

Análisis Discriminante Lineal

Oscar Elí Bonilla Morales

2022-05-25

Matriz

```
library(MASS)
```

Se cargan los datos iris

```
Z<-as.data.frame(iris)
```

Se define la matriz de datos y la variable

```
x<-Z[,1:4]  
y<-Z[,5]
```

Definir como n y p el numero de flores y variables

```
n<-nrow(x)  
p<-ncol(x)
```

Se aplica el Analisis discriminante lineal (LDA)

Cross validation (cv): clasificacion optima

```
lda.iris<-lda(y~.,data=x, CV=TRUE)
```

lda.iris\$class contiene las clasificaciones hechas por CV usando LDA.

```
(lda.iris$class)
```

```
## [1] setosa setosa setosa setosa setosa setosa  
## [7] setosa setosa setosa setosa setosa setosa  
## [13] setosa setosa setosa setosa setosa setosa  
## [19] setosa setosa setosa setosa setosa setosa  
## [25] setosa setosa setosa setosa setosa setosa  
## [31] setosa setosa setosa setosa setosa setosa
```

```
## [37] setosa      setosa      setosa      setosa      setosa      setosa
## [43] setosa      setosa      setosa      setosa      setosa      setosa
## [49] setosa      setosa      versicolor  versicolor  versicolor  versicolor
## [55] versicolor  versicolor  versicolor  versicolor  versicolor  versicolor
## [61] versicolor  versicolor  versicolor  versicolor  versicolor  versicolor
## [67] versicolor  versicolor  versicolor  versicolor  virginica   versicolor
## [73] versicolor  versicolor  versicolor  versicolor  versicolor  versicolor
## [79] versicolor  versicolor  versicolor  versicolor  versicolor  virginica
## [85] versicolor  versicolor  versicolor  versicolor  versicolor  versicolor
## [91] versicolor  versicolor  versicolor  versicolor  versicolor  versicolor
## [97] versicolor  versicolor  versicolor  versicolor  virginica   virginica
## [103] virginica   virginica   virginica   virginica   virginica   virginica
## [109] virginica   virginica   virginica   virginica   virginica   virginica
## [115] virginica   virginica   virginica   virginica   virginica   virginica
## [121] virginica   virginica   virginica   virginica   virginica   virginica
## [127] virginica   virginica   virginica   virginica   virginica   virginica
## [133] virginica   versicolor  virginica   virginica   virginica   virginica
## [139] virginica   virginica   virginica   virginica   virginica   virginica
## [145] virginica   virginica   virginica   virginica   virginica   virginica
## Levels: setosa versicolor virginica
```

Creacion de la tabla de clasificaciones buenas y malas

```
table.lda<-table(y,lda.iris$class)
table.lda
```

```
##
## y          setosa versicolor virginica
## setosa      50          0          0
## versicolor   0         48          2
## virginica    0          1         49
```

Proporcion de errores

```
mis.lda<- n-sum(y==lda.iris$class)
mis.lda/n
```

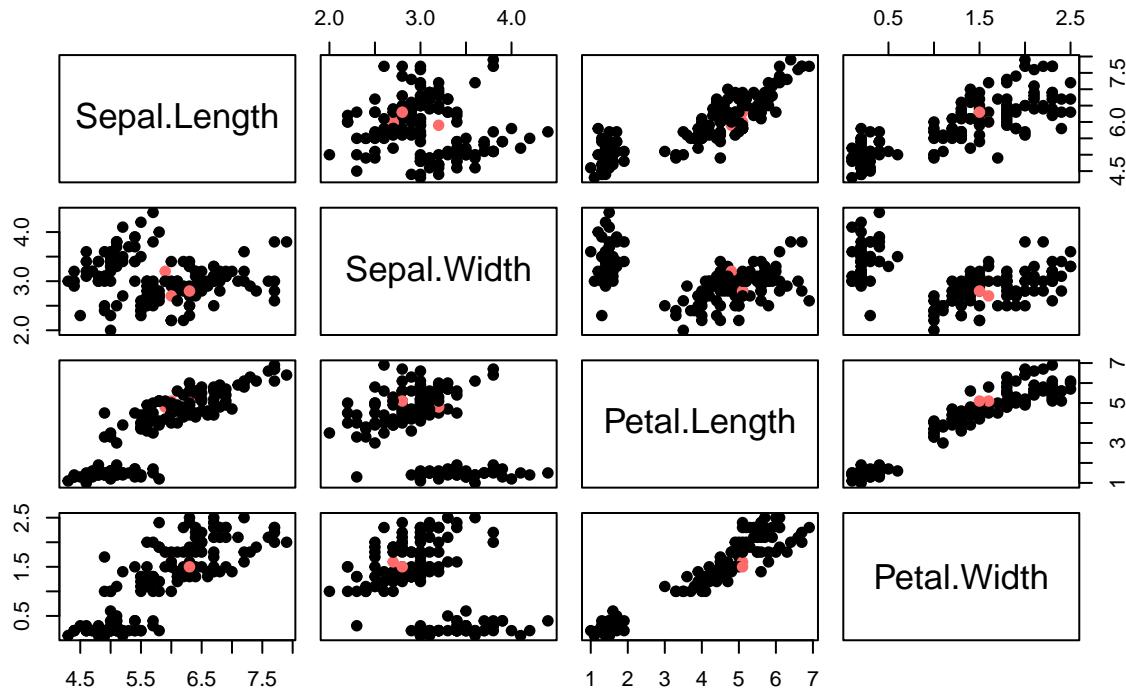
```
## [1] 0.02
```

scater plot

Buenas clasificaciones en negro y malas en rojo

```
col.lda.iris<-c("indianred1","black")[1*(y==lda.iris$class)+1]
pairs(x,main="Buena Clasificacion (negro), Mala Clasificacion (rojo)",
      pch=19,col=col.lda.iris)
```

Buena Clasificacion (negro), Mala Clasificacion (rojo)



Probabilidad de pertenencia a uno de los tres grupos

```
head(lda.iris$posterior)
```

```
##   setosa  versicolor  virginica
## 1      1 5.087494e-22 4.385241e-42
## 2      1 9.588256e-18 8.888069e-37
## 3      1 1.983745e-19 8.606982e-39
## 4      1 1.505573e-16 5.101765e-35
## 5      1 2.075670e-22 1.739832e-42
## 6      1 5.332271e-21 8.674906e-40
```

Grafico de probabilidades

```
plot(1:n, lda.iris$posterior[,1],
     main="Probabilidades a posterior",
     pch=20, col="cyan",
     xlab="Numero de observaciones", ylab="Probabilidades")
points(1:n, lda.iris$posterior[,2],
       pch=20, col="green")
points(1:n, lda.iris$posterior[,3],
       pch=20, col="orange")
```

Probabilidades a posterior

