

# Análisis Multivariante no Lineal

Para el uso de técnicas multivariantes no lineales vamos a instalar la librería **homals** (De Leew, J. and Patrick, M. (2009) Gifi Methods for Optimal Scaling in R: The Package homals. Journal of Statistical Software, 31(4), 1-20).

```
install.packages("homals", repos = "http://cran.r-project.org")
```

La cargamos:

```
library(homals)
```

El paquete **homals** contiene una única función que nos permite realizar las tres técnicas de análisis multivariante no lineal vistas hasta ahora.

## Análisis de Correspondencias Múltiples (HOMALS)

Aunque vamos a utilizar la función **homals**, existen otras opciones en diferentes librerías:

- **MCA** (FactoMineR)
- **mca** (MASS)
- **dudi.acm** (ade4)
- **mjca** (ca)

Vamos a ver un ejemplo con la función **homals** sobre los datos **senate** que contienen los votos de 100 senadores de EEUU sobre veinte temas seleccionados para la Acción Democrática. Los argumentos necesarios para poder ejecutarla son:

- **data**. Matriz de datos.
- **ndim**. Número de dimensiones a retener.
- **level**. Nivel de cuantificación. Los valores posibles son “nominal”, “ordinal”, “numerical” y “polynomial”.
- **active**. Índices de las variables que son activas para el cálculo.
- ...

La primera variable contiene la afiliación de los senadores así que no se activa.

```
data(senate)
reshomals <- homals(senate, ndim=3, active=c(FALSE, rep(TRUE, 20)))
summary(reshomals)
```

```
##
## Number of dimensions: 3
## Number of iterations: 3
##
## -----
##
## Variable: Party
## Loadings:
##          D1          D2          D3
## 1 -0.2144 -0.025  0.0168
## 2  0.0014 -0.041 -0.0435
```

```

##
## Category centroids:
##      D1      D2      D3
## (D)  0.0213  0.0023 -0.0018
## (I) -0.0169  0.0388  0.0444
## (R) -0.0216 -0.0033  0.0009
##
## Category quantifications (scores):
##      D1      D2      D3
## (D)  0.0213  0.0023 -0.0019
## (I) -0.0169  0.0388  0.0444
## (R) -0.0216 -0.0033  0.0009
##
## Lower rank quantifications (rank = 2):
##      1      2
## (D) -0.0994  0.0043
## (I)  0.0724 -0.9909
## (R)  0.1011  0.0188
##
## -----
##
## Variable: V1
## Loadings:
##      D1      D2      D3
## 1 0.1957 0.026 -0.0337
##
## Category centroids:
##      D1      D2      D3
## - -0.0168 -0.0023  0.0029
## +  0.0229  0.0029 -0.0040
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0168 -0.0022  0.0029
## +  0.0229  0.0030 -0.0039
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.0857
## +  0.1170
##
## -----
##
## Variable: V2
## Loadings:
##      D1      D2      D3
## 1 0.205 0.0162 -0.0415
##
## Category centroids:
##      D1      D2      D3
## - -0.0183 -0.0015  0.0037
## +  0.0230  0.0017 -0.0047
##
## Category quantifications (scores):

```

```

##          D1          D2          D3
## - -0.0183 -0.0014  0.0037
## +  0.0230  0.0018 -0.0047
##
## Lower rank quantifications (rank = 1):
##          1
## - -0.0891
## +  0.1123
##
## -----
##
## Variable: V3
## Loadings:
##          D1          D2          D3
## 1 0.2152 0.0149 -0.0264
##
## Category centroids:
##          D1          D2          D3
## - -0.0208 -0.0015  0.0025
## +  0.0223  0.0014 -0.0028
##
## Category quantifications (scores):
##          D1          D2          D3
## - -0.0208 -0.0014  0.0026
## +  0.0223  0.0015 -0.0027
##
## Lower rank quantifications (rank = 1):
##          1
## - -0.0966
## +  0.1036
##
## -----
##
## Variable: V4
## Loadings:
##          D1          D2          D3
## 1 0.1634 0.0368 0.0799
##
## Category centroids:
##          D1          D2          D3
## - -0.0245 -0.0056 -0.0119
## +  0.0108  0.0024  0.0053
##
## Category quantifications (scores):
##          D1          D2          D3
## - -0.0245 -0.0055 -0.0120
## +  0.0108  0.0024  0.0053
##
## Lower rank quantifications (rank = 1):
##          1
## - -0.1498
## +  0.0664
##
## -----

```

```

##
## Variable: V5
## Loadings:
##      D1      D2      D3
## 1 0.1774 0.0591 0.0418
##
## Category centroids:
##      D1      D2      D3
## - -0.0218 -0.0073 -0.0051
## +  0.0144  0.0047  0.0034
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0218 -0.0073 -0.0052
## +  0.0144  0.0048  0.0034
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.1232
## +  0.0810
##
## -----
##
## Variable: V6
## Loadings:
##      D1      D2      D3
## 1 0.2173 0.0258 -0.0205
##
## Category centroids:
##      D1      D2      D3
## - -0.0218 -0.0027  0.0020
## +  0.0216  0.0025 -0.0021
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0218 -0.0026  0.0021
## +  0.0216  0.0026 -0.0020
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.1005
## +  0.0995
##
## -----
##
## Variable: V7
## Loadings:
##      D1      D2      D3
## 1 0.2181 0.017 -0.0269
##
## Category centroids:
##      D1      D2      D3
## - -0.0215 -0.0018  0.0026
## +  0.0221  0.0016 -0.0028

```

```

##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0215 -0.0017  0.0027
## +  0.0221  0.0017 -0.0027
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.0985
## +  0.1015
##
## -----
##
## Variable: V8
## Loadings:
##      D1      D2      D3
## 1 0.2097 0.0232 -0.0396
##
## Category centroids:
##      D1      D2      D3
## - -0.0195 -0.0023  0.0036
## +  0.0226  0.0024 -0.0043
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0195 -0.0022  0.0037
## +  0.0226  0.0025 -0.0043
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.0928
## +  0.1078
##
## -----
##
## Variable: V9
## Loadings:
##      D1      D2      D3
## 1 0.2173 0.0258 -0.0205
##
## Category centroids:
##      D1      D2      D3
## - -0.0218 -0.0027  0.0020
## +  0.0216  0.0025 -0.0021
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0218 -0.0026  0.0021
## +  0.0216  0.0026 -0.0020
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.1005
## +  0.0995

```

```

##
## -----
##
## Variable: V10
## Loadings:
##      D1      D2      D3
## 1 0.2084 0.0399 0.0165
##
## Category centroids:
##      D1      D2      D3
## - -0.0227 -0.0044 -0.0018
## + 0.0191 0.0036 0.0015
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0227 -0.0044 -0.0018
## + 0.0191 0.0037 0.0015
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.1089
## + 0.0917
##
## -----
##
## Variable: V11
## Loadings:
##      D1      D2      D3
## 1 0.2168 0.0163 -0.0151
##
## Category centroids:
##      D1      D2      D3
## - -0.0218 -0.0017 0.0015
## + 0.0216 0.0015 -0.0015
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0218 -0.0016 0.0015
## + 0.0216 0.0016 -0.0015
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.1005
## + 0.0995
##
## -----
##
## Variable: V12
## Loadings:
##      D1      D2      D3
## 1 0.1857 0.0414 -0.025
##
## Category centroids:
##      D1      D2      D3

```

```

## - -0.0163 -0.0036 0.0018
## + 0.0214 0.0048 -0.0032
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0163 -0.0036 0.0022
## + 0.0214 0.0048 -0.0029
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.0875
## + 0.1153
##
## -----
##
## Variable: V13
## Loadings:
##      D1      D2      D3
## 1 0.2088 0.0303 -0.0441
##
## Category centroids:
##      D1      D2      D3
## - -0.0195 -0.0030 0.0040
## + 0.0227 0.0031 -0.0049
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0195 -0.0028 0.0041
## + 0.0227 0.0033 -0.0048
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.0933
## + 0.1085
##
## -----
##
## Variable: V14
## Loadings:
##      D1      D2      D3
## 1 0.1661 0.0662 -0.0509
##
## Category centroids:
##      D1      D2      D3
## - -0.0140 -0.0052 0.0040
## + 0.0209 0.0087 -0.0067
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0138 -0.0055 0.0042
## + 0.0211 0.0084 -0.0065
##
## Lower rank quantifications (rank = 1):
##      1

```

```

## - -0.0832
## + 0.1272
##
## -----
##
## Variable: V15
## Loadings:
##      D1      D2      D3
## 1 0.1631 0.0414 0.0719
##
## Category centroids:
##      D1      D2      D3
## - -0.0199 -0.0048 -0.0086
## + 0.0134 0.0036 0.0061
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0198 -0.0050 -0.0087
## + 0.0135 0.0034 0.0059
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.1213
## + 0.0827
##
## -----
##
## Variable: V16
## Loadings:
##      D1      D2      D3
## 1 0.1704 0.0651 0.0221
##
## Category centroids:
##      D1      D2      D3
## - -0.0168 -0.0065 -0.0022
## + 0.0173 0.0065 0.0022
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0168 -0.0064 -0.0022
## + 0.0173 0.0066 0.0022
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.0988
## + 0.1013
##
## -----
##
## Variable: V17
## Loadings:
##      D1      D2      D3
## 1 0.1854 0.0414 0.0406
##

```



```

## Category centroids:
##      D1      D2      D3
## - -0.0232 -0.0045 -0.0052
## +  0.0159  0.0041  0.0034
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0231 -0.0052 -0.0051
## +  0.0160  0.0036  0.0035
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.1247
## +  0.0864
##
## -----
##
## Variable: V18
## Loadings:
##      D1      D2      D3
## 1 0.1626 0.0337 0.0338
##
## Category centroids:
##      D1      D2      D3
## - -0.0233 -0.0053 -0.0047
## +  0.0121  0.0020  0.0027
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0234 -0.0049 -0.0049
## +  0.0120  0.0025  0.0025
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.1438
## +  0.0738
##
## -----
##
## Variable: V19
## Loadings:
##      D1      D2      D3
## 1 0.008 0.208 -0.0744
##
## Category centroids:
##      D1      D2      D3
## - -1e-03 -0.0222  0.0079
## +  7e-04  0.0195 -0.0070
##
## Category quantifications (scores):
##      D1      D2      D3
## - -9e-04 -0.0222  0.0079
## +  8e-04  0.0195 -0.0070
##

```

```

## Lower rank quantifications (rank = 1):
##      1
## - -0.1066
## +  0.0937
##
## -----
##
## Variable: V20
## Loadings:
##      D1      D2      D3
## 1 0.1299 0.0773 0.1321
##
## Category centroids:
##      D1      D2      D3
## - -0.0234 -0.0140 -0.0233
## +  0.0070  0.0041  0.0077
##
## Category quantifications (scores):
##      D1      D2      D3
## - -0.0232 -0.0138 -0.0236
## +  0.0073  0.0043  0.0074
##
## Lower rank quantifications (rank = 1):
##      1
## - -0.1787
## +  0.0558

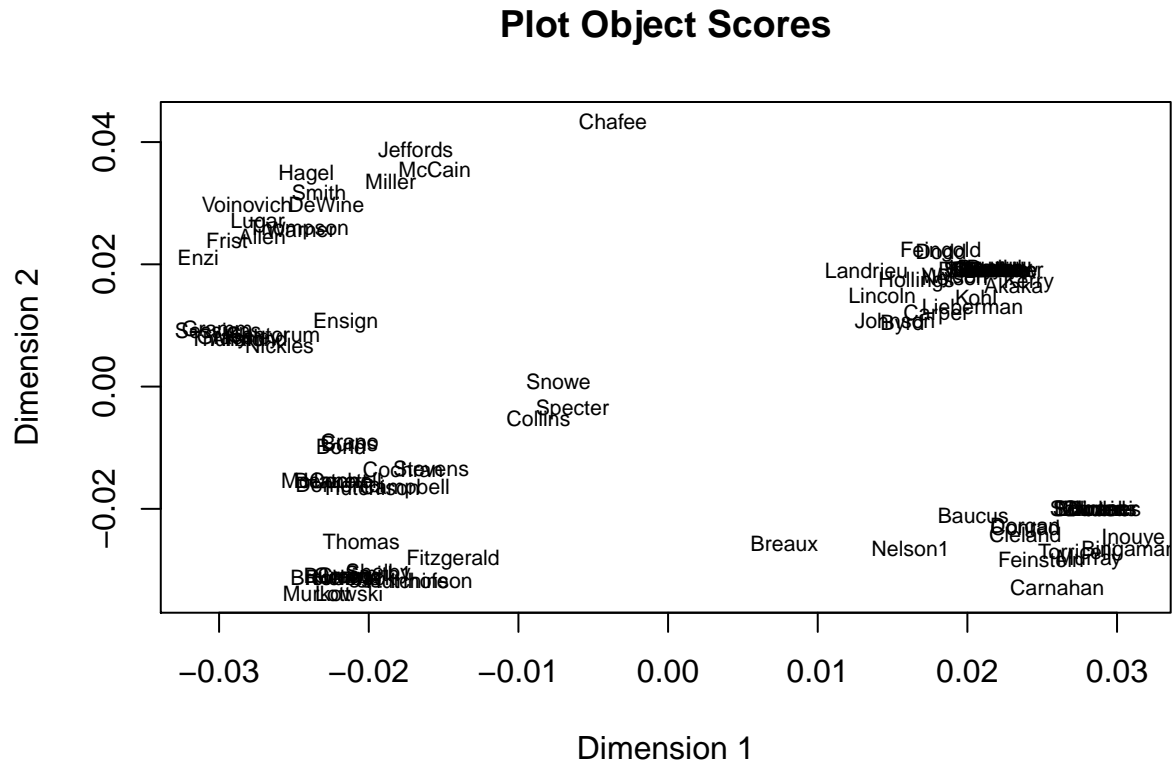
```

Esta librería ofrece una gran variedad de gráficos a representar a partir de los resultados obtenidos.

- **objplot**. Dibuja los scores de los objetos (filas) en dos y tres dimensiones.
- **catplot**. Dibuja la cuantificación de las categorías en rango reducido de cada variable por separado. Está disponible en tres dimensiones.
- **vorplot**. Dibuja un gráfico de las categorías con regiones de Voronoi.
- **jointplot**. Dibuja un gráfico conjunto de los scores de los objetos y de las cuantificaciones de las categorías. También está disponible en tres dimensiones.
- **graphplot**. La opción anterior pero con conexiones entre objetos y las correspondientes categorías de respuesta.
- **hullplot**. Para una variable específica se representan los scores de los objetos y se dibuja una línea poligonal cerrada respecto de cada categoría.
- **labplot**. Se representan los scores de los objetos para cada variable etiquetados con la categoría correspondiente. También se puede visualizar en tres dimensiones.
- **spanplot**. Dibuja los scores de los objetos para cada variable y los conecta entre sí mediante el camino más corto dentro de cada categoría.
- **starplot**. Representa los scores de los objetos en dos y tres dimensiones conectándolos con la categoría centroide.
- **lossplot**. Dibuja las cuantificaciones de las categorías en rango reducido frente a cada una de las variables por separado.
- **prjplot**. Se proyectan los scores de los objetos sobre la línea ortogonal a la dirección de la cuantificación.

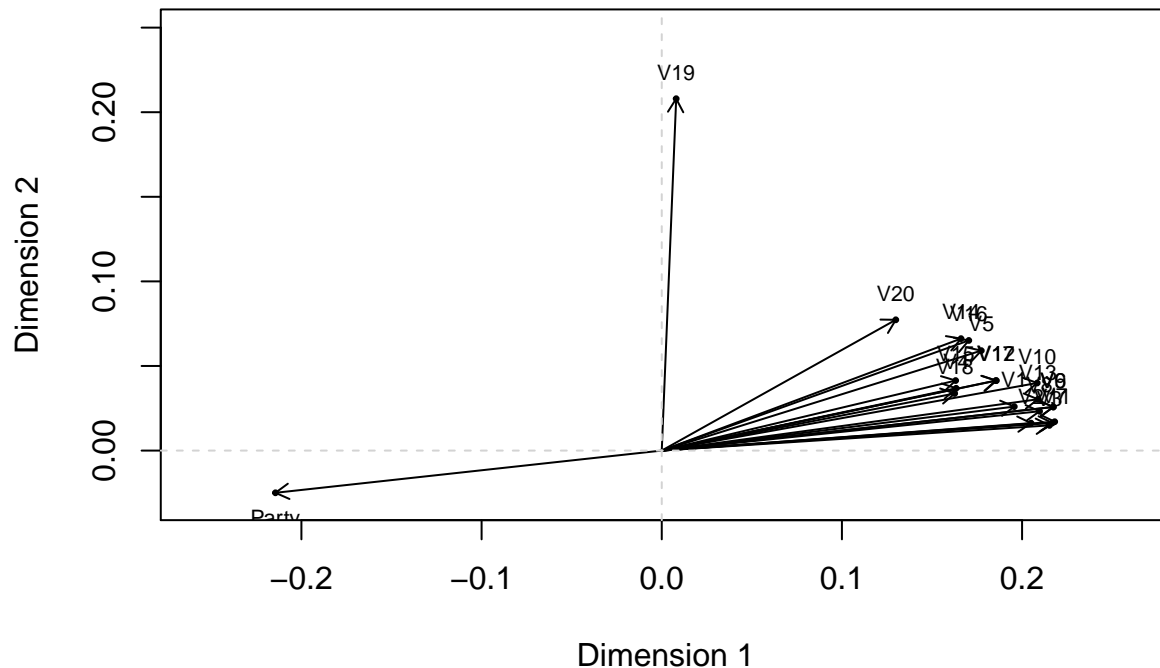
- `vecplot`. Se proyectan los scores de los objetos sobre la dirección de la cuantificación. Estos dos últimos sólo se pueden representar si la variable es de rango 1.
- `trfplot`. Representa las variables originales frente a las transformadas en la solución.
- `loadplot`. Se dibujan las cargas y se unen con el origen de coordenadas.
- `screeplot`. Scree plot de los valores propios.
- `dmplot`. Se dibujan las medidas discriminantes para cada variable.

```
plot(reshomals, plot.type="objplot")
```



```
plot(reshomals, plot.type="loadplot")
```

## Loadings plot



Es posible predecir la afiliación de cada senador asignando a cada uno la categoría más cercana. A partir de esta asignación y la verdadera pertenencia a los distintos partidos se puede obtener la matriz de confusión.

```
p.res <- predict(reshomals)
p.res$cl.table$Party
```

```
##      pre
## obs  (D) (I) (R)
##  (D)  49  1  0
##  (I)   0  1  0
##  (R)   0  9  40
```

Aproximadamente el 90 % de los senadores son bien clasificados.

## Componentes Principales Categóricas (PRINCALS)

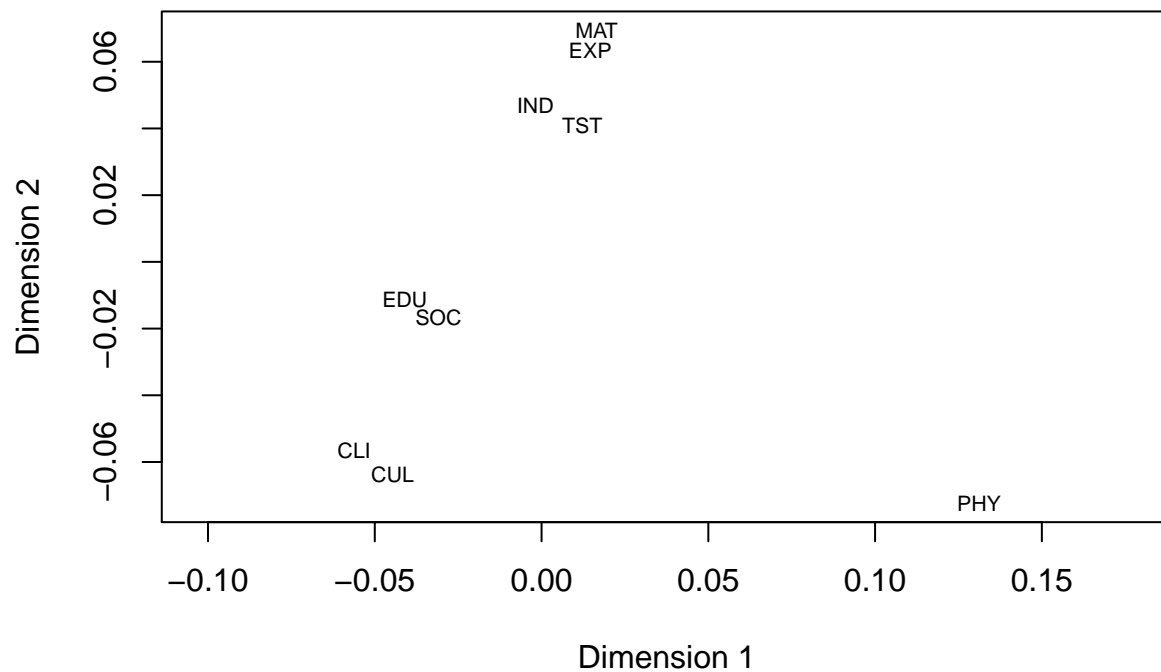
Vamos a ver un ejemplo de Componentes Principales Categóricas con la función `homals` de la librería `homals` sobre los datos `roskam`. Los datos aportan información de 39 psicólogos sobre el lugar que ocupa su trabajo en nueve áreas de investigación y docencia.

```
data(roskam)
res <- homals(roskam, rank = 1, level = "ordinal")

## Loss function increases in iteration 1

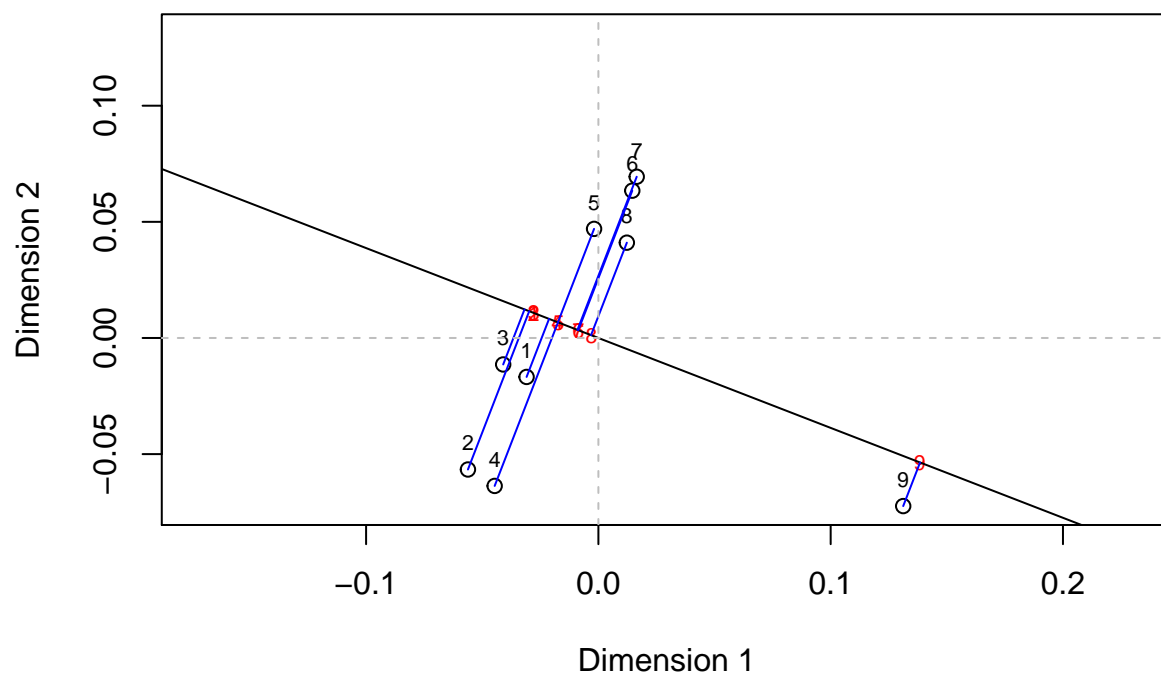
## Warning in homals(roskam, rank = 1, level = "ordinal"):
#res2<-princals(roskam, level = "ordinal")
plot(res, plot.type = "objplot", asp = 1)
```

## Plot Object Scores



```
plot(res, plot.type = "vecplot", var.subset = 2, asp = 1, main="Psicólogo nº 2")
```

## Psicólogo nº 2



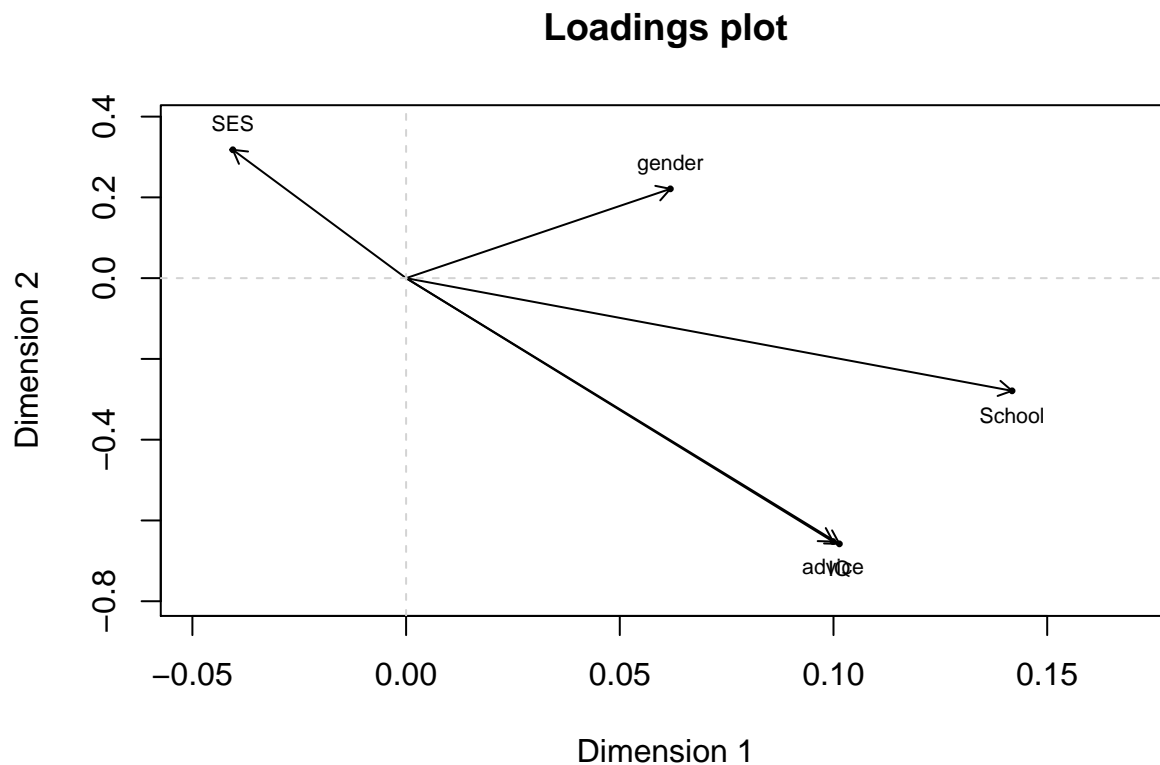
## Correlación Canónica No Lineal (OVERALS)

Cuando tenemos dos o más conjuntos de variables categóricas podemos utilizar el análisis de Correlación Canónica no Lineal. Utilizando el argumento `sets` de la función `homals` es posible emular la situación de un análisis de Correlación Canónica no Lineal. Vamos a realizar un ejemplo con los datos `galo` que contienen información de 1290 estudiantes sobre los que se han medido cinco variables.

```
data(galo)
res <- homals(galo, active = c(rep(TRUE, 4), FALSE), sets = list(c(1,2),c(3,4),5))

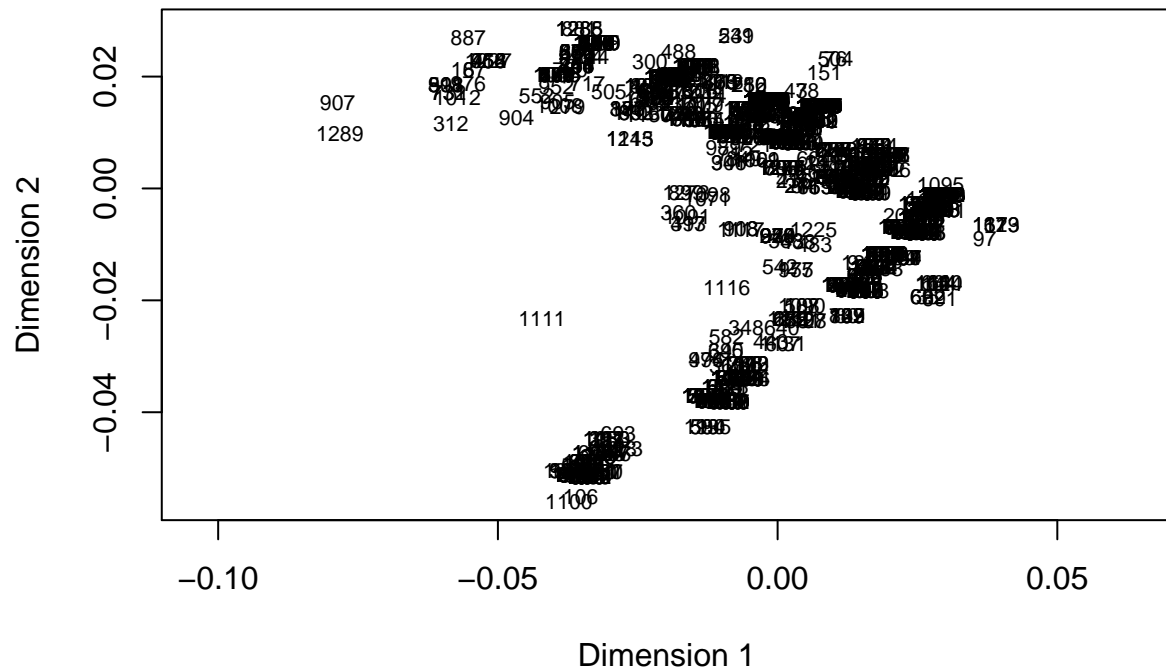
## Loss function increases in iteration 6

## Warning in homals(galo, active = c(rep(TRUE, 4), FALSE), sets = list(c(1, :
plot(res, plot.type="loadplot", asp=1)
```



```
plot(res, plot.type="objplot", asp=1)
```

## Plot Object Scores



```
plot(res, plot.type="dmpplot", asp=1)
```

## Discrimination Measures

