SOLUTION-HW2

TA:Ao Sun

1. Consider the two covariance matrix...

$$|\Sigma_1| = 1, |\Sigma_2| = 4, tr(\Sigma_1) = 20, tr(\Sigma_2) = 15$$

2. Suppose that...

define matrix A

$$A = \begin{bmatrix} 1, & 1, & 1 \\ 1, & -1, & -1 \end{bmatrix}$$

then $A\mathbf{y}$ is multivariate normal with variance

$$A\Sigma A' = \begin{bmatrix} 4\rho + 3, & -2\rho - 1 \\ -2\rho - 1, & 3 \end{bmatrix}$$

if $-2\rho - 1 = 0$, two random variables are independent, therefore, $\rho = -1/2$

3.Suppose ...

(a)

define $a = [2, -1, 3]'$,then

$$Z = a'\mathbf{y} \sim \mathcal{N}(16, 21)$$

(b)

define matrix $A = \begin{bmatrix} 1, & 1, & 1 \\ 1, & -1, & 2 \end{bmatrix}$, then

$$Z = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = A\mathbf{y} = \mathcal{N}(\begin{bmatrix} 9 \\ 9 \end{bmatrix}, \begin{bmatrix} 29, & -1 \\ -1, & 9 \end{bmatrix})$$

(c)

the distribution of $Y_2 \sim \mathcal{N}(2, 13)$

(d)

the joint distribution (Y_1, Y_3)

$$\begin{bmatrix} Y_1 \\ Y_3 \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 6, & -2 \\ -2, & 4 \end{bmatrix})$$

(e)

define matrix $A = \begin{bmatrix} 1, & 0, & 0 \\ 0, & 0, & 1 \\ 1/2, & 0, & 1/2 \end{bmatrix}$, then

$$\begin{bmatrix} Y_1 \\ Y_3 \\ \frac{1}{2}(Y_1 + Y_3) \end{bmatrix} = A\mathbf{y} \sim \mathcal{N}(\begin{bmatrix} 3 \\ 4 \\ 3.5 \end{bmatrix}, \begin{bmatrix} 6, & -2, & 2 \\ -2, & 4. & 1 \\ 2, & 1, & 1.5 \end{bmatrix})$$

(f)

first find the joint distribution (Y_1, Z_1, Y_2, Z_2), let A,

$$A = \begin{bmatrix} 1, & 0, & 0 \\ 1, & 1, & 1 \\ 0, & 1, & 0 \\ 1, & -1, & 2 \end{bmatrix}$$

then,

$$\begin{bmatrix} Y_1 \\ Z_1 \\ Y_2 \\ Z_2 \end{bmatrix} = A\mathbf{y}\mathcal{N} \sim (\begin{bmatrix} 3 \\ 9 \\ 2 \\ 9 \end{bmatrix}, \begin{bmatrix} 6, & 5, & 1, & 1 \\ 5, & 29, & 18, & -1 \\ 1, & 18, & 13, & -4 \\ 1, & -1, & -4, & 9 \end{bmatrix})$$

then mean of condition distribution is

$$\begin{bmatrix} 3 \\ 9 \end{bmatrix} + \frac{1}{101} \begin{bmatrix} 1 & 1 \\ 18 & -1 \end{bmatrix} \begin{bmatrix} 9 & 4 \\ 4 & -13 \end{bmatrix} \begin{bmatrix} Y_2 - 2 \\ Z_2 - 9 \end{bmatrix} = \begin{bmatrix} \frac{13}{101} Y_2 + \frac{17}{101} Z_2 + \frac{124}{101} \\ \frac{158}{101} Y_2 + \frac{59}{101} Z_2 + \frac{62}{101} \end{bmatrix}$$

$$(Y_1, Z_1)|(Y_2, Z_2) \sim (\begin{bmatrix} \frac{13}{101} Y_2 + \frac{17}{101} Z_2 + \frac{124}{101} \\ \frac{158}{101} Y_2 + \frac{59}{101} Z_2 + \frac{62}{101} \end{bmatrix}, \begin{bmatrix} \frac{576}{101}, & \frac{288}{101} \\ \frac{288}{101}, & \frac{144}{101} \end{bmatrix})$$

4. Let...

(a)

$$P(Y_2 \le a) = \begin{cases} P(Y_1 \le a), a \le -1, \\ P(Y_2 \le -1) + P(-1 < Y_2 \le a) = P(Y_1 \le -1) + P(-a \le Y_1 < 1) = P(Y_1 \le a), -1 < a < 1, \\ P(Y_2 \le -1) + P(-1 < Y_2 \le 1) + P(1 < Y_2 \le a) = P(Y_1 \le a), a > 1. \end{cases}$$

(b)

If (Y_1, Y_2) were bivariate normal, then it must be the case that $Y_1 + Y_2$ is normal. Note that

$$P(Y_1 + Y_2 = 0) = P(-1 \le Y_1 \le 1) \ne 0,$$

which leads to a contradiction.

5. Suppose...

(a) Check that...

Check:

$$\begin{aligned} (\mathbf{y} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu}) = & \left\{\mathbf{y}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{y}_2 - \boldsymbol{\mu}_2)\right\}' \\ & \times \left(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1} \left\{\mathbf{y}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{y}_2 - \boldsymbol{\mu}_2)\right\} \\ & + (\mathbf{y}_2 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}_{22}^{-1} (\mathbf{y}_2 - \boldsymbol{\mu}_2) \end{aligned}$$

partition

$$\mathbf{y} - \boldsymbol{\mu} = (\mathbf{y}_1' - \boldsymbol{\mu}_1', \mathbf{y}_2' - \boldsymbol{\mu}_2')'$$
$$\begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$$

With computation, we can get

$$\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \left(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1} & -\left(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1} \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\left(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1} & \boldsymbol{\Sigma}_{22}^{-1} + \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\left(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \end{pmatrix}$$

For simplicity, use $\mathbf{a}$ to represent $\left(\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1}$, so we ger

$$\boldsymbol{\Sigma}^{-1}=\begin{pmatrix} \mathbf{a} & -\mathbf{a}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{a} & \boldsymbol{\Sigma}_{22}^{-1}+\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{a}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \end{pmatrix}$$

Therefore, the original formula turns to

$$(\mathbf{y}_1'-\boldsymbol{\mu}_1',\mathbf{y}_2'-\boldsymbol{\mu}_2')\begin{pmatrix} \mathbf{a} & -\mathbf{a}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{a} & \boldsymbol{\Sigma}_{22}^{-1}+\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{a}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \end{pmatrix}(\mathbf{y}_1-\boldsymbol{\mu}_1,\mathbf{y}_2-\boldsymbol{\mu}_2)'$$

We can decompose $\boldsymbol{\Sigma}^{-1}$ as

$$\begin{pmatrix} \mathbf{a} & -\mathbf{a}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{a} & \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{a}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \end{pmatrix}+\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22}^{-1} \end{pmatrix}$$

It is clear to check that

$$(\mathbf{y}_1'-\boldsymbol{\mu}_1',\mathbf{y}_2'-\boldsymbol{\mu}_2')\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22}^{-1} \end{pmatrix}(\mathbf{y}_1-\boldsymbol{\mu}_1,\mathbf{y}_2-\boldsymbol{\mu}_2)=(\mathbf{y}_2'-\boldsymbol{\mu}_2')\,\boldsymbol{\Sigma}_{22}^{-1}\,(\mathbf{y}_2-\boldsymbol{\mu}_2)$$

Therefore, we can get the following conclusion,

$$\begin{aligned}(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu}) &= (\mathbf{y}_1'-\boldsymbol{\mu}_1',\mathbf{y}_2'-\boldsymbol{\mu}_2')\begin{pmatrix} \mathbf{a} & -\mathbf{a}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{a} & \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{a}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \end{pmatrix}(\mathbf{y}_1-\boldsymbol{\mu}_1,\mathbf{y}_2-\boldsymbol{\mu}_2) \\
&= (\mathbf{y}_1'-\boldsymbol{\mu}_1')\,\mathbf{a}\,(\mathbf{y}_1-\boldsymbol{\mu}_1)-(\mathbf{y}_2'-\boldsymbol{\mu}_2')\left(\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)\mathbf{a}\,(\mathbf{y}_1-\boldsymbol{\mu}_1) \\
&- (\mathbf{y}_1'-\boldsymbol{\mu}_1')\,\mathbf{a}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\,(\mathbf{y}_2-\boldsymbol{\mu}_2)+(\mathbf{y}_2'-\boldsymbol{\mu}_2')\,\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{a}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\,(\mathbf{y}_2-\boldsymbol{\mu}_2) \\
&= \left(\mathbf{y}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\,(\mathbf{y}_2-\boldsymbol{\mu}_2)\right)' \\
&\times \mathbf{a}\left(\mathbf{y}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\,(\mathbf{y}_2-\boldsymbol{\mu}_2)\right) \\
&+ (\mathbf{y}_2-\boldsymbol{\mu}_2)'\,\boldsymbol{\Sigma}_{22}^{-1}\,(\mathbf{y}_2-\boldsymbol{\mu}_2)\end{aligned}$$

(b) Derive the conditional density...

$$f\left(\mathbf{y}_1|\mathbf{y}_2\right)=\tfrac{f(\mathbf{y}_1,\mathbf{y}_2)}{f(\mathbf{y}_2)}$$
$$f(\mathbf{y})=\tfrac{1}{(2\pi)^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left\{\tfrac{1}{2}(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right\}$$

According to above conclusion,we can decomposition $f(\mathbf{y})=f\left(\mathbf{y}_1|\mathbf{y}_2\right)f(\mathbf{y_2})$

where

$$f\left(\mathbf{y}_2\right)=\frac{1}{(2\pi)^{\frac{p_2}{2}}\left|\Sigma_{22}\right|^{\frac{1}{2}}}\exp\left\{\frac{1}{2}\left(\mathbf{y}_2-\boldsymbol{\mu}_2\right)'\boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{y}_2-\boldsymbol{\mu}_2\right)\right\}$$

and

$$\begin{aligned}f\left(\mathbf{y}_1|\mathbf{y}_2\right)=&\frac{1}{(2\pi)^{\frac{p_1}{2}}\left|\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right|^{\frac{1}{2}}} \\
&\times\exp\left\{\frac{1}{2}\left(\mathbf{y}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{y}_2-\boldsymbol{\mu}_2\right)\right)'\mathbf{a}\left(\mathbf{y}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{y}_2-\boldsymbol{\mu}_2\right)\right)\right\}\end{aligned}$$

As a result

$$\mathbf{y}_1|\mathbf{y}_2\sim N\left(\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{y}_2-\boldsymbol{\mu}_2\right),\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)$$

6. For one sample case, prove that the likelihood ratio test leads to...

Suppose $y = (y_1, \cdots, y_n)$ is a $N_p(\mu, \Sigma)$ sample. The testing of interest is

$$H_0 : \mu = \mu_0 \leftrightarrow H_1 : \mu \neq \mu_0.$$

Note that the unrestricted maximum likelihood is

$$\begin{aligned}
\mathcal{L}(\Theta) &= \sup_{\mu,\Sigma} \mathcal{L}(y; \mu, \Sigma) \\
&= \mathcal{L}\left(y; \bar{\mu}, \hat{\Sigma}\right) \\
&= \frac{1}{(2\pi)^{np}|\hat{\Sigma}|^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_i-\bar{\mu})'\hat{\Sigma}^{-1}(\mathbf{y}_i-\bar{\mu})} \\
&= \frac{1}{(2\pi)^{np}|\hat{\Sigma}|^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}\operatorname{tr}\left[\hat{\Sigma}^{-1}(\mathbf{y}_i-\bar{\mu})(\mathbf{y}_i-\bar{\mu})'\right]} \\
&= \frac{1}{(2\pi)^{np}|\hat{\Sigma}|^{\frac{n}{2}}} e^{-\frac{1}{2}\operatorname{tr}\left[\hat{\Sigma}^{-1}\sum_{i=1}^{n}(\mathbf{y}_i-\bar{\mu})(\mathbf{y}_i-\bar{\mu})'\right]} \\
&= \frac{1}{(2\pi)^{np}|\hat{\Sigma}|^{\frac{n}{2}}} e^{-\frac{np}{2}}
\end{aligned}$$

Similarly, the restricted maximum likelihood is

$$\mathcal{L}(\Theta_0) = \sup_{\Sigma} \mathcal{L}(y; \mu_0, \Sigma) = \mathcal{L}\left(y; \mu_0, \hat{\Sigma}_0\right) = \frac{1}{(\sqrt{2\pi})^{np}\det(\hat{\Sigma}_0)^{n/2}} \exp\left(-\frac{np}{2}\right), \quad \hat{\Sigma}_0 = \frac{1}{n}\sum_{j=1}^{n}(y_j-\mu_0)(y_j-\mu_0)'.$$

Therefore, the generalized likelihood ratio is

$$\lambda = \frac{\mathcal{L}(\Theta)}{\mathcal{L}(\Theta_0)} = \left\{\frac{\det\left(\hat{\Sigma}_0\right)}{\det\left(\hat{\Sigma}\right)}\right\}^{n/2} = \left[\frac{\det\left\{\hat{\Sigma} + (\bar{y}-\mu_0)(\bar{y}-\mu_0)'\right\}}{\det\left(\hat{\Sigma}\right)}\right]^{n/2}$$

$$= \left\{\det\left(I_p + (\bar{y}-\mu_0)(\bar{y}-\mu_0)'\hat{\Sigma}^{-1}\right)\right\}^{n/2} = \left\{1 + (\bar{y}-\mu_0)'\hat{\Sigma}^{-1}(\bar{y}-\mu_0)\right\}^{n/2} = \left(1 + \frac{T^2}{n-1}\right)^{n/2},$$

$$\begin{aligned}
\lambda &= \frac{\frac{1}{(2\pi)^{np}|\hat{\Sigma}|^{\frac{n}{2}} e^{-\frac{np}{2}}}}{\frac{1}{(2\pi)^{np}|\tilde{\Sigma}|^{\frac{n}{2}} e^{-\frac{np}{2}}}} \\
&= \frac{|\hat{\Sigma}|^{-\frac{n}{2}}}{|\tilde{\Sigma}|^{-\frac{n}{2}}} \\
&= \frac{\left|\sum_{i=1}^{n}(\mathbf{y}_i - \boldsymbol{\mu}_0)(\mathbf{y}_i - \boldsymbol{\mu}_0)'\right|^{-\frac{n}{2}}}{\left|\sum_{i=1}^{n}(\mathbf{y}_i - \overline{\mathbf{y}})(\mathbf{y}_i - \overline{\mathbf{y}})'\right|^{-\frac{n}{2}}} \\
&= \frac{\left|\sum_{i=1}^{n}(\mathbf{y}_i - \overline{\mathbf{y}})(\mathbf{y}_i - \overline{\mathbf{y}})' + \sum_{i=1}^{n}(\overline{\mathbf{y}} - \boldsymbol{\mu}_0)(\overline{\mathbf{y}} - \boldsymbol{\mu}_0)'\right|^{-\frac{n}{2}}}{\left|\sum_{i=1}^{n}(\mathbf{y}_i - \overline{\mathbf{y}})(\mathbf{y}_i - \overline{\mathbf{y}})'\right|^{-\frac{n}{2}}}
\end{aligned}$$

where numerator can be written as

$$\begin{vmatrix} \sum_{i=1}^{n}(\mathbf{y}_i - \overline{\mathbf{y}})(\mathbf{y}_i - \overline{\mathbf{y}})' & -\sqrt{n}(\overline{\mathbf{y}} - \boldsymbol{\mu}_0) \\ \sqrt{n}(\overline{\mathbf{y}} - \boldsymbol{\mu}_0)' & I \end{vmatrix}^{-\frac{n}{2}}$$

then,

$$\begin{aligned}
\lambda &= \frac{\left|\sum_{i=1}^{n}\left(\mathbf{y}_i-\overline{\mathbf{y}}\right)\left(\mathbf{y}_i-\overline{\mathbf{y}}\right)'\right|^{-\frac{n}{2}}\left|I+n\left(\overline{\mathbf{y}}-\boldsymbol{\mu}_0\right)'\left(\sum_{i=1}^{n}\left(\mathbf{y}_i-\overline{\mathbf{y}}\right)\left(\mathbf{y}_i-\overline{\mathbf{y}}\right)'\right)^{-1}\left(\overline{\mathbf{y}}-\boldsymbol{\mu}_0\right)\right|^{-\frac{n}{2}}}{\left|\sum_{i=1}^{n}\left(\mathbf{y}_i-\overline{\mathbf{y}}\right)\left(\mathbf{y}_i-\overline{\mathbf{y}}\right)'\right|^{-\frac{1}{2}}} \\
&= \left|I+n\left(\overline{\mathbf{y}}-\boldsymbol{\mu}_0\right)'\left(\sum_{i=1}^{n}\left(\mathbf{y}_i-\overline{\mathbf{y}}\right)\left(\mathbf{y}_i-\overline{\mathbf{y}}\right)'\right)^{-1}\left(\overline{\mathbf{y}}-\boldsymbol{\mu}_0\right)\right|^{-\frac{n}{2}} \\
&= \left[1+\left(\overline{\mathbf{y}}-\boldsymbol{\mu}_0\right)'\left(\frac{(n-1)S}{n}\right)^{-1}\left(\overline{\mathbf{y}}-\boldsymbol{\mu}_0\right)\right]^{-\frac{n}{2}} \\
&= \left(1+\frac{1}{n-1}T^2\right)^{-\frac{n}{2}}
\end{aligned}$$

So, the corresponding rejection region is

$$\{Y:\lambda\le\lambda^*\}=\left\{Y:\left(1+\frac{1}{n-1}T^2\right)^{-\frac{n}{2}}\le\lambda^*\right\}=\{Y:T^2>c^*\}$$

as desired.

7. (R exercise.) The world's 10 largest companies (2005 database) yield...

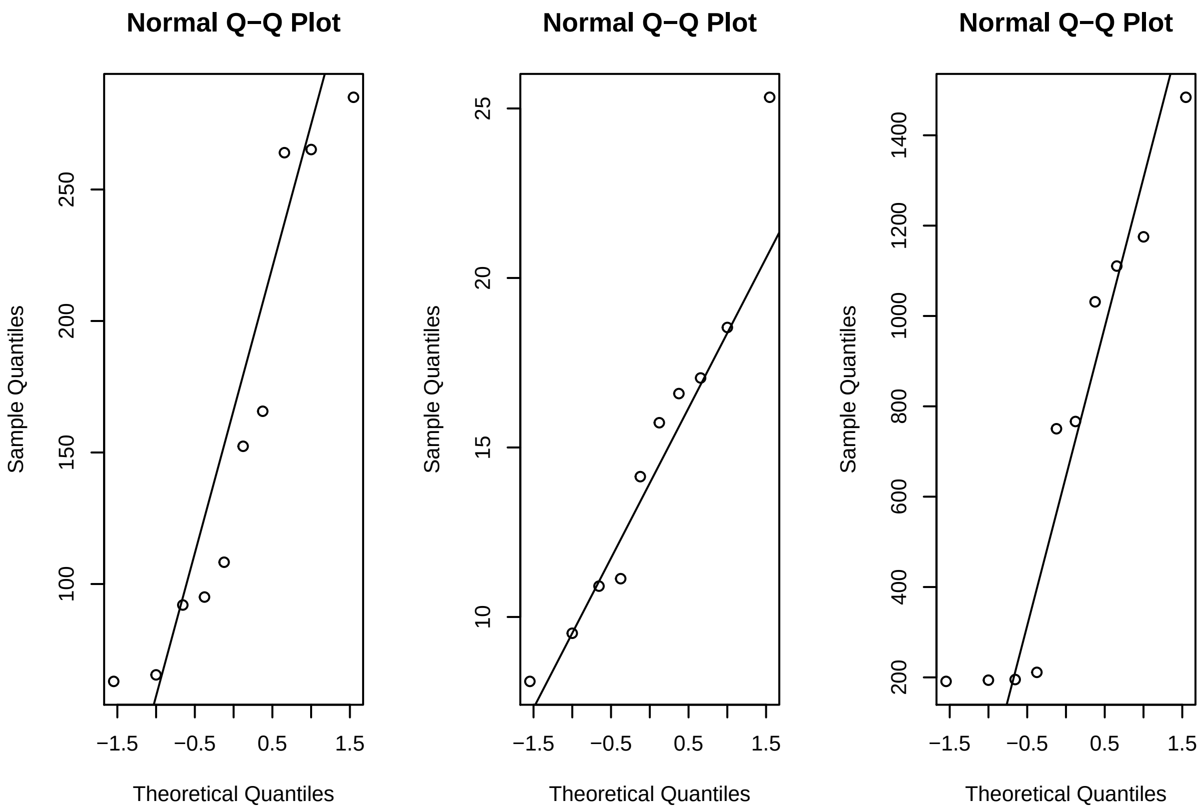
```
rm(list = ls())
y <- read.table("company.txt")
y
```

```
##        V1    V2      V3
## 1  108.28 17.05 1484.10
## 2  152.36 16.59  750.33
## 3   95.04 10.91  766.42
## 4   65.45 14.14 1110.46
## 5   62.97  9.52 1031.29
## 6  263.99 25.33  195.26
## 7  265.19 18.54  193.83
## 8  285.06 15.73  191.11
## 9   92.01  8.10 1175.16
## 10 165.68 11.13  211.15
```

For all the three variables:

(a) Construct individual QQ plots to investigate univariate normality. Interpret the output.

```
par(mfrow = c(1, 3))
qqnorm(y$V1)
qqline(y$V1)
qqnorm(y$V2)
qqline(y$V2)
qqnorm(y$V3)
qqline(y$V3)
```



As illustrated in the graphs above, most points were close to the corresponding Q-Q line, which seems to indicate a fine normality. There still exist, however, some concerns: (1) X_1 and X_3 performed slightly thinner tails than the normal; (2) one sample point of X_2 is large outlier.

(b) Conduct formal statistical tests for the individual normality. Explain the results.

```
shapiro.test(y$V1)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  y$V1
## W = 0.85909, p-value = 0.07444
```

```
shapiro.test(y$V2)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  y$V2
## W = 0.94221, p-value = 0.5778
```

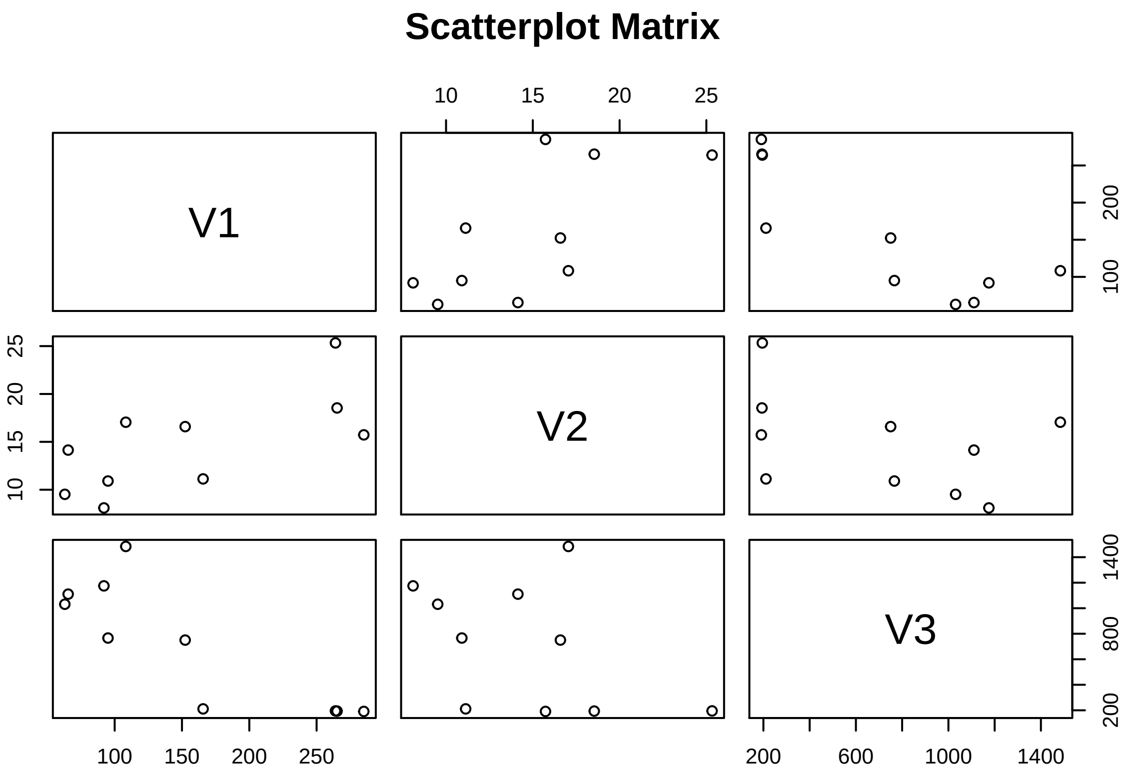
```
shapiro.test(y$V3)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  y$V3
## W = 0.86969, p-value = 0.09914
```

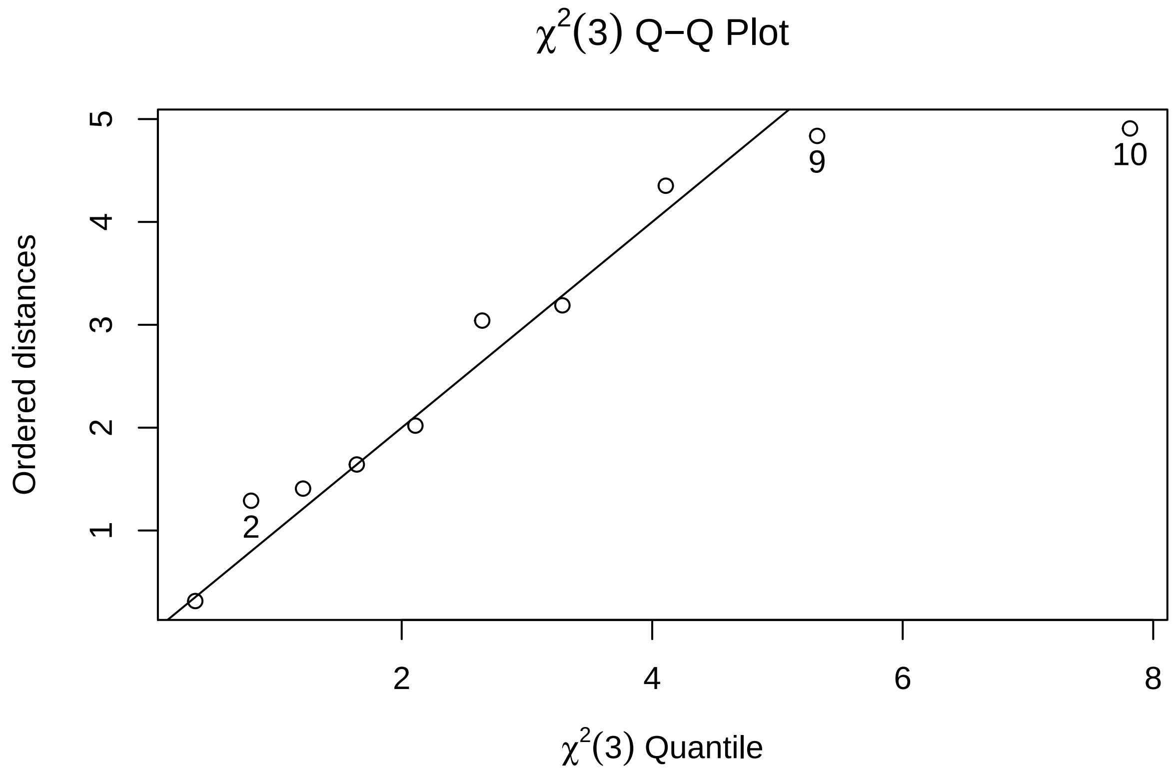
Based on the Shapiro-Wilks test, the normality of all the three variables cannot be respectively rejected with level 0.05, which coincides with the normal Q-Q plots. Moreover, the normality of X_1 and X_3 can be respectively rejected with level 0.1, which coincides with the concern (1) in (a). It seems that a single large outlier did not affect the result of the Shapiro-Wilk test since p-value of normality for X_2 was relatively high.

(c) Check the multivariate normality of...

```
pairs(y, main = "Scatterplot Matrix")
```




```
cm <- colMeans(y)
S <- cov(y)
d <- apply(y, 1, function(y) t(y - cm) %*% solve(S) %*% (y - cm))
plot(qc <- qchisq((1:nrow(y) - 1/2) / nrow(y), df = 3), sd <- sort(d),
     xlab = expression(paste(chi^2, (3), " Quantile")), ylab = "Ordered distances")
oups <- which(rank(abs(qc - sd), ties.method = "random") > nrow(y) - 3)
text(qc[oups], sd[oups] - 0.25, oups)
abline(0, 1)
title(expression(paste(chi^2, (3), " Q-Q Plot")))
```



Based on the scatterplot, the linear relationship between all pairs of the three variables were at least moderate. Based on the $\chi^2(3)$ Q-Q plot, $d_i = (x_i - \bar{x})'S^{-1}(x_i - \bar{x})$ were overall close to $\chi^2(3)$ except for d_{10}, which is corresponding to Toyota Motor. Thus, there is no powerful evidence to state that $(X_1, X_2, X_3)'$ are not jointly normal.

8. (R exercise) Recall the relationship between the hypothesis testing and the confidence interval, i.e. the conclusion of a test can be directly obtained from the related confidence interval. For the multivariate case, the confidence interval becomes the confidence region.

(a) Analogous to the definition of confidence interval, define the...

An $1-\alpha$ confidence region for μ is any random region $C(y)$ in $\mathbb{R}^p$ that safisfies $P_\mu(\mu \in C(y)) \geq 1-\alpha$. By the argument in question 6, an $1-\alpha$ confidence region when the covariance matrix is unknown can be given by

$$\left\{\mu : (\bar{y}-\mu)'S^{-1}(\bar{y}-\mu) \leq \frac{p(n-1)}{n(n-p)}F_{1-\alpha}(p, n-p)\right\}.$$

(b) For the sweat data (Page 20 in the slides, and data attached as sweat.dat), suppose we only have the information of the first two variables with mean...

```
rm(list = ls())
y <- read.table("sweat.dat")
y <- data.frame(V1 = y$V1, V2 = y$V2)
y
```

```
##     V1   V2
## 1  3.7 48.5
## 2  5.7 65.1
## 3  3.8 47.2
```

```
## 4  3.2 53.2
## 5  3.1 55.5
## 6  4.6 36.1
## 7  2.4 24.8
## 8  7.2 33.1
## 9  6.7 47.4
## 10 5.4 54.1
## 11 3.9 36.9
## 12 4.5 58.8
## 13 3.5 27.8
## 14 4.5 40.2
## 15 1.5 13.5
## 16 8.5 56.4
## 17 4.5 71.6
## 18 6.5 52.8
## 19 4.1 44.1
## 20 5.5 40.9
```

```r
n <- nrow(y)
p <- ncol(y)
cm <- colMeans(y)
cm
```

```
##    V1    V2
##  4.64 45.40
```

```r
S <- cov(y)
S.inv <- solve(S)
S.inv
```

```
##              V1           V2
## V1  0.42055021 -0.021070829
## V2 -0.02107083  0.006061007
```

```r
RHS <- p * (n - 1) / n / (n - p) * qf(.95, p, n - p)
RHS
```

```
## [1] 0.3752033
```

Using the computation results above, an 95% region estimate for μ is given by

$$\left\{\mu : (\mu_1 - 4.64, \quad \mu_2 - 45.40) \begin{pmatrix} 0.421 & -0.021 \\ -0.021 & 0.006 \end{pmatrix} \begin{pmatrix} \mu_1 - 4.64 \\ \mu_2 - 45.40 \end{pmatrix} \leq 0.375 \right\}.$$

(c) Describe the confidence region geometrically using...

```r
S.eig <- eigen(S)
S.eig
```

```
## eigen() decomposition
## $values
## [1] 200.295978   2.371812
##
## $vectors
##           [,1]       [,2]
## [1,] 0.0506399 -0.9987170
## [2,] 0.9987170  0.0506399
```

```r
acos(S.eig$vectors[1]) * 180 / pi
```

```
## [1] 87.09731
```

The computation results above decompose S as

$$S = T\Lambda T', \quad \Lambda = \begin{pmatrix} 200.30 & 0 \\ 0 & 2.37 \end{pmatrix}, \quad T = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \quad \varphi = 87.10^\circ.$$

```r
semiaxes <- sqrt(S.eig$values * RHS)
semiaxes
```

```
## [1] 8.6690082 0.9433512
```

Consequently, the confidence region can be written as

$$\{\mu : x'Dx \leq 1\}, \quad D^{-1} = (0.375)\Lambda = \begin{pmatrix} (8.67)^2 & 0 \\ 0 & (0.94)^2 \end{pmatrix}, \quad x = T'(\mu - \bar{y}) \iff \mu = \bar{y} + Tx.$$

To sum up, the confidence region is an 87.10° rotated ellipse centered at $\bar{y} = (4.64, 45.40)'$ with semi-axes 8.67 and 0.94.

(d) Consturct the 95% univariate confidence interval for each variable.

```r
cv <- qt(.975, n - 1)
L1 <- cm[1] - cv * sqrt(S[1] / n)
R1 <- cm[1] + cv * sqrt(S[1] / n)
L2 <- cm[2] - cv * sqrt(S[4] / n)
R2 <- cm[2] + cv * sqrt(S[4] / n)
L1
```

```
##      V1
## 3.84584
```

```r
R1
```

```
##      V1
## 5.43416
```

```r
L2
```

```
##       V2
## 38.78478
```

```r
R2
```

```
##       V2
## 52.01522
```

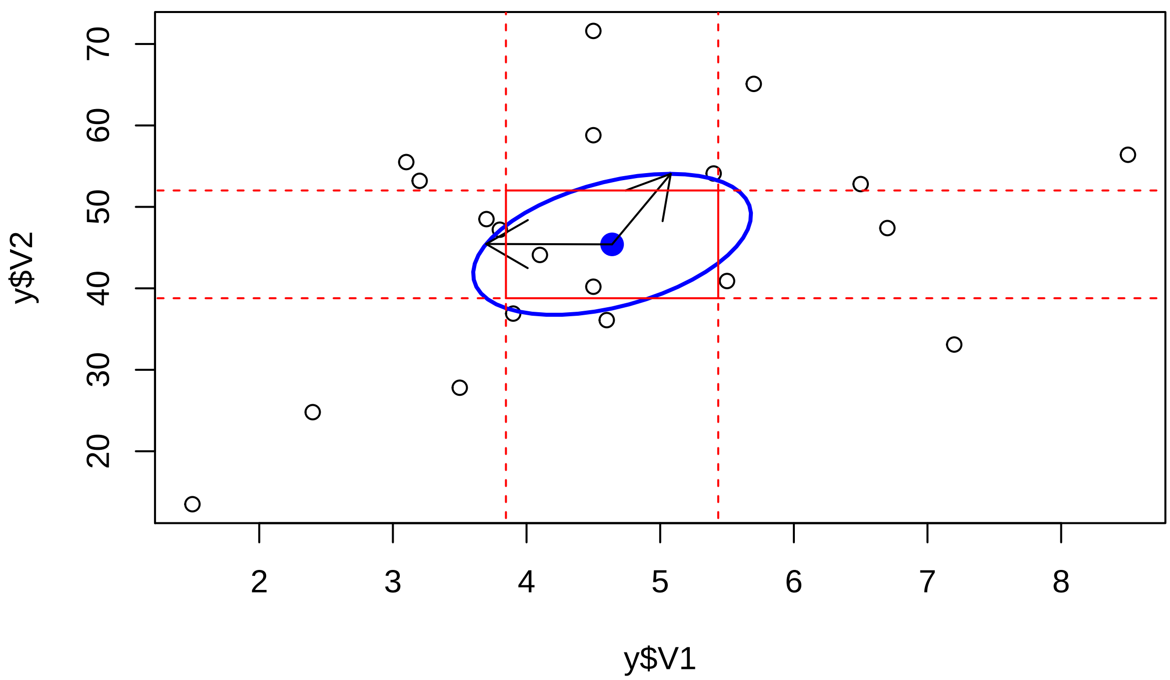
Using the computation results above, 95% confidence interval for μ_1 is $[3.85, 5.43]$ and 95% confidence interval for μ_2 is $[38.78, 52.02]$.

(e) Condsider the test...

```r
library(car)
```

```
## Loading required package: carData
```

```r
plot(x = y$V1, y = y$V2, type = "p")
ellipse(center = cm, shape = S, radius = sqrt(RHS))
arrows(cm[1], cm[2],
       cm[1] + semiaxes[1] * S.eig$vectors[,1][1],
       cm[2] + semiaxes[1] * S.eig$vectors[,1][2])
arrows(cm[1], cm[2],
       cm[1] + semiaxes[2] * S.eig$vectors[,2][1],
       cm[2] + semiaxes[2] * S.eig$vectors[,2][2])
lines(c(L1, L1), c(0, L2), col = "red", lty = 2)
lines(c(L1, L1), c(L2, R2), col = "red")
lines(c(L1, L1), c(R2, 80), col = "red", lty = 2)
lines(c(R1, R1), c(0, L2), col = "red", lty = 2)
lines(c(R1, R1), c(L2, R2), col = "red")
lines(c(R1, R1), c(R2, 80), col = "red", lty = 2)
lines(c(0, L1), c(L2, L2), col = "red", lty = 2)
lines(c(L1, R1), c(L2, L2), col = "red")
lines(c(R1, 10), c(L2, L2), col = "red", lty = 2)
lines(c(0, L1), c(R2, R2), col = "red", lty = 2)
lines(c(L1, R1), c(R2, R2), col = "red")
lines(c(R1, 10), c(R2, R2), col = "red", lty = 2)
```



As illustrated in the figure above, to find a μ_0 such that the multivariate test rejects the null hypothesis but both univariate tests accept is equivalent to find a point that is outside the ellipse but inside the rectangular. One choice is $\mu_0 = (5.2, 40)'$.

```r
mu0 = c(5.2, 40)
(mu0 - cm) %*% S.inv %*% (mu0 - cm)
```

```
##           [,1]
## [1,] 0.4360599
```

Indeed, $5.2 \in [3.85, 5.43]$, $40 \in [38.78, 52.02]$, and

$$\begin{pmatrix} 5.2 - 4.64, & 40 - 45.40 \end{pmatrix} \begin{pmatrix} 0.421 & -0.021 \\ -0.021 & 0.006 \end{pmatrix} \begin{pmatrix} 5.2 - 4.64 \\ 40 - 45.40 \end{pmatrix} = 0.436 > 0.375.$$