Training in ade4 in R - Module I: Basic methods

Multiple correspondence analysis

Stéphane Dray

2023-12-03

Data

We will analyze the `doubs` data set (see `?doubs`)

```
library(ade4)
library(adegraphics)
data(doubs)
names(doubs)
```

```
## [1] "env"     "fish"    "xy"      "species"
```

```
names(doubs$env)
```

```
##  [1] "dfs" "alt" "slo" "flo" "pH"  "har" "pho" "nit" "amm" "oxy" "bdo"
```

Tranformation into categorical variables

```
fenv <- apply(doubs$env, 2, cut, breaks = 4, labels = 1:4)
fenv <- as.data.frame(fenv, stringsAsFactors = TRUE)
```

Multiple Correspondence Analysis

- Perform MCA
- Display the barplot of eigenvalues

Graphical representation of MCA results

- Plot the results using the `plot` function

PCA scores on the geographical map

- Draw maps of PCA scores on the first two axes
- Interpret the maps to describe the environmental structure of the river

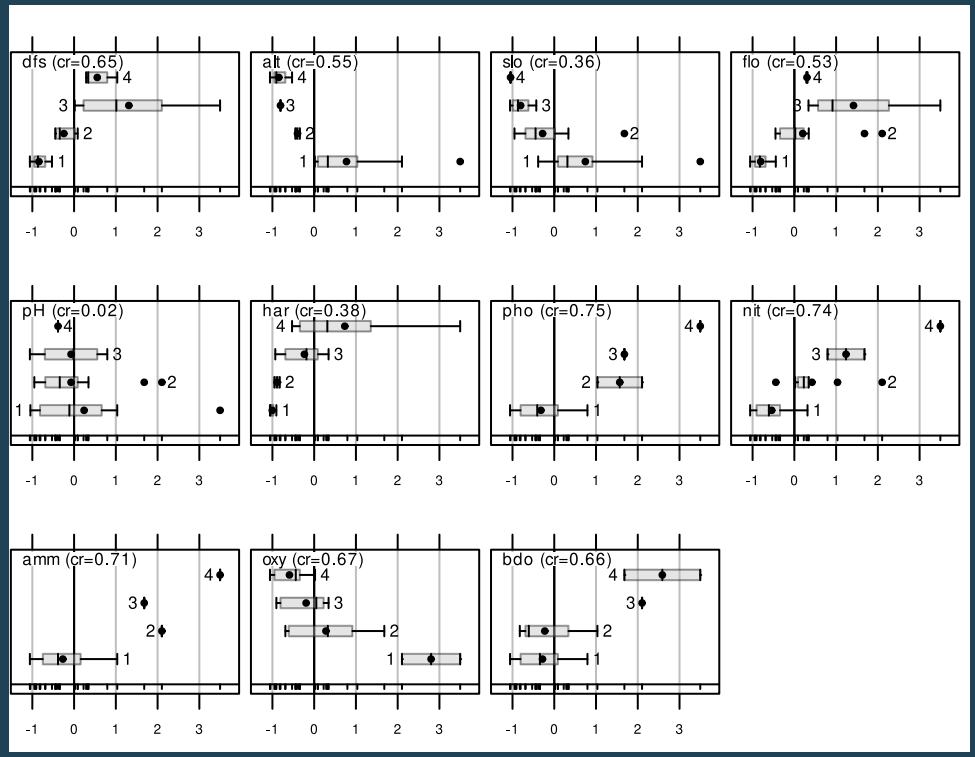
A look to variables

- Which variables are the most discriminated by the first axes

A look to variables

The generic function `score` provides an optimal representation of the maximized criteria

```
score(acm1, type = "boxplot")
```



Hill-Smith analysis

- Build a table mixing quantitative and categorical variables

```
menv <- cbind(fenv[, 1:6], doubs$env[, 7:11])
```

- Perform Hill-Smith analysis

Graphical representation

```
score(hs1, type = "boxplot")
```