Homework 4

Luo Bingjun 2017013573 Software 71

2019/4/28

8.1

Solve the equation

$$|\Sigma - \lambda I| = 0$$

and we can obtain

$$\lambda_1 = 6, e_1 = (\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})'$$

$$\lambda_2 = 1, e_2 = (-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})'$$

Hence

$$Y_1 = e_1'X = \frac{2}{\sqrt{5}}X_1 + \frac{1}{\sqrt{5}}X_2, Var(Y_1) = \lambda_1 = 6$$

$$Y_2 = e_2'X = -\frac{1}{\sqrt{5}}X_1 + \frac{2}{\sqrt{5}}X_2, Var(Y_2) = \lambda_2 = 1$$

the proportion of the first principal component is $\frac{Var(Y_1)}{Var(Y_1)+Var(Y_2)} = \frac{6}{7} = 0.86$.

8.2

From Σ we can obtain

$$D = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$$

Hence

$$\rho = D^{-1/2}\Sigma D^{-1/2} = \begin{bmatrix} 1 & \frac{\sqrt{10}}{5} \\ \frac{\sqrt{10}}{5} & 1 \end{bmatrix}$$

(a)

Solve the equation

$$|\rho - \lambda I| = 0$$

and we can obtain

$$\lambda_1 = 1 + \sqrt{\frac{2}{5}}, e_1 = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})'$$

$$\lambda_2 = 1 - \sqrt{\frac{2}{5}}, e_1 = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})'$$

Hence

$$Y_1 = e_1'Z = \frac{\sqrt{2}}{2}Z_1 + \frac{\sqrt{2}}{2}Z_2$$

$$Y_2 = e_2'Z = \frac{\sqrt{2}}{2}Z_1 - \frac{\sqrt{2}}{2}Z_2$$

the proportion of Y_1 is

$$\frac{\lambda_1}{p} = \frac{1}{2} + \frac{1}{\sqrt{10}} = 0.82$$

(b)

No, because Z_1 and Z_2 in 8.2 is regarded as contribute equally because they have the same variance while X_1 and X_2 in 8.1 cannot be regarded as contribute equally for their different variance.

(c)

$$\rho_{Y_1,Z_1} = e_{11}\sqrt{\lambda_1} = 0.90$$
$$\rho_{Y_1,Z_2} = e_{12}\sqrt{\lambda_1} = 0.90$$
$$\rho_{Y_2,Z_1} = e_{21}\sqrt{\lambda_2} = 0.43$$

8.14

The analysis result:

```
data=read.table("T5-1.DAT")
n = dim(data)[1]
p = dim(data)[2]
colnames(data) <- c("X1","X2","X3")
pca=prcomp(data)
pca
```

```
## Standard deviations (1, .., p=3):
## [1] 14.158477  2.128753  1.140786
##
## Rotation (n x k) = (3 x 3):
##            PC1         PC2         PC3
## X1  0.05084144 -0.57370364 -0.81748351
## X2  0.99828352  0.05302042  0.02487655
## X3 -0.02907156  0.81734508 -0.57541452
```

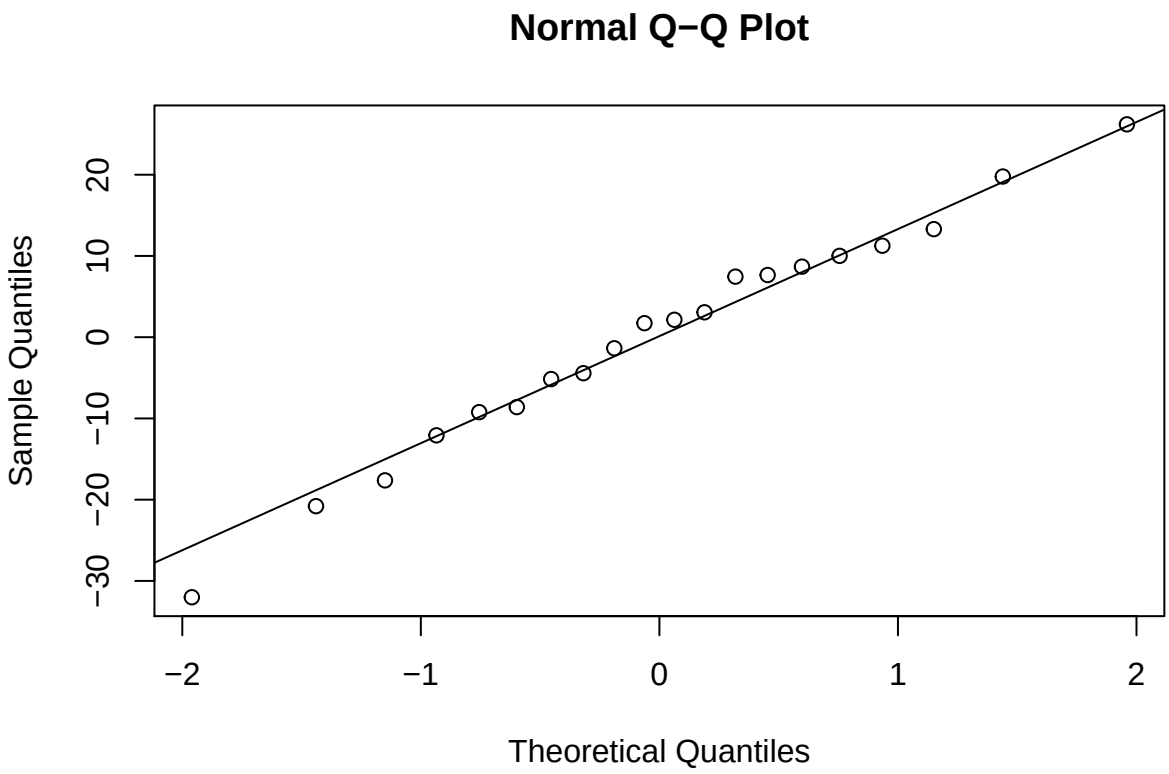
We should keep one principle component. The proportion of total variance explained by PC1 is

```
(pca$sdev^2)[1]/sum(pca$sdev^2)
```

```
## [1] 0.9717251
```

The Q-Q plot of PC1:

```
qqnorm(pca$x[,1])
qqline(pca$x[,1])
```



There appears to be no suspect observations in the Q-Q plot.