判别分析

R 语言

徐平峰

长春工业大学

2021/3/30

二次判別

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 判別

```
library(MASS)
(qd2<-qda(Species~., data=iris))
```

```
## Call:
## qda(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
```

```
p.qd2<-predict(qd2)
table(iris$Species, p.qd2$class)
```

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

```
names(p.qd2)
```

```
## [1] "class"      "posterior"
```

```
head(p.qd2$class)
```

```
## [1] setosa setosa setosa setosa setosa setosa
## Levels: setosa versicolor virginica
```

```
head(p.qd2$posterior)
```

```
##   setosa    versicolor      virginica
## 1      1 4.918517e-26 2.981541e-41
## 2      1 7.655808e-19 1.311032e-34
## 3      1 1.552279e-21 3.380440e-36
## 4      1 8.300396e-19 8.541858e-32
## 5      1 3.365614e-27 2.010147e-41
## 6      1 1.472533e-26 1.271928e-40
```

```
newdata=data.frame(Sepal.Length=5.9,Sepal.Width=3.9,Petal.Length=
(y1=predict(qd2, newdata = newdata))#利用二次判别函数进行预测
```

```
## $class
## [1] setosa
## Levels: setosa versicolor virginica
##
## $posterior
##    setosa    versicolor    virginica
## 1       1 1.345826e-15 5.19788e-30
```

```
y1$class#新样本的判别类别
```

```
## [1] setosa
## Levels: setosa versicolor virginica
```

```
y1$posterior#新样本在各个类别内的后验概率
```

```
##    setosa    versicolor    virginica
## 1       1 1.345826e-15 5.19788e-30
```

贝叶斯二次判别

- 贝叶斯判别

```
library(MASS)
(qd3<-qda(Species~., data=iris, prior=c(0.3,0.3,0.4)))
```

```
## Call:
## qda(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
```

- 贝叶斯判别

```
p.qd3<-predict(qd3)
head(p.qd3$class)
```

```
# [1] setosa setosa setosa setosa setosa setosa
# Levels: setosa versicolor virginica
```

```
table(iris$Species, p.qd3$class)#混淆矩阵
```

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

```
newdata=data.frame(Sepal.Length=5.9,Sepal.Width=3.9,Petal.Length=
predict(qd3, newdata = newdata)#对新的值进行预测
```

```
# $class
# [1] setosa
# Levels: setosa versicolor virginica
#
# $posterior
#   setosa   versicolor    virginica
# 1      1 1.345826e-15 6.930507e-30
```

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。