Modern Applyed Statistics(Chap 11)

Library Package

```
library(MASS)
```

Data input

```
# 1. Iris data
data(iris3)
head(iris)
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1         3.5          1.4         0.2  setosa
## 2          4.9         3.0          1.4         0.2  setosa
## 3          4.7         3.2          1.3         0.2  setosa
## 4          4.6         3.1          1.5         0.2  setosa
## 5          5.0         3.6          1.4         0.2  setosa
## 6          5.4         3.9          1.7         0.4  setosa
```

```
# 2. Leptograpsus variegatus crabs data
data("crabs")
head(crabs)
```

```
##   sp sex index   FL  RW   CL   CW  BD
## 1  B   M     1  8.1 6.7 16.1 19.0 7.0
## 2  B   M     2  8.8 7.7 18.1 20.8 7.4
## 3  B   M     3  9.2 7.8 19.0 22.4 7.7
## 4  B   M     4  9.6 7.9 20.1 23.1 8.2
## 5  B   M     5  9.8 8.0 20.3 23.0 8.2
## 6  B   M     6 10.8 9.0 23.0 26.5 9.8
```

1. Visualization Methods

1-1) Principal component analysis

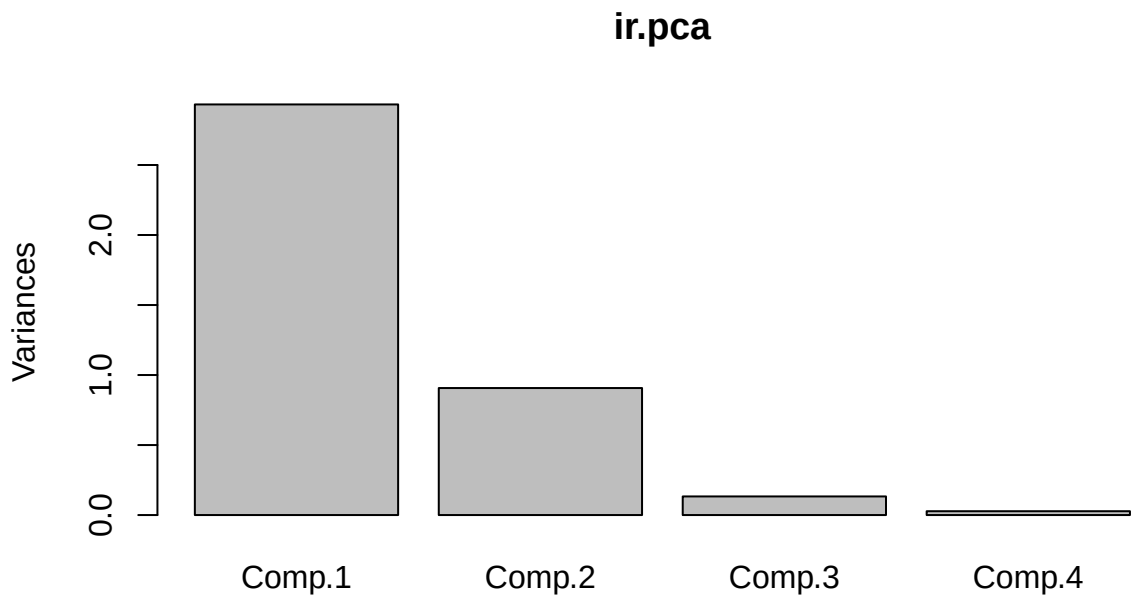
```
# Iris Data
ir <- rbind(iris3[,,1], iris3[,,2], iris3[,,3])
ir.species <- factor(c(rep("s", 50), rep("c", 50), rep("v", 50)))
(ir.pca <- princomp(log(ir), cor = T))
```

```
## Call:
## princomp(x = log(ir), cor = T)
##
## Standard deviations:
##    Comp.1    Comp.2    Comp.3    Comp.4
## 1.7124583 0.9523797 0.3647029 0.1656840
##
##  4  variables and  150 observations.
```

```
summary(ir.pca)
```

```
## Importance of components:
##                           Comp.1    Comp.2     Comp.3    Comp.4
## Standard deviation     1.7124583 0.9523797 0.36470294 0.1656840
## Proportion of Variance 0.7331284 0.2267568 0.03325206 0.0068628
## Cumulative Proportion  0.7331284 0.9598851 0.99313720 1.0000000
```

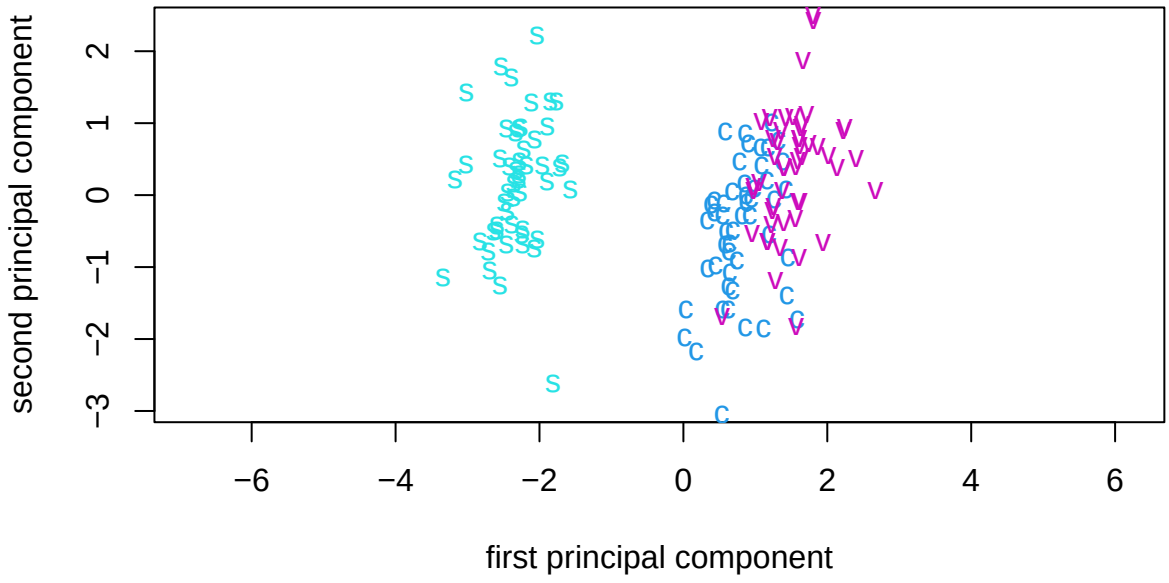
```
plot(ir.pca)
```



```
loadings(ir.pca)
```

```
##
## Loadings:
##          Comp.1 Comp.2 Comp.3 Comp.4
## Sepal L.  0.504  0.455  0.709  0.191
## Sepal W. -0.302  0.889 -0.331
## Petal L.  0.577         -0.219 -0.786
## Petal W.  0.567         -0.583  0.580
##
##                Comp.1 Comp.2 Comp.3 Comp.4
## SS loadings      1.00   1.00   1.00   1.00
## Proportion Var   0.25   0.25   0.25   0.25
## Cumulative Var   0.25   0.50   0.75   1.00
```

```
ir.pc <- predict(ir.pca)
eqscplot(ir.pc[, 1:2], type = "n",
         xlab = "first principal component",
         ylab = "second principal component")
text(ir.pc[, 1:2], labels = as.character(ir.species),
     col = 3 + as.integer(ir.species))
```



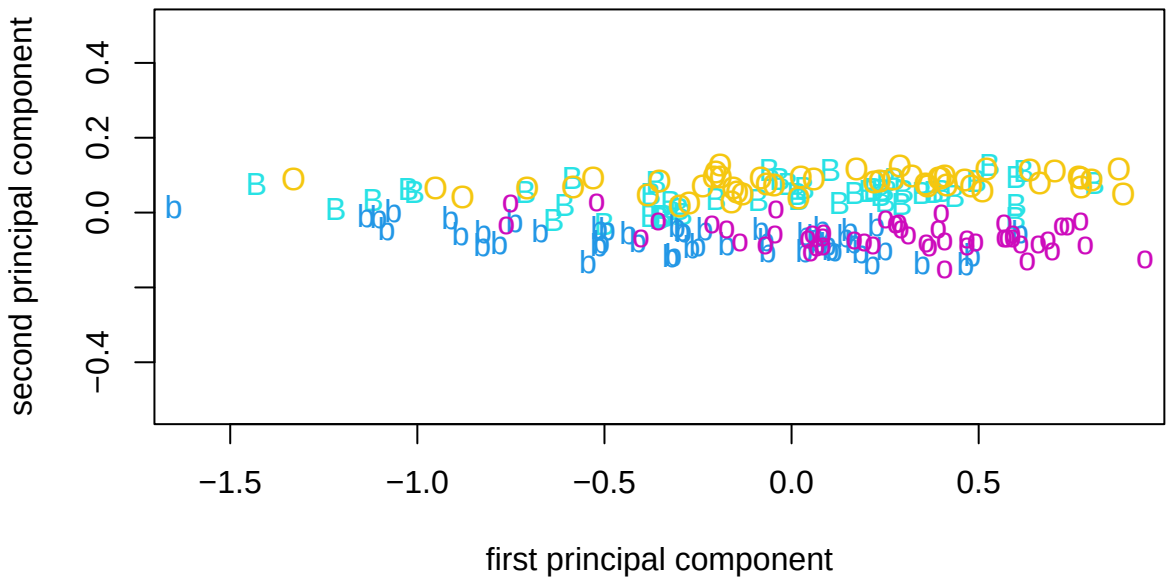

```
# Leptograpsus variegatus crabs dat
lcrabs <- log(crabs[,4:8])
crabs.grp <- factor(c("B", "b", "O", "o")[rep(1:4, each = 50)])
(lcrabs.pca <- princomp(lcrabs))
```

```
## Call:
## princomp(x = lcrabs)
##
## Standard deviations:
##      Comp.1      Comp.2      Comp.3      Comp.4      Comp.5
## 0.516640451 0.074653581 0.047914392 0.024804021 0.009052189
##
##  5  variables and  200 observations.
```

```
loadings(lcrabs.pca)
```

```
##
## Loadings:
##    Comp.1 Comp.2 Comp.3 Comp.4 Comp.5
## FL  0.452  0.157  0.438  0.752  0.114
## RW  0.387 -0.911
## CL  0.453  0.204 -0.371        -0.784
## CW  0.440        -0.672         0.591
## BD  0.497  0.315  0.458 -0.652  0.136
##
##                Comp.1 Comp.2 Comp.3 Comp.4 Comp.5
## SS loadings       1.0    1.0    1.0    1.0    1.0
## Proportion Var    0.2    0.2    0.2    0.2    0.2
## Cumulative Var    0.2    0.4    0.6    0.8    1.0
```

```
lcrabs.pc <- predict(lcrabs.pca)
dimnames(lcrabs.pc) <- list(NULL, paste("PC", 1:5, sep = ""))
eqscplot(lcrabs.pc[, 1:2], type = "n",
         xlab = "first principal component",
         ylab = "second principal component")
text(lcrabs.pc[, 1:2], labels = as.character(crabs.grp),
     col = 3 + as.integer(crabs.grp))
```


second principal component
0.4
0.2
0.0
−0.4
−1.5
−1.0
−0.5
0.0
0.5
first principal component