

Análisis Discriminante cuadrático

Introducción

En el análisis discriminante cuadrático no se parte del supuesto de que los grupos tienen matrices de covarianzas iguales. Como con el análisis discriminante lineal, una observación se clasifica en el grupo que tiene la distancia al cuadrado más pequeña. Sin embargo, la distancia al cuadrado no se simplifica en una función lineal, de allí el nombre de análisis discriminante cuadrático.

A diferencia de la distancia lineal, la distancia cuadrática no es simétrica. En otras palabras, la función discriminante cuadrática del grupo i evaluada con la media del grupo j no es igual a la función discriminante cuadrática del grupo j evaluada con la media del grupo i . En los resultados, la distancia cuadrática se llama distancia al cuadrado generalizada. Si el determinante de la matriz de covarianzas del grupo de muestra es menor que uno, la distancia al cuadrado generalizada puede ser negativa [Minitab].

La función `qda_function` que proporcionaste anteriormente devuelve una lista con cuatro elementos: `media`, `covarianza`, `prior` y `log.prior`. Cada uno de estos elementos contiene información sobre las medias, las matrices de covarianza, las probabilidades a priori y el logaritmo natural de las probabilidades a priori para cada grupo en los datos, respectivamente.

- `media`: Este elemento de la lista es una lista que contiene las medias de cada variable predictora para cada grupo. Cada elemento de esta lista es un vector numérico que contiene las medias de cada variable predictora para un grupo específico.
- `covarianza`: Este elemento de la lista es una lista que contiene las matrices de covarianza de las variables predictoras para cada grupo. Cada elemento de esta lista es una matriz numérica que muestra cómo las variables predictoras están relacionadas entre sí dentro de un grupo específico.
- `prior`: Este elemento de la lista es un vector numérico que contiene las probabilidades a priori de cada grupo (es decir, la proporción de observaciones en cada grupo).
- `log.prior`: Este elemento de la lista es un vector numérico que contiene el logaritmo natural de las probabilidades a priori de cada grupo.

En resumen, la función `qda_function` devuelve una lista con cuatro elementos, cada uno de los cuales contiene información sobre diferentes aspectos de los datos. Esta información se puede utilizar para clasificar nuevas observaciones en uno de los grupos utilizando la regla de Bayes.

```
{r, eval=FALSE} devtools::install_github("bjesusvh/DBCA")
```

```
##Objetivos
```

- Desarrollar un paquete en R para poder realizar el analisis discriminatorio cuadratico, para poder obtener un resultado de calculos mas eficas y rapido.
- Ilustrar mediante un ejemplo practicando el uso del paquete
- Definir la función qda_function qda_function <- function(x, grouping) { # Calcular las medias por grupo media <- by(x, grouping, colMedia)

Calcular la matriz de covarianza por grupo

```
covariances <- by(x, grupo, cov)
```

Calcular el número de observaciones por grupo

```
n <- table(grupo)
```

Calcular las probabilidades a priori por grupo

```
prior <- n / sum(n)
```

Calcular el logaritmo de las probabilidades a priori por grupo

```
log.prior <- log(prior)
```

Crear una lista para almacenar los resultados

```
resultados <- list(media = media, covarianza = covarianza, prior = prior, log.prior = log.prior)
```

Asignar la clase 'qda' al resultado

```
class(resultados) <- 'qda'
```

Devolver el resultado

```
return(resultados) }
```

Crear un marco de datos con variables predictoras

```
x <- data.frame( x1 = c(1, 2, 3, 4, 5, 6), x2 = c(6, 5, 4, 3, 2, 1), x3 = c(1, 2, 1, 2, 1, 2) )
```

Crear un vector que indica a qué grupo pertenece cada observación

```
grouping <- c('A', 'A', 'B', 'B', 'C', 'C')
```

Ejecutar la función qda_function con los datos creados

```
qda_result <- qda_function(x, grouping)
```

Ver el resultado

```
qda_resultados
```

Ejemplo de aplicación del paquete

ejemplo

Crear un marco de datos con variables predictoras

```
x <- data.frame( x1 = c(1, 2, 3, 4, 5, 6), x2 = c(6, 5, 4, 3, 2, 1), x3 = c(1, 2, 1, 2, 1, 2) )
```

Crear un vector que indica a qué grupo pertenece cada observación

```
grupo <- c('A', 'A', 'B', 'B', 'C', 'C')
```

```
resultados <- qda(x, grupo) print(resultados)
```

Conclusiones

Este paquete que se realizó, realmente cumple con los objetivos, ya que este programa te da resultados sobre análisis discriminante cuadrático, siendo así un programa fácil y eficaz de usar.