs H_a : not all zero
Using SAS, we have

$$SSE = \begin{pmatrix} 948.0150 & 267.5750 \\ 267.5750 & 606.8363 \end{pmatrix}$$

and using matrix algebra with R, we have

$$SSTotal = (Y - \bar{Y})(Y - \bar{Y})' = \begin{pmatrix} 2277.44 & 1230.04 \\ 1230.04 & 1304.64 \end{pmatrix}$$

where each row of $\bar{Y}$ is $(\bar{Y}_1, \bar{Y}_2)$.
The likelihood ratio statistic is

$$\Lambda = \frac{|SSE|}{|SSTotal|} = 0.3454$$

The observed Bartlett corrected test statistic is

$$-(25 - 2 - 1 - (2 - 2 + 1 + 1)/2)\log 0.3454 = 22.3249$$

and the corresponding p-value is $= P(\chi^2_2 > 13.3249) \approx 0$
Extremely small p-value $\Rightarrow$ strong evidence that the model is significant

d. H_0 : no X_1 effect vs $H_0 : X_1$ effect exists
From SAS, the observed Wilks' Lambda statistic is $\Lambda^*_{obs} = 0.8787$ and the corresponding p-value is 0.2573
Large p-value $\Rightarrow$ no evidence that X_1 effect exists
H_0 : no X_2 effect vs $H_0 : X_2$ effect exists
From SAS, the observed Wilks' Lambda statistic is $\Lambda^*_{obs} = 0.8552$ and the corresponding p-value is 0.1935
Large p-value $\Rightarrow$ no evidence that X_2 is significant

e. Let $\underset{\sim}{x}_0 = (1, 192, 152)$, then $\hat{\underset{\sim}{y}} = \underset{\sim}{x}_0\hat{\beta} = (190.0842, 154.2086)$ and $m = 2, r = 2$ and $n = 25$ Using matrix algebra from R, we have $\hat{Y} = (190.0842, 154.2086)'$,

$$\widehat{var}(\hat{Y}) = \underset{\sim}{x}_0(X'X)^{-1}\underset{\sim}{x}_0'\left(\frac{E}{n-r-1}\right) = \begin{pmatrix} 8.0212 & 2.2640 \\ 2.2640 & 5.1345 \end{pmatrix} = \begin{pmatrix} \widehat{var}(\hat{Y}_1) & \widehat{cov}(\hat{Y}_1, \hat{Y}_2) \\ \widehat{cov}(\hat{Y}_1, \hat{Y}_2) & \widehat{var}(\hat{Y}_2) \end{pmatrix}$$

For working Hotelling, the table value is

$$\sqrt{\frac{m(n-r-1)}{n-r-m}F_{m,n-r-m,\alpha}} = 2.6951$$

Therefore, the required confidence interval is

$$\begin{array}{ll} \text{mean of } Y_1: & 190.0842 \pm 2.6951\sqrt{8.0212} \Leftrightarrow (185.1381, 195.0111) \\ \text{mean of } Y_2: & 154.2086 \pm 2.6951\sqrt{5.1345} \Leftrightarrow (150.4678, 158.3668) \end{array}$$

f. Let $\underset{\sim}{Y}_0$ be the new response at $\underset{\sim}{x}_0$. We have $\hat{\underset{\sim}{Y}}_0 = (190.0842, 154.2086)$ and

$$\widehat{var}(\hat{Y}_0) = \underset{\sim}{x}_0(X'X)^{-1}\underset{\sim}{x}_0'\left(\frac{E}{n-r-1}\right) + \left(\frac{E}{n-r-1}\right) = \begin{pmatrix} 46.4464 & 13.1094 \\ 13.1094 & 29.7309 \end{pmatrix}$$

Therefore, the required prediction interval is

$$\begin{array}{ll} \text{new response of } Y_1: & 190.0842 \pm 2.6951\sqrt{46.4464} \Leftrightarrow (171.7068, 208.4424) \\ \text{new response of } Y_2: & 154.2086 \pm 2.6951\sqrt{29.7309} \Leftrightarrow (139.7217, 169.1128) \end{array}$$

Question 2:

a. From SAS, we can obtain the sample covariance matrix and also its eigenvalues and eigenvectors. The five principle components are:

$$
\begin{aligned}
Z_1 &= 0.4728X_1 + 0.3919X_2 + 0.4875X_3 + 0.4677X_4 + 0.4080X_5 \\
Z_2 &= -0.5763X_1 - 0.1083X_2 + 0.0960X_3 - 0.1206X_4 + 0.7952X_5 \\
Z_3 &= -0.44167X_1 + 0.4524X_2 - 0.4795X_3 + 0.6195X_4 - 0.0887X_5 \\
Z_4 &= 0.2286X_1 + 0.6558X_2 - 0.3690X_3 - 0.5803X_4 + 0.2115X_5 \\
Z_5 &= -0.4672X_1 + 0.4472X_2 + 0.6222X_3 - 0.2146X_4 - 0.3854X_5
\end{aligned}
$$

b. Using 3 principle components will cover 92.37% of total variance

c. From R, we can obtain the sample correlation matrix and also its eigenvalues and eigenvectors. The five principle components are:

$$
\begin{aligned}
Z_1 &= 0.4418X_1 + 0.4536X_2 + 0.4728X_3 + 0.4536X_4 + 0.4120X_5 \\
Z_2 &= -0.2006X_1 - 0.4281X_2 + 0.3679X_3 - 0.3935X_4 + 0.6974X_5 \\
Z_3 &= -0.6786X_1 + 0.3491X_2 - 0.3754X_3 + 0.3345X_4 + 0.4059X_5 \\
Z_4 &= 0.2125X_1 + 0.6055X_2 - 0.2581X_3 - 0.7010X_4 + 0.1735X_5 \\
Z_5 &= 0.5088X_1 - 0.3500X_2 - 0.6584X_3 + 0.1900X_4 + 0.3860X_5
\end{aligned}
$$

Using 3 principle components will cover 92.06% of total correlation

Question 3:

a. Using SAS, the 2 canonical correlations are$\hat{\rho}_1^* = 0.8587$ and $\hat{\rho}_2^* = 0.4126$.

b. $H_0 : \rho_1^* = \rho_2^* = 0$ vs H_a : not all zero.
From SAS, the p-value from the Wilks Lambda test is 0.1052.
Large p-value implies no evidence against H_0.

c. $H_0 : \rho_1^* = 0$ vs $H_a : \rho_1^* \neq 0$.
From SAS, the p-value from the Wilks Lambda test is 0.1052.
Large p-value implies no evidence against $\rho_1^* = 0$.
$H_0 : \rho_2^* = 0$ vs $H_a : \rho_2^* \neq 0$.
From SAS, the p-value from the Wilks Lambda test is 0.5204.
Large p-value implies no evidence against $\rho_2^* = 0$.

Question 4:

a. Since

$$\begin{aligned}
-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{\mu}_1)'\Sigma^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu}_1) &= -\frac{1}{2}\underset{\sim}{x}'\Sigma^{-1}\underset{\sim}{x}+\underset{\sim}{\mu}_1'\Sigma^{-1}\underset{\sim}{x}--\frac{1}{2}\underset{\sim}{\mu}_1'\Sigma^{-1}\underset{\sim}{\mu}_1 \\
-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{\mu}_2)'\Sigma^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu}_2) &= -\frac{1}{2}\underset{\sim}{x}'\Sigma^{-1}\underset{\sim}{x}+\underset{\sim}{\mu}_2'\Sigma^{-1}\underset{\sim}{x}--\frac{1}{2}\underset{\sim}{\mu}_2'\Sigma^{-1}\underset{\sim}{\mu}_2 \\
(\underset{\sim}{\mu}_1-\underset{\sim}{\mu}_2)'\Sigma^{-1}(\underset{\sim}{\mu}_1+\underset{\sim}{\mu}_2) &= \underset{\sim}{\mu}_1\Sigma^{-1}\underset{\sim}{\mu}_1-\underset{\sim}{\mu}_2\Sigma^{-1}\underset{\sim}{\mu}_2
\end{aligned}$$

Left side of the equation can be simplified to

$$\begin{aligned}
LS &= (\underset{\sim}{\mu}_1'-\underset{\sim}{\mu}_2')\Sigma^{-1}\underset{\sim}{x}-\frac{1}{2}(\underset{\sim}{\mu}_1\Sigma^{-1}\underset{\sim}{\mu}_1-\underset{\sim}{\mu}_2\Sigma^{-1}\underset{\sim}{\mu}_2) \\
&= (\underset{\sim}{\mu}_1-\underset{\sim}{\mu}_2)'\Sigma^{-1}\underset{\sim}{x}-\frac{1}{2}(\underset{\sim}{\mu}_1-\underset{\sim}{\mu}_2)'\Sigma^{-1}(\underset{\sim}{\mu}_1+\underset{\sim}{\mu}_2) = RS
\end{aligned}$$

b1. You can do the plot and it is 2 identical isoscles triangles (shifted) intersecting at $x = 0.025$.

b2. For $p_1 = p_2$ and $c(1|2) = c(2|1)$, a new data point x_0 will be classified into π_1 if

$$\frac{f_1(x_0)}{f_2(x_0)} \geq 1 \quad \Leftrightarrow \quad f_1(x_0) = f_2(x_0)$$

In this case, we can solve the equation and we have

$$\begin{cases} x_0 \in \pi_1 & \text{if} \quad x_0 < 0.25 \\ x_0 \in \pi_2 & \text{if} \quad x_0 > 0.25 \end{cases}$$

And if $x = 0$, then it can be in both populations.

b3. For $p_1 = 0.2$ and $c(1|2) = c(2|1)$, a new data point x_0 will be classified into π_1 if

$$\frac{f_1(x_0)}{f_2(x_0)} \geq \frac{1-0.2}{0.2} \quad \Leftrightarrow \quad f_1(x_0) = 4f_2(x_0)$$

In this case, we can solve the equation and we have $x_0 \in \pi_1$ if $x_0 < -0.3333$ or $x_0 > 1.6667$. Since x_0 cannot be greater than 1, thus we have

$$\begin{cases} x_0 \in \pi_1 & \text{if} \quad x_0 < -0.3333 \\ x_0 \in \pi_2 & \text{if} \quad x_0 > -0.3333 \end{cases}$$

And if $x = -0.3333$, then it can be in both populations.

Question 5:

a. From the data set, we can obtain

$$\underset{\sim}{\bar{x}}_1 = \begin{pmatrix} 48.2167 \\ 49.2750 \\ 50.1333 \\ 51.1750 \end{pmatrix}, \quad \underset{\sim}{\bar{x}}_2 = \begin{pmatrix} 49.2000 \\ 50.1500 \\ 51.2250 \\ 51.8625 \end{pmatrix}, \quad S_{pooled} = \begin{pmatrix} 6.3520 & 6.2847 & 5.8252 & 5.6136 \\ 6.2847 & 6.6035 & 6.2406 & 6.0921 \\ 5.8252 & 6.2406 & 6.9845 & 7.1321 \\ 5.6136 & 6.0921 & 7.1321 & 7.7534 \end{pmatrix}$$

With $\underset{\sim}{\hat{a}}' = (\underset{\sim}{\bar{x}}_1 - \underset{\sim}{\bar{x}}_2)' S_{pooled}^{-1}$, we have

$$\bar{y}_1 = \underset{\sim}{\hat{a}}' \underset{\sim}{\bar{x}}_1 = -6.6855, \quad \bar{y}_2 = \underset{\sim}{\hat{a}}' \underset{\sim}{\bar{x}}_2 = -7.4395.$$

Hence, for the case $p_1 = p_2$ and $c(1|2) = c(2|1)$, a new observation $\underset{\sim}{x}$ is classified into A if $\hat{y} = \underset{\sim}{\hat{a}}' \underset{\sim}{x} \geq \frac{1}{2}(\bar{y}_1 + \bar{y}_2) = -7.0625$.

b. for the case $p_1 = 0.25$ and $c(1|2) = \frac{1}{2} c(2|1)$, a new observation $\underset{\sim}{x}$ is classified into A if $\hat{y} = \underset{\sim}{\hat{a}}' \underset{\sim}{x} \geq \frac{1}{2}(\bar{y}_1 + \bar{y}_2) + \log((p_2/p_1)(c(1|2)/c(2|1)) = -6.6571$.

c. For a given population π_i is the random variable $\underset{\sim}{X}$ is distributed as $N_4(\underset{\sim}{\mu}_i, \Sigma)$ for $i = 1, 2$.
Since $\underset{\sim}{\mu}_i$ and Σ are unknown, we use $\underset{\sim}{\bar{x}}_i$ and S_{pooled} to estimate them.
With $p_1 = 0.6$, the posterior density of π_1 given an observation $\underset{\sim}{x}$ is

$$p(\pi_1 | \underset{\sim}{x}) = \frac{0.6 f(\underset{\sim}{x} | \pi_1)}{0.6 f(\underset{\sim}{x} | \pi_1) + 0.4 f(\underset{\sim}{x} | \pi_2)}$$

and similarly, the posterior density of π_2 given an observation $\underset{\sim}{x}$ is

$$p(\pi_2 | \underset{\sim}{x}) = \frac{0.4 f(\underset{\sim}{x} | \pi_2)}{0.6 f(\underset{\sim}{x} | \pi_1) + 0.4 f(\underset{\sim}{x} | \pi_2)} = 1 - p(\pi_1 | \underset{\sim}{x})$$

Hence, a new observation $\underset{\sim}{x}$ is classified into A if $p(\pi_1 | \underset{\sim}{x}) \geq p(\pi_2 | \underset{\sim}{x}))$

d. With the new data $\underset{\sim}{x} = (50, 48, 47, 49)'$, $\hat{y} = \underset{\sim}{\hat{a}}' \underset{\sim}{x} = 2.4549$
Using the classification rule in part (a), we have $\hat{y} \geq -7.0625$, this implies $\underset{\sim}{x}$ is classified into population A.
Using the classification rule in part (b), we have $\hat{y} \geq -6.6571$, this implies $\underset{\sim}{x}$ is classified into population A.
Using the classification rule in part (c), we have $p(\pi_1 | \underset{\sim}{x}) \approx 1 \geq p(\pi_2 | \underset{\sim}{x}) \approx 2$, and hence, $\underset{\sim}{x}$ is classified into population A.