Homework 5

Problem 1

(a)

First, we have $\hat{\mathbf{B}} = \left(\hat{\boldsymbol{\beta}}_{(1)}, \hat{\boldsymbol{\beta}}_{(2)}, \dots, \hat{\boldsymbol{\beta}}_{(p)}\right) = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{B} + \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\Xi$, where $\Xi = (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \dots, \boldsymbol{\epsilon}_n)'$ and $\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \dots, \boldsymbol{\epsilon}_n$ are multinormally distributed with $\boldsymbol{\epsilon}_i \sim N\left(\mathbf{0}, \Sigma = (\sigma_{ij})_{n\times n}\right)$. Thus,

$$\begin{aligned} Cov\left(\hat{\boldsymbol{\beta}}_{(i)}, \hat{\boldsymbol{\beta}}_{(j)}\right) &= Cov\left(\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\hat{\boldsymbol{\epsilon}}_{(i)}, \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\hat{\boldsymbol{\epsilon}}_{(j)}\right) \\ &= \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\sigma_{ij}I_n\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1} = \sigma_{ij}\left(\mathbf{X}'\mathbf{X}\right)^{-1}, \end{aligned}$$

which leads to

$$Cov\left(\mathbf{x}_0'\hat{\boldsymbol{\beta}}_{(i)}, \mathbf{x}_0'\hat{\boldsymbol{\beta}}_{(j)}\right) = \sigma_{ij}\mathbf{x}_0'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{x}_0.$$

Now, we have

$$\hat{\mathbf{y}}_0 = \hat{\mathbf{B}}'\mathbf{x}_0 + \boldsymbol{\epsilon}_0 = \left(\mathbf{x}_0'\hat{\boldsymbol{\beta}}_{(1)}, \mathbf{x}_0'\hat{\boldsymbol{\beta}}_{(2)}, \dots, \mathbf{x}_0'\hat{\boldsymbol{\beta}}_{(p)}\right)'\boldsymbol{\epsilon}_0,$$

and

$$\begin{aligned} Var\left(\hat{\mathbf{y}}_0\right) &= Var\left(\hat{\mathbf{B}}'\mathbf{x}_0\right) + Var\left(\boldsymbol{\epsilon}_0\right) \\ &= \mathbf{x}_0'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{x}_0\Sigma + \Sigma \\ &= \left(\mathbf{x}_0'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{x}_0 + 1\right)\Sigma. \end{aligned}$$

By the fact that $E\left[\hat{\mathbf{y}}_0\right] = \mathbf{B}'\mathbf{x}_0$ and $\hat{\boldsymbol{\beta}}_{(i)}$'s are multi-normally distributed, we know that

$$\hat{\mathbf{y}}_0 \sim N\left(\mathbf{B}'\mathbf{x}_0, \left(\mathbf{x}_0'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{x}_0 + 1\right)\Sigma\right).$$

(b)

Note that by the same argument as above, we have

$$E\left[\mathbf{y}_0\right] = \hat{\mathbf{B}}'\mathbf{x}_0 \sim N\left(\mathbf{B}'\mathbf{x}_0, \mathbf{x}_0'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{x}_0\Sigma\right),$$

and thus

$$\left(\boldsymbol{x}_0'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\boldsymbol{x}_0\right)^{-1/2}\Sigma^{-1/2}\left(E\left[\mathbf{y}_0\right] - \mathbf{B}'\mathbf{x}_0\right) \sim N\left(0, I_p\right).$$

Note that

$$\Sigma^{-1/2}E\Sigma^{-1/2} = \Sigma^{-1/2}\left(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\right)'\left(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\right)\Sigma^{-1/2} = \Sigma^{-1/2}\boldsymbol{\Xi}'\left(I - H\right)\boldsymbol{\Xi}\Sigma^{-1/2},$$

where $H = \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'$ and $\text{rank}\left(I - H\right) = n - q - 1$. By spectral decomposition, we have

$$I - H = Q\begin{pmatrix} I_{n-q-1} & 0 \\ 0 & 0 \end{pmatrix}Q',$$

where $Q = (q_{ij})_{n\times n}$ is orthogonal. So,

$$\Sigma^{-1/2}E\Sigma^{-1/2} = \left(\Sigma^{-1/2}\boldsymbol{\epsilon}_1, \Sigma^{-1/2}\boldsymbol{\epsilon}_2, \cdots, \Sigma^{-1/2}\boldsymbol{\epsilon}_n\right) Q \begin{pmatrix} I_{n-q-1} & 0 \\ 0 & 0 \end{pmatrix} Q' \begin{pmatrix} \boldsymbol{\epsilon}_1'\Sigma^{-1/2} \\ \boldsymbol{\epsilon}_2'\Sigma^{-1/2} \\ \vdots \\ \boldsymbol{\epsilon}_n'\Sigma^{-1/2} \end{pmatrix},$$

where $\boldsymbol{\epsilon}_i$ is the ith row of Ξ and denoted by column vector. Now, $\Sigma^{-1/2}\boldsymbol{\epsilon}_i \sim N(0, I_p)$ and $Cov\left(\Sigma^{-1/2}\boldsymbol{\epsilon}_i, \Sigma^{-1/2}\boldsymbol{\epsilon}_j\right) = \mathbf{0}$ for $i \neq j$. Now, denote

$$(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n) = \left(\Sigma^{-1/2}\boldsymbol{\epsilon}_1, \Sigma^{-1/2}\boldsymbol{\epsilon}_2, \cdots, \Sigma^{-1/2}\boldsymbol{\epsilon}_n\right) Q,$$

we have

$$Cov(\boldsymbol{e}_i, \boldsymbol{e}_j) = Cov\left(\sum_{s=1}^n q_{si}\Sigma^{-1/2}\boldsymbol{\epsilon}_s, \sum_{t=1}^n q_{tj}\Sigma^{-1/2}\boldsymbol{\epsilon}_t\right) = \sum_{s=1}^n q_{si}q_{sj} Cov(\boldsymbol{\epsilon}_s, \boldsymbol{\epsilon}_s) = \delta_{ij} I_p,$$

where $\delta_{ij} = 1$ if $i = j$; $\delta_{ij} = 0$ otherwise. Now, we can conclude that $\boldsymbol{e}_i$'s are i.i.d. and $\boldsymbol{e}_i \sim N(\mathbf{0}, I_p)$. Thus,

$$\Sigma^{-1/2}E\Sigma^{-1/2} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n) \begin{pmatrix} I_{n-q-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_1' \\ \boldsymbol{e}_2' \\ \vdots \\ \boldsymbol{e}_n' \end{pmatrix} = \sum_{i=1}^{n-q-1} \boldsymbol{e}_i\boldsymbol{e}_i' \sim W_p(n-q-1, I_p).$$

Therefore, according to the definition of Hotelling t distribution, we have

$$\begin{aligned} &(E[\mathbf{y}_0] - \mathbf{B}'\mathbf{x}_0)' \Sigma^{-1/2} \left(\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0\right)^{-1/2} \left(\frac{\Sigma^{-1/2}E\Sigma^{-1/2}}{n-q-1}\right)^{-1} \left(\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0\right)^{-1/2} \Sigma^{-1/2} (E[\mathbf{y}_0] - \mathbf{B}'\mathbf{x}_0) \\ &\quad \sim T^2(n-q-1), \end{aligned}$$

i.e.,

$$(E[\mathbf{y}_0] - \mathbf{B}'\mathbf{x}_0)' \left(\frac{\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0 E}{n-q-1}\right)^{-1} (E[\mathbf{y}_0] - \mathbf{B}'\mathbf{x}_0) \sim T^2(n-q-1).$$

Therefore, the $100(1-\alpha)\%$ confidence ellipse is

$$(E[\mathbf{y}_0] - \mathbf{B}'\mathbf{x}_0)' \left(\frac{E}{n-q-1}\right)^{-1} (E[\mathbf{y}_0] - \mathbf{B}'\mathbf{x}_0) \leq T_\alpha^2(n-q-1)\left(\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0\right).$$

(c)

We know that, by Cauchy's inequality, for all $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^q$ and L is non-singular,

$$(\boldsymbol{a}'\boldsymbol{b})^2 \leq \left((L'\boldsymbol{a})'(L'\boldsymbol{a})\right)\left(\left(L^{-1}\boldsymbol{b}\right)'\left(L^{-1}\boldsymbol{b}\right)\right) = (\boldsymbol{a}'LL'\boldsymbol{a})\left(\boldsymbol{b}'(LL')^{-1}\boldsymbol{b}\right).$$

By Cholesky decomposition, there exists L such that $E/(n-q-1) = LL'$. Now, we let $\boldsymbol{b} = E[\mathbf{y}_0] - \mathbf{B}'\mathbf{x}_0$, so we have

$$\frac{(\boldsymbol{a}'(E[\mathbf{y}_0] - \mathbf{B}'\mathbf{x}_0))^2}{\boldsymbol{a}'E\boldsymbol{a}/(n-q-1)} \leq (E[\mathbf{y}_0] - \mathbf{B}'\mathbf{x}_0)' \left(\frac{E}{n-q-1}\right)^{-1} (E[\mathbf{y}_0] - \mathbf{B}'\mathbf{x}_0),$$

for all $\boldsymbol{a} \in \mathbb{R}^q$. Since with probability $1-\alpha$, it holds that

$$(E[\mathbf{y}_0] - \mathbf{B}'\mathbf{x}_0)' \left(\frac{E}{n-q-1}\right)^{-1} (E[\mathbf{y}_0] - \mathbf{B}'\mathbf{x}_0) \leq T_\alpha^2(n-q-1)\left(\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0\right),$$

we conclude that for all $\boldsymbol{a} \in \mathbb{R}^q$, with probability $1-\alpha$, it holds that

$$\frac{\left(\boldsymbol{a}'\left(E\left[\mathbf{y}_0\right]-\mathbf{B}'\mathbf{x}_0\right)\right)^2}{\boldsymbol{a}'E\boldsymbol{a}/\left(n-q-1\right)} \leq T_\alpha^2\left(n-q-1\right)\left(\mathbf{x}_0'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{x}_0\right).$$

Now, let $\boldsymbol{a}$ be the ith standard unit vector successively, we obtain that, for all $1 \leq i \leq q$, with probability $1-\alpha$,

$$\mathbf{B}'\mathbf{x}_{0i} - \sqrt{T_\alpha^2\left(n-q-1\right)s_{ii}\left(\mathbf{x}_0'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{x}_0\right)} \leq E\left[\mathbf{y}_{0i}\right] \leq \mathbf{B}'\mathbf{x}_{0i} + \sqrt{T_\alpha^2\left(n-q-1\right)s_{ii}\left(\mathbf{x}_0'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{x}_0\right)},$$

where s_{ii} is the ith diagonal element of $E/\left(n-q-1\right)$.

(d)

We know that

$$\left(\hat{\boldsymbol{y}}_0-\mathbf{B}'\boldsymbol{x}_0\right)'\left(\frac{\left(\boldsymbol{x}_0'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\boldsymbol{x}_0+1\right)E}{n-q-1}\right)^{-1}\left(\hat{\boldsymbol{y}}_0-\mathbf{B}'\boldsymbol{x}_0\right) \sim T^2\left(n-q-1\right),$$

so adopting the same argument, we have for all $1 \leq i \leq q$, with probability $1-\alpha$,

$$\mathbf{B}'\mathbf{x}_{0i} - \sqrt{T_\alpha^2\left(n-q-1\right)s_{ii}\left(\mathbf{x}_0'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{x}_0+1\right)} \leq \hat{\boldsymbol{y}}_{0i} \leq \mathbf{B}'\mathbf{x}_{0i} + \sqrt{T_\alpha^2\left(n-q-1\right)s_{ii}\left(\mathbf{x}_0'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{x}_0+1\right)},$$

where s_{ii} is the ith diagonal element of $E/\left(n-q-1\right)$.

2. For the covariance matrix of...

(a) Determine the population principal components...

$$|\lambda I-\Sigma| = \begin{vmatrix}\lambda-5 & -2\\ -2 & \lambda-2\end{vmatrix} = (\lambda-6)(\lambda-1) \implies \lambda_1=6, \lambda_2=1,$$

$$\begin{pmatrix}\lambda_1-5 & -2\\ -2 & \lambda_1-2\end{pmatrix}e_1=0 \quad \text{s.t.} \quad e_1'e_1=1 \implies e_1=\begin{pmatrix}2/\sqrt{5}\\ 1/\sqrt{5}\end{pmatrix} \implies Z_1=\frac{2}{\sqrt{5}}Y_1+\frac{1}{\sqrt{5}}Y_2,$$

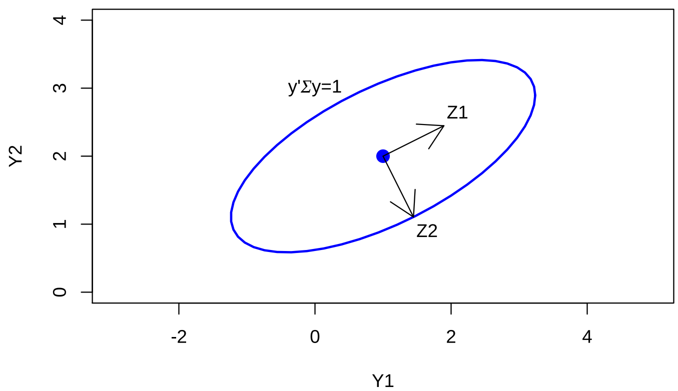
$$\begin{pmatrix}\lambda_2-5 & -2\\ -2 & \lambda_2-2\end{pmatrix}e_2=0 \quad \text{s.t.} \quad e_2'e_2=1 \implies e_2=\begin{pmatrix}1/\sqrt{5}\\ -2/\sqrt{5}\end{pmatrix} \implies Z_2=\frac{1}{\sqrt{5}}Y_1-\frac{2}{\sqrt{5}}Y_2.$$

(b) Compute the proportion of total population variance explained by the first principal component...

$$\frac{var(Z_1)}{\mathrm{Tr}(\Sigma)} = \frac{\lambda_1}{\lambda_1+\lambda_2} = \frac{6}{7}.$$

(c) Suppose the original variables follows bivariate normal distribution with mean vector...

```
rm(list = ls())
library(car)
plot(0, 0, type = "n", xlim = c(-2, 4), ylim = c(0, 4), xlab = "Y1", ylab = "Y2", asp = 1)
ellipse(center = c(1, 2), shape = matrix(c(5, 2, 2, 2), nrow = 2), radius = 1)
arrows(1, 2, 1 + 2 / sqrt(5), 2 + 1 / sqrt(5))
arrows(1, 2, 1 + 1 / sqrt(5), 2 - 2 / sqrt(5))
text(0, 3, expression(paste("y'",Sigma,"y=1")))
text(1 + 2 / sqrt(5) + 0.2, 2 + 1 / sqrt(5) + 0.2, "Z1")
text(1 + 1 / sqrt(5) + 0.2, 2 - 2 / sqrt(5) - 0.2, "Z2")
```



(d) Convert the covariance matrix to a correlation matrix...

$$P = D_s^{-1}\Sigma D_s^{-1} = \begin{pmatrix} 1 & \sqrt{2}/\sqrt{5} \\ \sqrt{2}/\sqrt{5} & 1 \end{pmatrix} \implies |\tilde{\lambda}I - \Sigma| = \tilde{\lambda}^2 - 2\tilde{\lambda} + \frac{3}{5} \implies \tilde{\lambda}_1 = 1 + \frac{\sqrt{2}}{\sqrt{5}}, \tilde{\lambda}_2 = 1 - \frac{\sqrt{2}}{\sqrt{5}},$$

$$\begin{pmatrix} \tilde{\lambda}_1 - 1 & -\sqrt{2}/\sqrt{5} \\ -\sqrt{2}/\sqrt{5} & \tilde{\lambda}_1 - 1 \end{pmatrix} \tilde{e}_1 = 0 \quad \text{s.t.} \quad \tilde{e}_1'\tilde{e}_1 = 1 \implies \tilde{e}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \implies V_1 = \frac{1}{\sqrt{2}}W_1 + \frac{1}{\sqrt{2}}W_2,$$

$$\begin{pmatrix} \tilde{\lambda}_2 - 1 & -\sqrt{2}/\sqrt{5} \\ -\sqrt{2}/\sqrt{5} & \tilde{\lambda}_2 - 1 \end{pmatrix} \tilde{e}_2 = 0 \quad \text{s.t.} \quad \tilde{e}_2'\tilde{e}_2 = 1 \implies \tilde{e}_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \implies V_2 = \frac{1}{\sqrt{2}}W_1 - \frac{1}{\sqrt{2}}W_2.$$

(e) Compare the components calculated in (d) with those obtained in (b). Are they the same? Should they be?

$$\begin{pmatrix} W_1 \\ W_2 \end{pmatrix} = D_s^{-1} \begin{pmatrix} Y_1 - \mu_1 \\ Y_2 - \mu_2 \end{pmatrix} \implies V_1 = \frac{1}{\sqrt{10}}(Y_1 - \mu_1) + \frac{1}{2}(Y_2 - \mu_2), V_2 = \frac{1}{\sqrt{10}}(Y_1 - \mu_1) - \frac{1}{2}(Y_2 - \mu_2).$$

Thus, v is different from z. Actually, they are not supposed to be the same since $D_s^{-1}\Sigma D_s^{-1}$, the covariance matrix of w, is similar to Σ, the covariance matrix of y, if and only if $D_s = I$.

(f) Find the correlation matrix...

$$corr(z, y) = corr\left(\begin{pmatrix} e_1' \\ e_2' \end{pmatrix} y, y\right) = \begin{pmatrix} \lambda_1^{-1/2} & 0 \\ 0 & \lambda_2^{-1/2} \end{pmatrix} \begin{pmatrix} e_1' \\ e_2' \end{pmatrix} \sum_{j=1}^{2} \lambda_j e_j e_j' D_s^{-1} = \begin{pmatrix} \lambda_1^{1/2} e_1' \\ \lambda_2^{1/2} e_2' \end{pmatrix} D_s^{-1} = \begin{pmatrix} 2\sqrt{6}/5 & \sqrt{3}/\sqrt{5} \\ 1/5 & -2/\sqrt{5} \end{pmatrix},$$

$$corr(v, y) = corr\left(\begin{pmatrix} \tilde{e}_1' \\ \tilde{e}_2' \end{pmatrix} w, D_s w + \mu\right) = \begin{pmatrix} \tilde{\lambda}_1^{-1/2} & 0 \\ 0 & \tilde{\lambda}_2^{-1/2} \end{pmatrix} \begin{pmatrix} \tilde{e}_1' \\ \tilde{e}_2' \end{pmatrix} \sum_{j=1}^{2} \tilde{\lambda}_j \tilde{e}_j \tilde{e}_j' D_s D_s^{-1}$$

$$= \begin{pmatrix} \tilde{\lambda}_1^{1/2} \tilde{e}_1' \\ \tilde{\lambda}_2^{1/2} \tilde{e}_2' \end{pmatrix} = \begin{pmatrix} \tilde{\lambda}_1^{1/2}/\sqrt{2} & \tilde{\lambda}_1^{1/2}/\sqrt{2} \\ \tilde{\lambda}_2^{1/2}/\sqrt{2} & -\tilde{\lambda}_2^{1/2}/\sqrt{2} \end{pmatrix},$$

$$corr(z, v) = corr\left(\begin{pmatrix} e_1' \\ e_2' \end{pmatrix} y, \begin{pmatrix} \tilde{e}_1' \\ \tilde{e}_2' \end{pmatrix} D_s^{-1}(y - \mu)\right) = \begin{pmatrix} \lambda_1^{-1/2} & 0 \\ 0 & \lambda_2^{-1/2} \end{pmatrix} \begin{pmatrix} e_1' \\ e_2' \end{pmatrix} \sum_{j=1}^{2} \lambda_j e_j e_j' D_s^{-1}(\tilde{e}_1, \tilde{e}_2) \begin{pmatrix} \tilde{\lambda}_1^{-1/2} & 0 \\ 0 & \tilde{\lambda}_2^{-1/2} \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1^{1/2} e_1' \\ \lambda_2^{1/2} e_2' \end{pmatrix} D_s^{-1}(\tilde{\lambda}_1^{1/2}\tilde{e}_1, \tilde{\lambda}_2^{1/2}\tilde{e}_2) \begin{pmatrix} \tilde{\lambda}_1^{-1} & 0 \\ 0 & \tilde{\lambda}_2^{-1} \end{pmatrix} = corr(z, y) corr(y, v) \begin{pmatrix} \tilde{\lambda}_1^{-1} & 0 \\ 0 & \tilde{\lambda}_2^{-1} \end{pmatrix}.$$

Thus, there might exist sort of transitivity property among z, y and v.

Problem 3

(a)

```
R = matrix(c(1,0.505,0.569,0.602,0.621,0.603,0.505,1,0.422,
             0.467,0.482,0.450,0.569,0.422,1,0.926,0.877,0.878,
             0.602,0.467,0.926,1,0.874,0.894,0.621,0.482,0.877,
             0.874,1,0.937,0.603,0.450,0.878,0.894,0.937,1),
           nrow = 6,ncol = 6,byrow = TRUE)
evalues=eigen(R)$values
evectors=eigen(R)$vectors
L = eigen(R)$vectors[,1:3]%*%diag(sqrt(evalues[1:3]))
print(round(L,4))
```

```
##         [,1]    [,2]    [,3]
## [1,] -0.7405  0.3500  0.5733
## [2,] -0.6042  0.7206 -0.3399
## [3,] -0.9288 -0.2329 -0.0748
## [4,] -0.9434 -0.1744 -0.0670
## [5,] -0.9476 -0.1428 -0.0446
## [6,] -0.9447 -0.1888 -0.0468
```

(b)

```
psi = diag(diag(R-L%*%t(L)))
print(round(psi,5))
```

```
##         [,1]    [,2]   [,3]   [,4]    [,5]    [,6]
## [1,] 0.00051 0.00000 0.0000 0.0000 0.00000 0.00000
## [2,] 0.00000 0.00018 0.0000 0.0000 0.00000 0.00000
## [3,] 0.00000 0.00000 0.0775 0.0000 0.00000 0.00000
## [4,] 0.00000 0.00000 0.0000 0.0751 0.00000 0.00000
## [5,] 0.00000 0.00000 0.0000 0.0000 0.07967 0.00000
## [6,] 0.00000 0.00000 0.0000 0.0000 0.00000 0.06976
```

(c)

Communalities are the diaganol elements of LL'.

```
diag(L%*%t(L))
```

```
## [1] 0.9994939 0.9998164 0.9225019 0.9248977 0.9203343 0.9302392
```

(d)

```
total = sum(evalues)
diag(L%*%t(L))/total
```

```
## [1] 0.1665823 0.1666361 0.1537503 0.1541496 0.1533891 0.1550399
```

(e)

```
print(round(R-L%*%t(L)-psi, 5))
```

```
##          [,1]     [,2]     [,3]     [,4]     [,5]     [,6]
## [1,]  0.00000  0.00025  0.00557  0.00285 -0.00518 -0.00368
## [2,]  0.00025  0.00000  0.00318 -0.00006 -0.00277 -0.00060
## [3,]  0.00557  0.00318  0.00000  0.00414 -0.03973 -0.04688
## [4,]  0.00285 -0.00006  0.00414  0.00000 -0.04786 -0.03326
## [5,] -0.00518 -0.00277 -0.03973 -0.04786  0.00000  0.01278
## [6,] -0.00368 -0.00060 -0.04688 -0.03326  0.01278  0.00000
```

Comment: all the entries of the residual matrix are pretty small, so we can conclude that the variables can be explained by three factors.

(f)

F_1^*: age; F_2^*: skull breadth; F_3^*: skull length.

Problem 4

(a)

```
library(psych)
data = read.table("pollution.dat")
S = cov(data)
R = cor(data)
S_evalues = eigen(S)$values
S_evalues
```

```
## [1] 304.2578640  28.2761046  11.4644830   2.5243296   1.2795247   0.5287288
## [7]   0.2096157
```

```
R_evalues = eigen(R)$values
R_evalues
```

```
## [1] 2.3367826 1.3860007 1.2040659 0.7270865 0.6534765 0.5366888 0.1558989
```

According to the values of the eigenvalues, we choose one factor for covariance-based method and six for correlation-based method.

```
fac.S = principal(data, nfactors=1, rotate="none", covar=F, cor="cov")
fac.S
```

```
## Principal Components Analysis
## Call: principal(r = data, nfactors = 1, rotate = "none", covar = F,
##     cor = "cov")
## Unstandardized loadings (pattern matrix) based upon covariance matrix
##      PC1      h2    u2     H2     U2
## V1 -0.18 3.1e-02  2.47 0.0123 0.9877
## V2 17.32 3.0e+02  0.38 0.9987 0.0013
## V3  0.25 6.0e-02  1.46 0.0395 0.9605
## V4 -0.08 6.8e-03  1.18 0.0057 0.9943
```

```
## V5  0.42 1.8e-01 11.18 0.0158 0.9842
## V6  1.96 3.8e+00 27.13 0.1241 0.8759
## V7  0.04 1.7e-03  0.48 0.0035 0.9965
##
##                     PC1
## SS loadings    304.26
## Proportion Var   0.87
##
##  Standardized loadings (pattern matrix)
##    V   PC1     h2     u2
## V1 1 -0.11 0.0123 0.9877
## V2 2     1 0.9987 0.0013
## V3 3   0.2 0.0395 0.9605
## V4 4 -0.08 0.0057 0.9943
## V5 5  0.13 0.0158 0.9842
## V6 6  0.35 0.1241 0.8759
## V7 7  0.06 0.0035 0.9965
##
##                 PC1
## SS loadings    6.11
## Proportion Var 0.87
##
## Mean item complexity =  1
## Test of the hypothesis that 1 component is sufficient.
##
## The root mean square of the residuals (RMSR) is  1.27
##  with the empirical chi square  2857.9  with prob <  0
##
## Fit based upon off diagonal values = 1
```

```r
fac.R = principal(data, nfactors=6, rotate="none", covar=F, cor="cor")
fac.R
```

```
## Principal Components Analysis
## Call: principal(r = data, nfactors = 6, rotate = "none", covar = F,
##     cor = "cor")
## Standardized loadings (pattern matrix) based upon correlation matrix
##      PC1   PC2   PC3   PC4   PC5   PC6   h2      u2 com
## V1 -0.36  0.33  0.71  0.15  0.45  0.16 0.99 0.00909 3.1
## V2  0.31 -0.62  0.25  0.66 -0.13  0.00 1.00 0.00002 2.8
## V3  0.84 -0.01 -0.12  0.00  0.46  0.08 0.95 0.05340 1.6
## V4  0.58  0.51 -0.45  0.25 -0.05  0.33 0.97 0.03312 4.0
## V5  0.76  0.24  0.22 -0.04  0.04 -0.55 0.98 0.01779 2.2
## V6  0.50 -0.67  0.18 -0.43  0.06  0.24 0.97 0.02712 3.2
## V7  0.49  0.36  0.59 -0.12 -0.46  0.20 0.98 0.01536 4.0
##
##                        PC1  PC2  PC3  PC4  PC5  PC6
## SS loadings           2.34 1.39 1.20 0.73 0.65 0.54
## Proportion Var        0.33 0.20 0.17 0.10 0.09 0.08
## Cumulative Var        0.33 0.53 0.70 0.81 0.90 0.98
## Proportion Explained  0.34 0.20 0.18 0.11 0.10 0.08
```

```
## Cumulative Proportion 0.34 0.54 0.72 0.83 0.92 1.00
##
## Mean item complexity =  3
## Test of the hypothesis that 6 components are sufficient.
##
## The root mean square of the residuals (RMSR) is  0.02
##  with the empirical chi square  0.8  with prob <  NA
##
## Fit based upon off diagonal values = 0.99
```

(b)

```r
d = data[c(1,2,5,6)]
principal(d, nfactors=1, rotate="none", covar=F)
```

```
## Principal Components Analysis
## Call: principal(r = d, nfactors = 1, rotate = "none", covar = F)
## Standardized loadings (pattern matrix) based upon correlation matrix
##      PC1   h2   u2 com
## V1 -0.56 0.32 0.68   1
## V2  0.65 0.42 0.58   1
## V5  0.48 0.23 0.77   1
## V6  0.77 0.59 0.41   1
##
##                 PC1
## SS loadings    1.56
## Proportion Var 0.39
##
## Mean item complexity =  1
## Test of the hypothesis that 1 component is sufficient.
##
## The root mean square of the residuals (RMSR) is  0.2
##  with the empirical chi square  19.81  with prob <  5e-05
##
## Fit based upon off diagonal values = -0.03
```

```r
principal(d, nfactors=2, rotate="none", covar=F)
```

```
## Principal Components Analysis
## Call: principal(r = d, nfactors = 2, rotate = "none", covar = F)
## Standardized loadings (pattern matrix) based upon correlation matrix
##      PC1   PC2   h2   u2 com
## V1 -0.56  0.24 0.38 0.62 1.4
## V2  0.65  0.52 0.69 0.31 1.9
## V5  0.48 -0.74 0.77 0.23 1.7
## V6  0.77  0.20 0.63 0.37 1.1
##
##                        PC1  PC2
## SS loadings           1.56 0.91
## Proportion Var        0.39 0.23
## Cumulative Var        0.39 0.62
```

```
## Proportion Explained  0.63 0.37
## Cumulative Proportion 0.63 1.00
##
## Mean item complexity =  1.5
## Test of the hypothesis that 2 components are sufficient.
##
## The root mean square of the residuals (RMSR) is  0.21
##  with the empirical chi square  22.56  with prob <  NA
##
## Fit based upon off diagonal values = -0.17
```

(c)

```r
factanal(d, factors=1, rotation = "none")
```

```
##
## Call:
## factanal(x = d, factors = 1, rotation = "none")
##
## Uniquenesses:
##    V1    V2    V5    V6
## 0.895 0.832 0.946 0.405
##
## Loadings:
##    Factor1
## V1 -0.324
## V2  0.410
## V5  0.232
## V6  0.771
##
##                Factor1
## SS loadings      0.921
## Proportion Var   0.230
##
## Test of the hypothesis that 1 factor is sufficient.
## The chi square statistic is 0.15 on 2 degrees of freedom.
## The p-value is 0.93
```

(d)

Based on principle component method, we have

$$L = (-0.56, 0.65, 0.48, 0.77)'$$

; based on MLE, we have

$$L = (-0.324, 0.410, 0.232, 0.771)'$$

, so based on different methods, we can obtain different results. And SS loadings for principle component method is larger than that for MLE method.

(e)

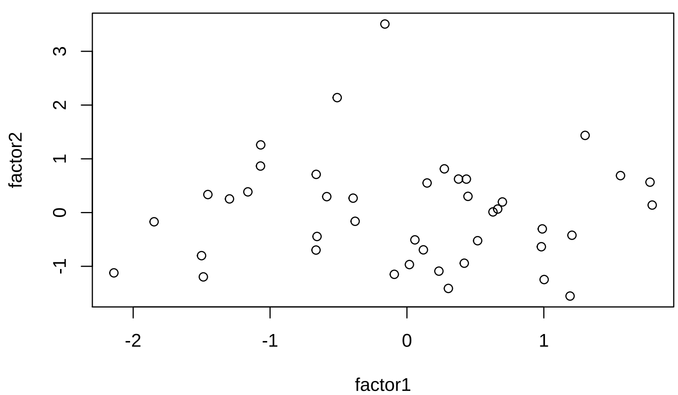
```
fac <- principal(d, nfactors=2, rotate="varimax", covar=F)
fac$loadings
```

```
##
## Loadings:
##    RC1    RC2
## V1 -0.313 -0.528
## V2  0.828
## V5         0.875
## V6  0.739  0.295
##
##                  RC1   RC2
## SS loadings    1.332 1.133
## Proportion Var 0.333 0.283
## Cumulative Var 0.333 0.616
```

We can see that RC1 and RC2 can account for 61.6% of the variance. And RC1 can be explained as the strength of the sun, and RC2 can be explained as the intensity of NO2.

(f)

```
factor1 = fac$scores[,1]
factor2 = fac$scores[,2]
plot(factor1,factor2)
```



The points in the first (third) quadrant face the strong(weak) sun and intensive(thin) NO2; the points in the second (forth) quadrant face the weak(strong) sun and intensive(thin) NO2.