

# 判别分析

R 语言

徐平峰

长春工业大学

2021/3/30

# Fisher线性判别

# 天气数据

- 利用今天和昨天湿度差x1, 气温差x2, 预报今天x1=8.1, x2=2.0是否下雨
- 1 为下雨, 2为晴天

```
(w3=read.csv("E:/teaching_plan_notes/msa11091083/rmd/MVAPureData/
```

```
#      G      x1      x2
# 1  1 -1.9   3.2
# 2  1 -6.9   0.4
# 3  1  5.2   2.0
# 4  1  5.0   2.5
# 5  1  7.3   0.0
# 6  1  6.8  12.7
# 7  1  0.9  -5.4
# 8  1 -12.5 -2.5
# 9  1  1.5   1.3
# 10 1  3.8   6.8
# 11 2  0.2   6.2
# 12 2 -0.1   7.5
# 13 2  0.4  14.6
# 14 2  2.7   8.3
# 15 2  2.1   0.8
# 16 2 -4.6   4.3
# 17 2 -1.7  10.9
# 18 2 -2.6  13.1
# 19 2  2.6  12.8
```

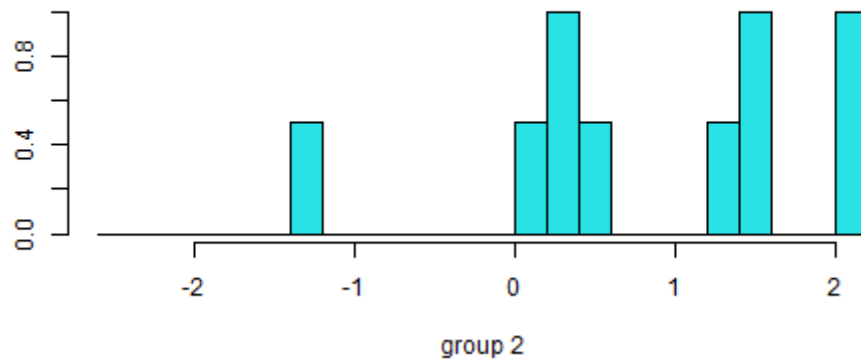
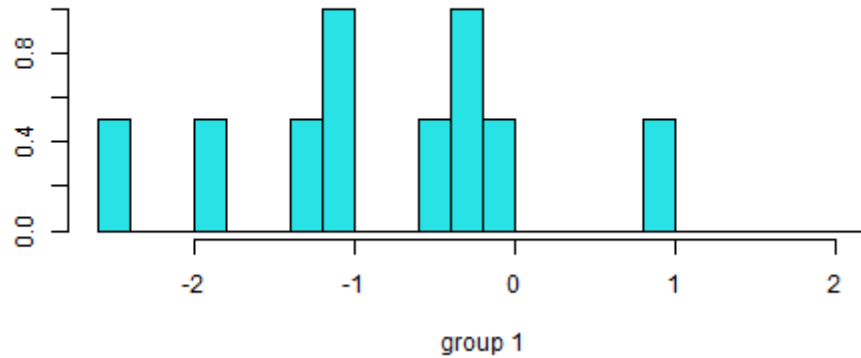
- Fisher判别

```
library(MASS)
(ld1<-lda(G~x1+x2, data=w3))
```

```
# call:
# lda(G ~ x1 + x2, data = w3)
#
# Prior probabilities of groups:
#   1   2
# 0.5 0.5
#
# Group means:
#      x1    x2
# 1  0.92 2.10
# 2 -0.38 8.85
#
# Coefficients of linear discriminants:
#           LD1
# x1 -0.1035305
# x2  0.2247957
```

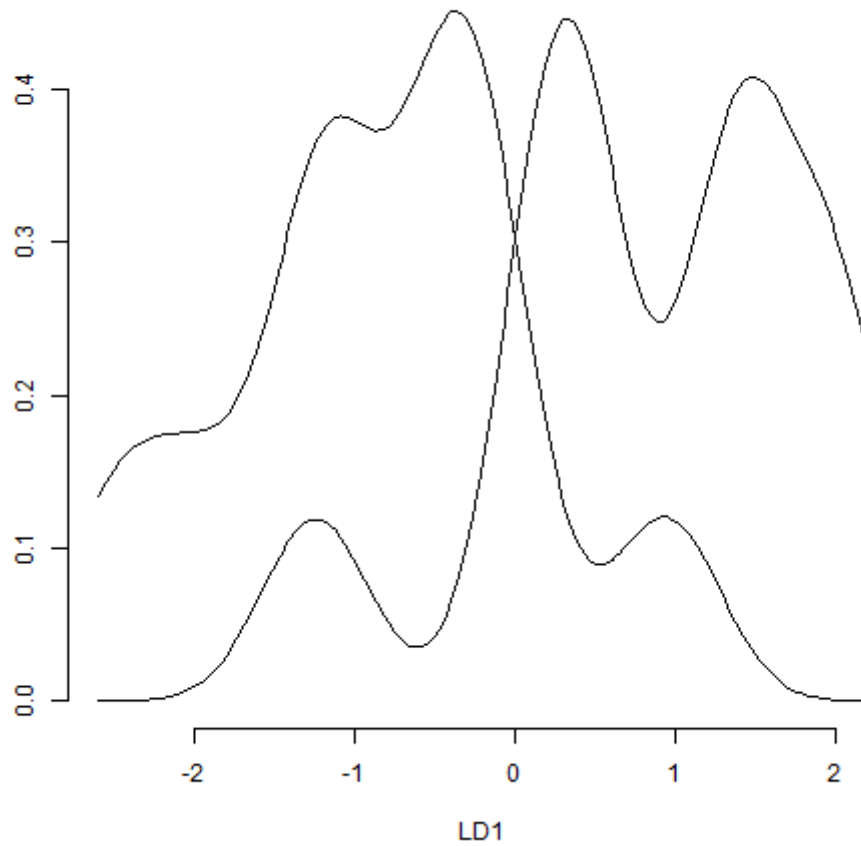
- Fisher判别

```
plot(ld1) #两类样本经过线性变换后的直方图
```



- Fisher判别

```
plot(ld1, type="density", dimen=1) #两类样本经过线性变换后的密度图
```



- Fisher判别

```
p.ld1<-predict(ld1)
p.ld1$class
```

```
# [1] 1 1 1 1 1 2 1 1 1 1 2 2 2 2 1 2 2 2 2 2
# Levels: 1 2
```

```
table(w3$G, p.ld1$class)#混淆矩阵
```

```
#
#      1 2
#      1 9 1
#      2 1 9
```

```
predict(ld1, newdata=data.frame(x1=8.1, x2=2.0))#对新的值进行预测
```

```
# $class
# [1] 1
# Levels: 1 2
#
# $posterior
#           1           2
# 1 0.9327428 0.06725717
#
# $x
#           LD1
# 1 -1.591809
```

# 鸢尾花

- R 中iris（鸢尾花）数据，三种不同的鸢尾花的 150 个样品的花瓣、花萼长、宽的数据。

```
dim(iris)
```

```
## [1] 150  5
```

```
summary(iris)
```

```
##      Sepal.Length      Sepal.width      Petal.Length      Petal.width
##  Min.       :4.300    Min.       :2.000    Min.       :1.000    Min.       :0.100
##  1st Qu.:5.100      1st Qu.:2.800    1st Qu.:1.600    1st Qu.:0.300
##  Median :5.800      Median :3.000    Median :4.350    Median :1.300
##  Mean    :5.843      Mean    :3.057    Mean    :3.758    Mean    :1.199
##  3rd Qu.:6.400      3rd Qu.:3.300    3rd Qu.:5.100    3rd Qu.:1.800
##  Max.    :7.900      Max.    :4.400    Max.    :6.900    Max.    :2.500
##           Species
##  setosa      :50
##  versicolor:50
##  virginica   :50
##
##
##
```



- Fisher 判别

```
library(MASS)
(ld2<-lda(Species~., data=iris))
```

```
## call:
## lda(Species ~ ., data = iris)
##
## Prior probabilities of groups:
##      setosa versicolor virginica
## 0.3333333 0.3333333 0.3333333
##
## Group means:
##      Sepal.Length Sepal.width Petal.Length Petal.width
## setosa           5.006         3.428         1.462         0.246
## versicolor       5.936         2.770         4.260         1.326
## virginica        6.588         2.974         5.552         2.026
##
## Coefficients of linear discriminants:
##              LD1          LD2
## Sepal.Length 0.8293776 0.02410215
## Sepal.width  1.5344731 2.16452123
## Petal.Length -2.2012117 -0.93192121
## Petal.width  -2.8104603 2.83918785
##
## Proportion of trace:
##      LD1      LD2
## 0.9912 0.0088
```

```
names(ld2)
```

```
## [1] "prior" "counts" "means" "scaling" "lev" "svd" "N"  
## [8] "call" "terms" "xlevels"
```

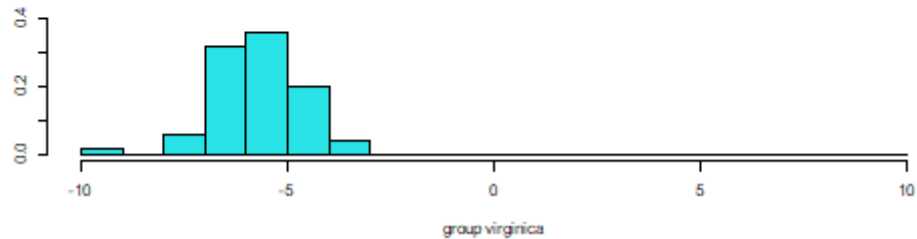
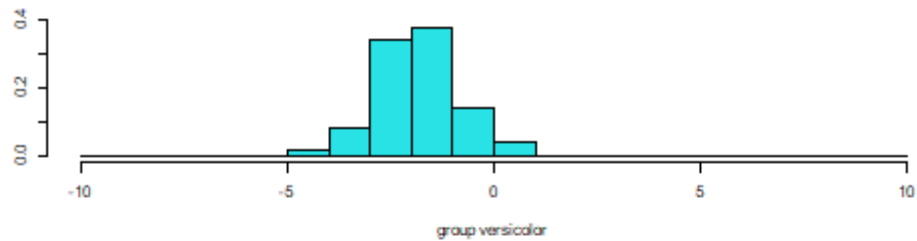
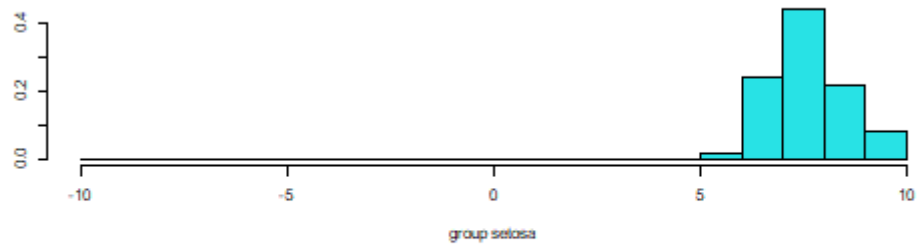
```
ld2$scaling
```

```
##  
## LD1 LD2  
## Sepal.Length 0.8293776 0.02410215  
## Sepal.width 1.5344731 2.16452123  
## Petal.Length -2.2012117 -0.93192121  
## Petal.width -2.8104603 2.83918785
```

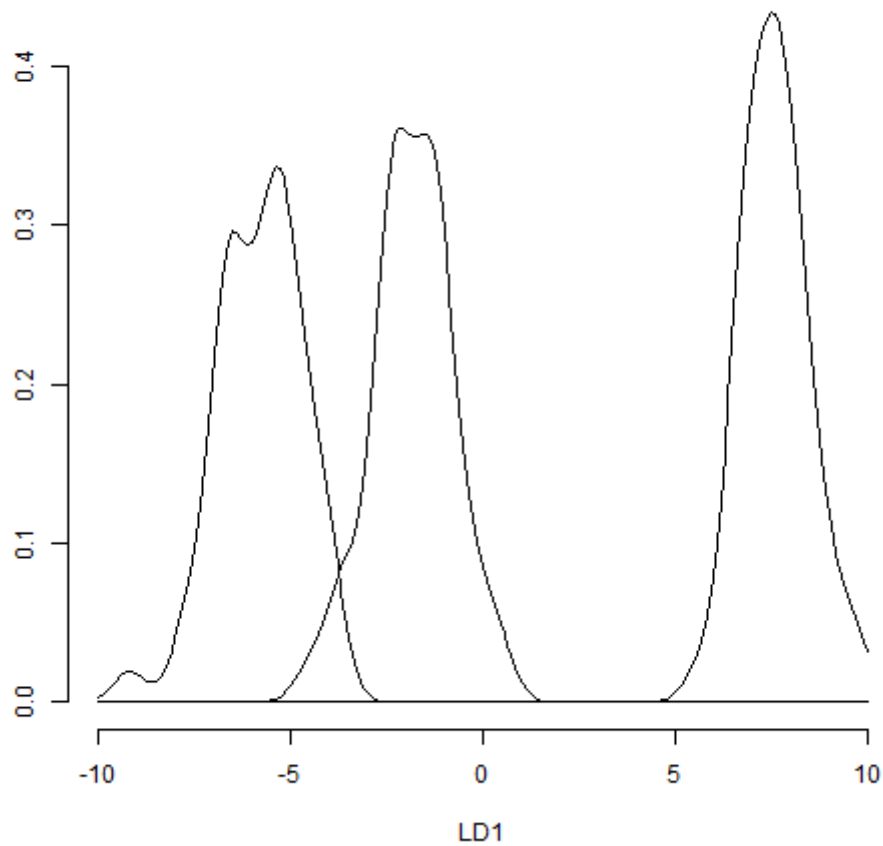
```
p.ld2<-predict(ld2)  
table(iris$Species, p.ld2$class)
```

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

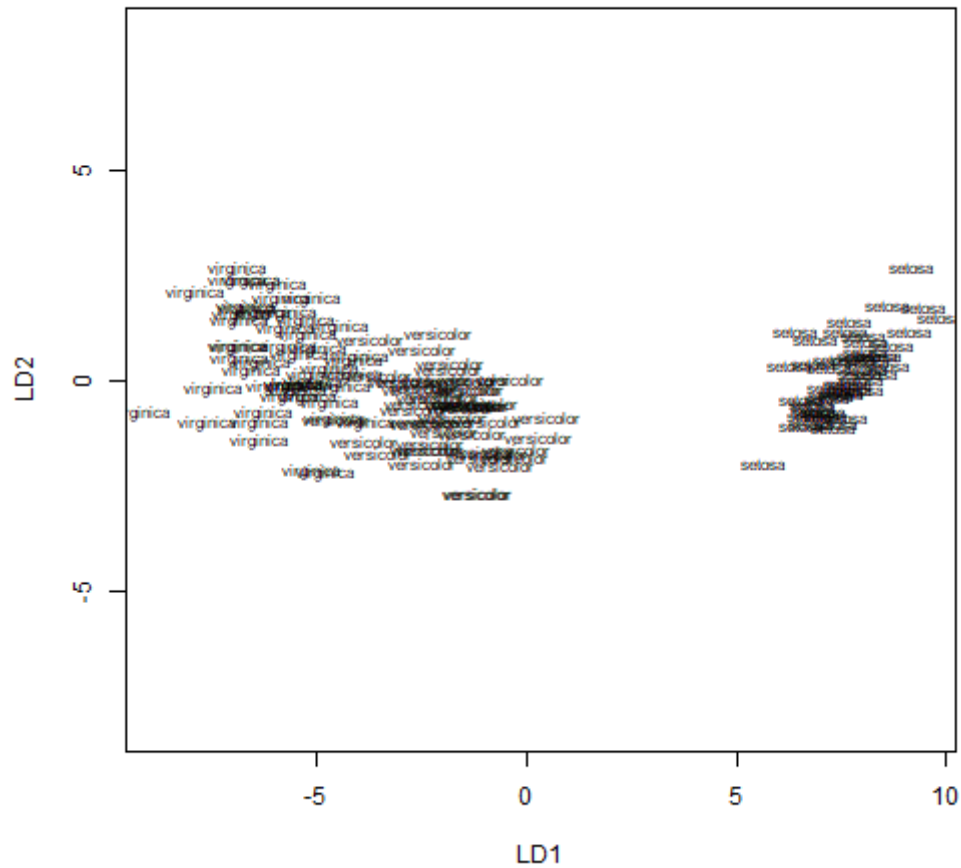
```
plot(ld2, dimen = 1) #第一判别函数变换后的直方图
```



```
plot(ld2, type="density",dimen = 1)#第一判别函数变换后的密度的
```



```
plot(ld2, dimen = 2)
```



```
(y1=predict(ld2, newdata=data.frame(Sepal.Length=5.9, Sepal.Width=
```

```
## $class
## [1] setosa
## Levels: setosa versicolor virginica
##
## $posterior
##      setosa      versicolor      virginica
## 1         1 8.026883e-15 3.046389e-32
##
## $x
##      LD1
## 1 6.332088
```

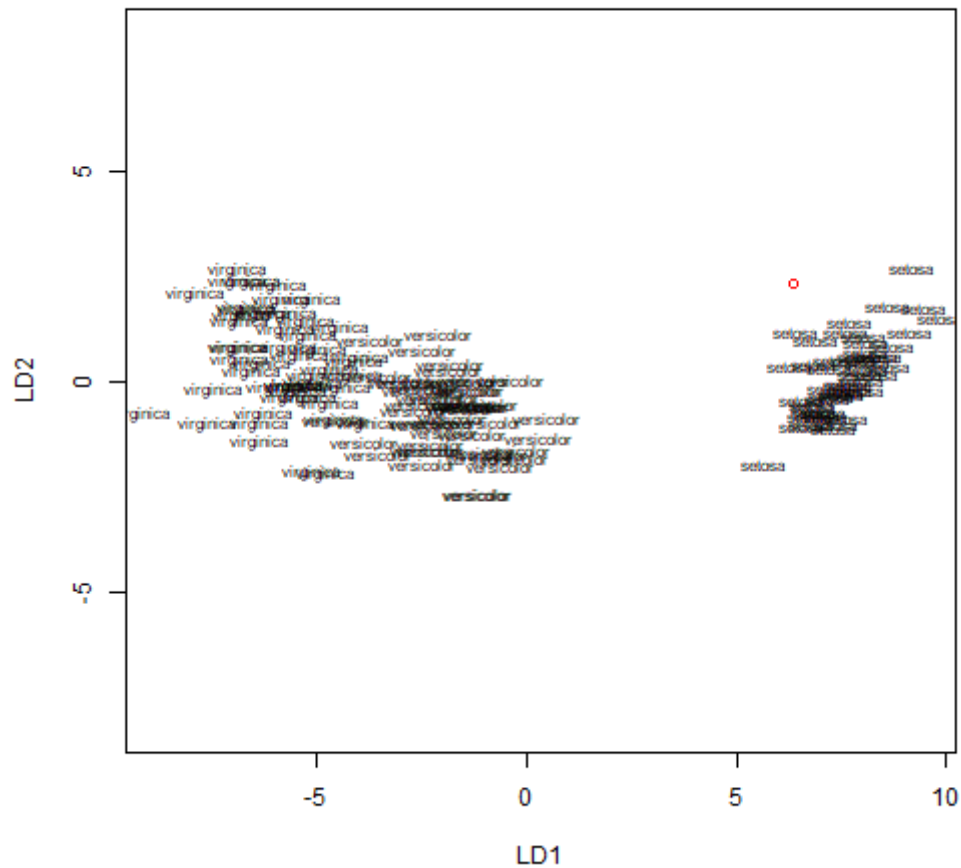
```
y1$x#新样本的线性变换后的值
```

```
##      LD1
## 1 6.332088
```

```
(y2=predict(ld2, newdata=data.frame(Sepal.Length=5.9,Sepal.Width=
```

```
## $class
## [1] setosa
## Levels: setosa versicolor virginica
##
## $posterior
##      setosa      versicolor      virginica
## 1          1 7.003845e-16 5.468919e-32
##
## $x
##           LD1           LD2
## 1 6.332088 2.329871
```

```
plot(ld2, dimen = 2)  
points(y2$x, col="red")
```





# 贝叶斯判别

- 贝叶斯判别

```
library(MASS)
(ld3<-lda(G~x1+x2, data=w3, prior=c(0.3,0.7)))
```

```
## call:
## lda(G ~ x1 + x2, data = w3, prior = c(0.3, 0.7))
##
## Prior probabilities of groups:
##      1      2
## 0.3 0.7
##
## Group means:
##      x1      x2
## 1  0.92  2.10
## 2 -0.38  8.85
##
## Coefficients of linear discriminants:
##              LD1
## x1 -0.1035305
## x2  0.2247957
```

- 贝叶斯判别

```
p.ld3<-predict(ld3)
p.ld3$class
```

```
# [1] 2 2 1 1 1 2 1 2 1 2 2 2 2 2 1 2 2 2 2 2
# Levels: 1 2
```

```
table(w3$G, p.ld3$class) #混淆矩阵
```

```
#
#      1 2
#      1 5 5
#      2 1 9
```

```
predict(ld3, newdata=data.frame(x1=8.1, x2=2.0)) #对新的值进行预测
```

```
# $class
# [1] 1
# Levels: 1 2
#
# $posterior
#           1           2
# 1 0.8559816 0.1440184
#
# $x
#           LD1
# 1 -1.922201
```

- 贝叶斯判别

```
(lda4<-lda(Species~., data=iris, prior=c(0.3,0.3, 0.4)))
```

```
## Call:
## lda(Species ~ ., data = iris, prior = c(0.3, 0.3, 0.4))
##
## Prior probabilities of groups:
##      setosa versicolor  virginica
##      0.3      0.3      0.4
##
## Group means:
##      Sepal.Length Sepal.width Petal.Length Petal.width
## setosa      5.006      3.428      1.462      0.246
## versicolor  5.936      2.770      4.260      1.326
## virginica   6.588      2.974      5.552      2.026
##
## Coefficients of linear discriminants:
##      LD1      LD2
## Sepal.Length 0.8291406 0.03121085
## Sepal.width  1.5158620 2.17759551
## Petal.Length -2.1931422 -0.95075619
## Petal.width  -2.8346951 2.81499172
##
## Proportion of trace:
##      LD1      LD2
## 0.9914 0.0086
```

# 蟹蟹

本幻灯片由 R 包 **xaringan** 生成；

查克拉来自于 **remark.js**、**knitr**、以及 **R Markdown**。