

# TP2

*You ZUO*

*2019/9/23*

## exo1

```
library(mclust)

## Package 'mclust' version 5.4.3
## Type 'citation("mclust")' for citing this R package in publications.

library(nnet)
data("iris")
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

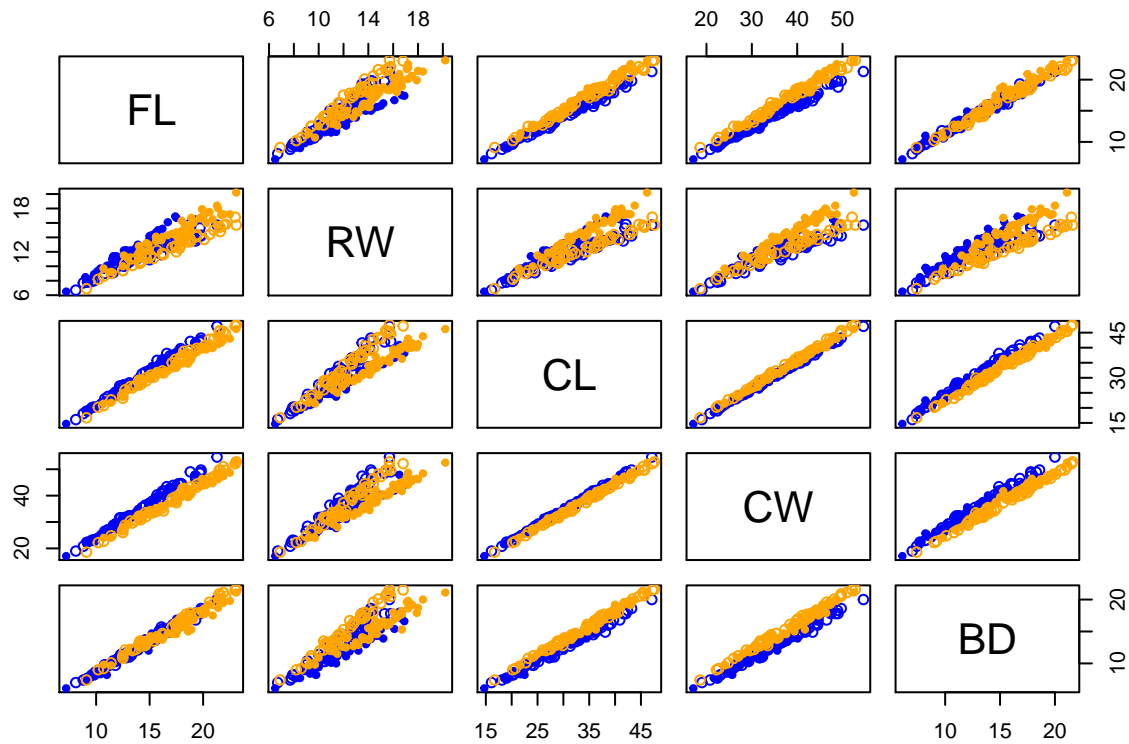
C <- class.ind(cl = iris$Species)
X <- matrix(iris[, -5])

mu_T <- sapply(iris[, 1:4], mean)

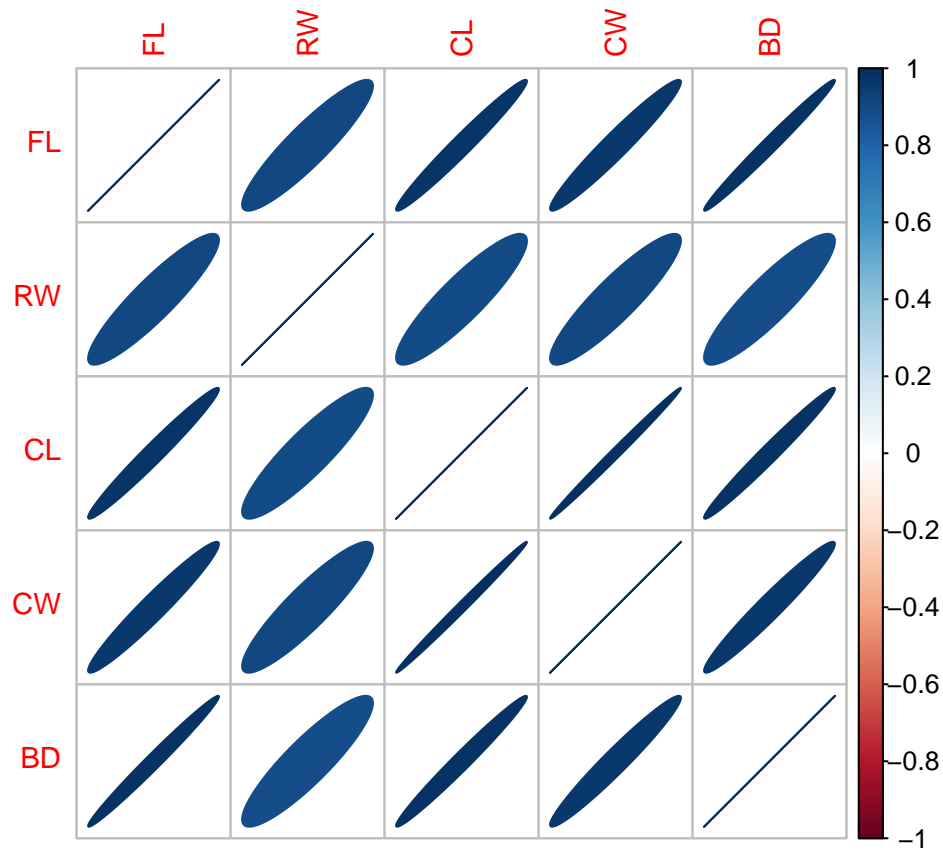
n_k <- sapply(1:3, function(i){
  sum(C[, i])
})
mu_k <- t(as.matrix(C)) %*% as.matrix(iris[, 1:4]) / n_k
```

## exo4

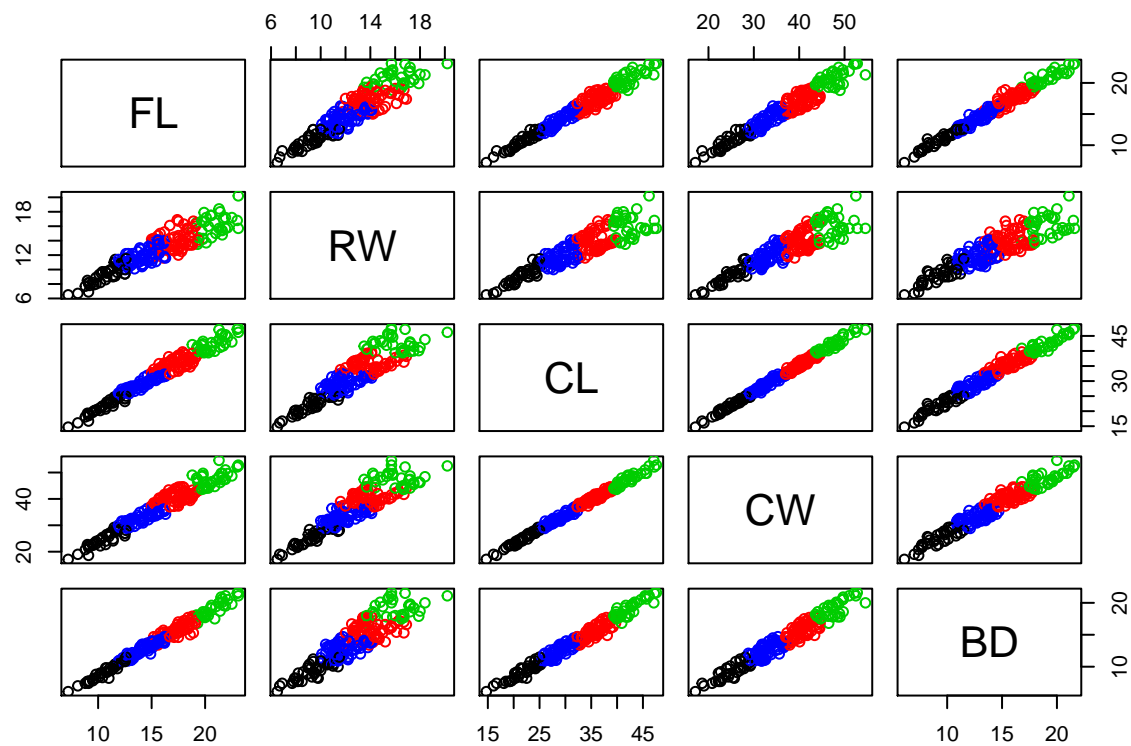
```
library(MASS)
data(crabs)
help(crabs)
pairs(x = crabs[, 4:8], col = c("blue", "orange")[crabs$sp], pch = c(20, 21)[crabs$sex])
```



```
cor <- cor(crabs[,4:8])
corrplot::corrplot(cor, method = "ellipse", tl.cex = 0.9)
```

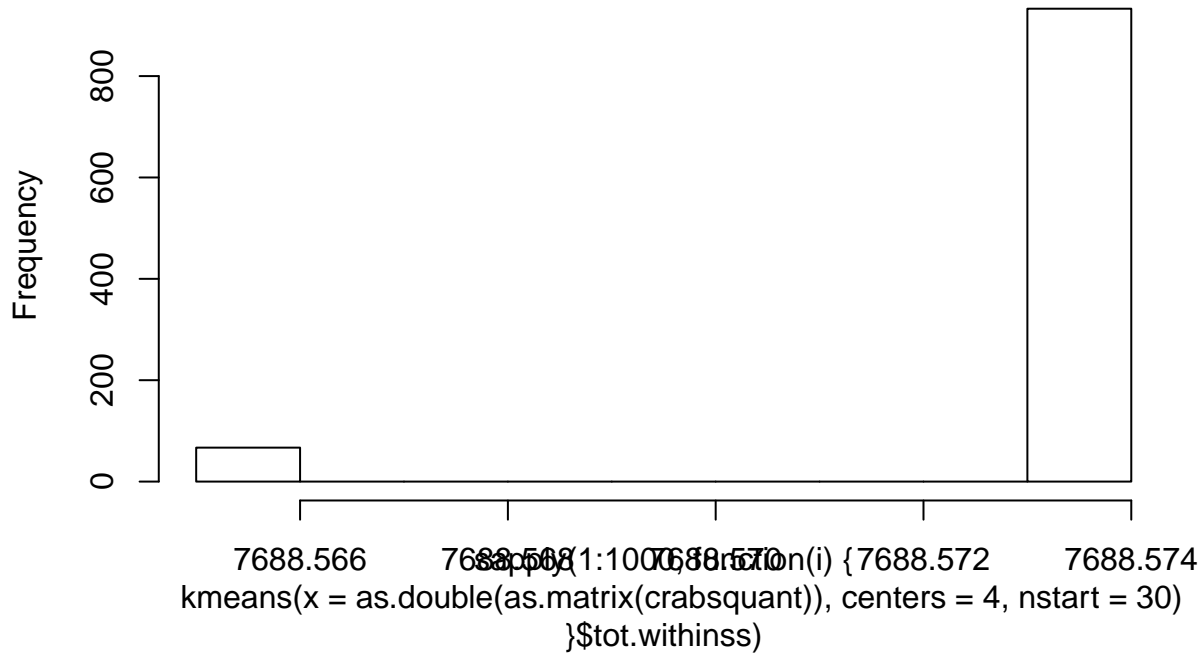


```
km4 <- kmeans(x = crabs[,4:8], nstart = 1, centers = 4)
plot(x = crabs[,4:8], col = km4$cluster)
points(km4$centers, col = 1:5, pch = 8)
```



```
crabsquant <- crabs[,4:8]
hist(sapply(1:1000, function(i){kmeans(x = as.double(as.matrix(crabsquant)), centers = 4, nstart = 30)}))
```

# **Histogram of `supply(1:1000, function(i) { kmeans(x = as.double(as.matrix(crabsquant)), centers = 4, nstart = }$tot.withinss)`**



```
crabsquant2 <- crabsquant[, -3] / crabsquant$CL
names(crabsquant2) <- paste(names(crabsquant[, -3]), "/CL")

kmeans(crabsquant2, 4, nstart = 30)$cluster -> cluster
natural.classes <- paste(crabs$sex, crabs$sp, sep = "-")
table(cluster, natural.classes)
```

```
##      natural.classes
## cluster F-B F-O M-B M-O
##      1  0  0  0 44
##      2  0  0 40  0
##      3  0 50  0  6
##      4 50  0 10  0
```