Training in ade4 in R - Module I: Basic methods

Correspondence analysis

Stéphane Dray

2023-12-06

Data

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

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

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

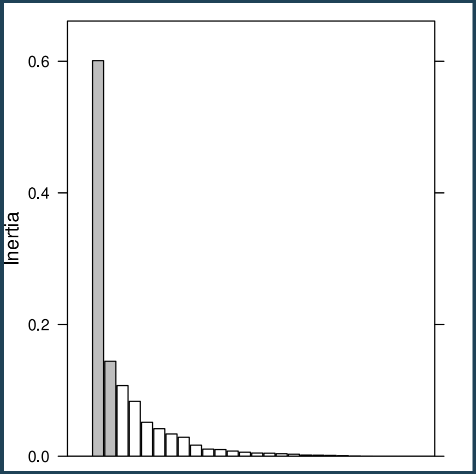
```
names(doubs$fish)
```

```
##  [1] "Cogo" "Satr" "Phph" "Neba" "Thth" "Teso" "Chna" "Chto" "Lele" "Lece"
## [12] "Spbi" "Gogo" "Eslu" "Pefl" "Rham" "Legi" "Scer" "Cyca" "Titi" "Abbr"
## [23] "Acce" "Ruru" "Blbj" "Alal" "Anan"
```

Correspondence Analysis

- Perform CA
- Display the barplot of eigenvalues

```r
coa1 <- dudi.coa(doubs$fish, scannf = FALSE)
screeplot(coa1, main = " ", xlab = " ")
```



Inertia statistics

- Compute the percentage of variation explained by the first COA axes

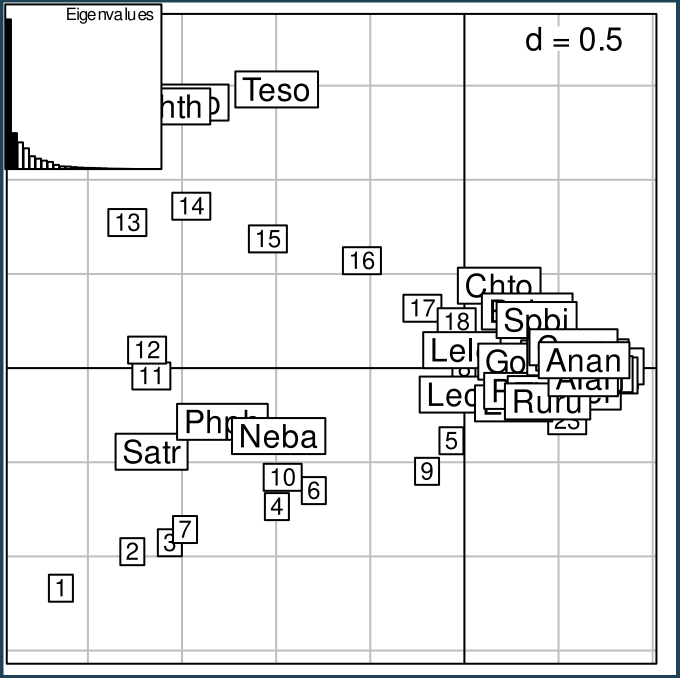
```r
summary(coa1)
```

```
## Class: coa dudi
## Call: dudi.coa(df = doubs$fish, scannf = FALSE)
##
## Total inertia: 1.167
##
## Eigenvalues:
##     Ax1     Ax2     Ax3     Ax4     Ax5
## 0.60099 0.14437 0.10729 0.08337 0.05158
##
## Projected inertia (%):
##     Ax1     Ax2     Ax3     Ax4     Ax5
##  51.503  12.372   9.195   7.145   4.420
##
## Cumulative projected inertia (%):
##     Ax1   Ax1:2   Ax1:3   Ax1:4   Ax1:5
##   51.50   63.87   73.07   80.21   84.63
##
## (Only 5 dimensions (out of 26) are shown)
```

Graphical representation of CA results

- Plot the results using the `biplot` function

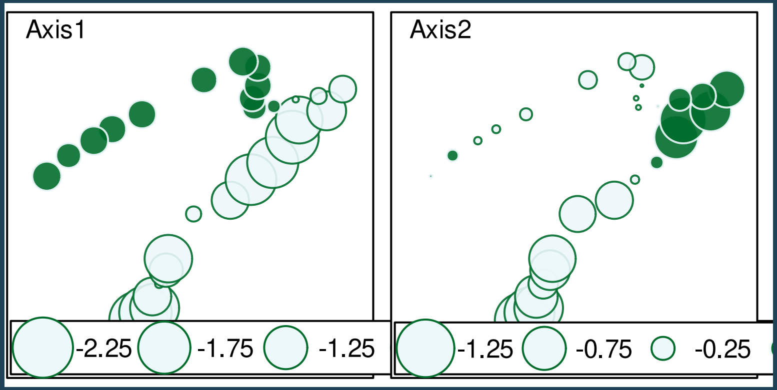
```
biplot(coa1)
```



CA scores on the geographical map

- Draw maps of CA scores on the first two axes
- Interpret the maps to describe how the fish communities vary along the river

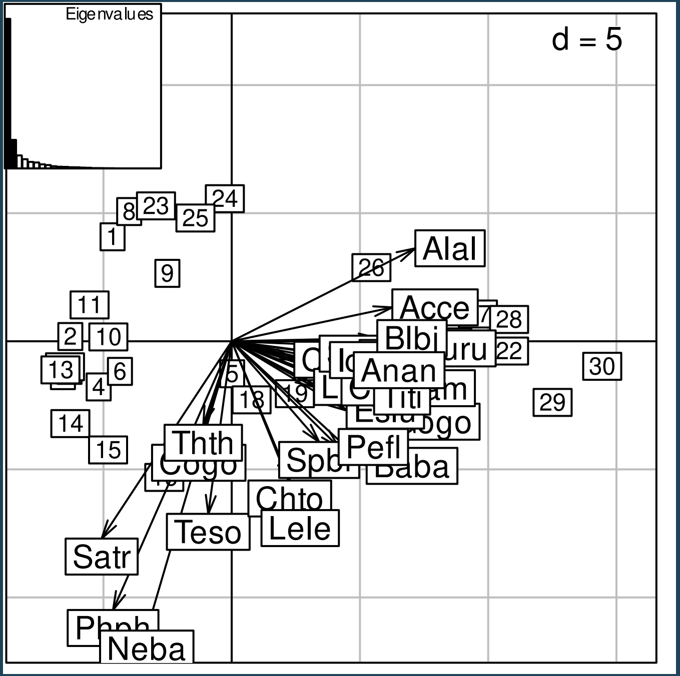
```
mypal <- colorRampPalette(c("#EDF8FB", "#006D2C"))
s.value(doubs$xy, coa1$li[, 1:2], pgrid.draw = FALSE,
    porigin.draw = FALSE, method = "size", symbol = "circle",
    col = mypal(2), ppoints.cex = 1)
```



Principal Component Analysis

PCA can also be applied on the abundance table. Perform PCA on `doubs$fish` table. Should we scale or not?

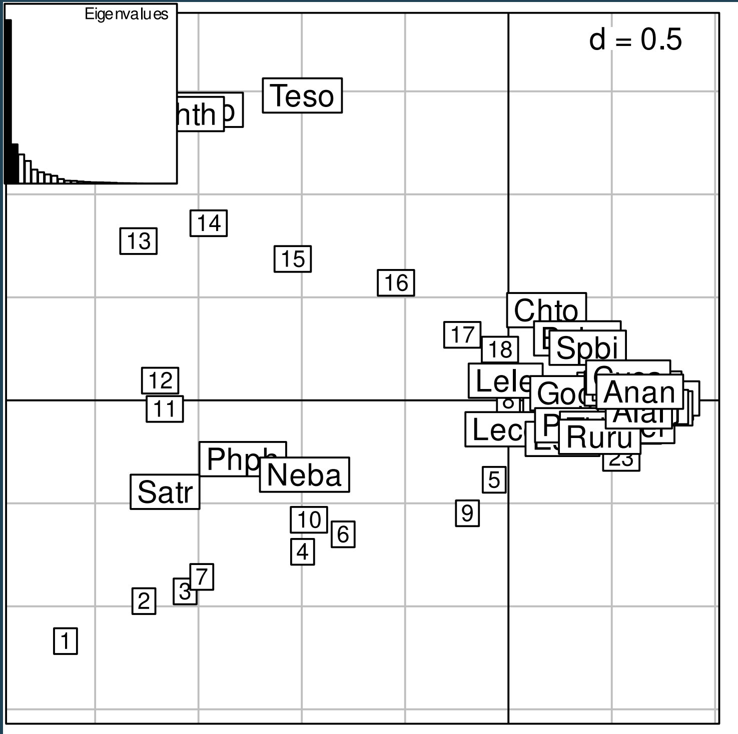
```
pca1 <- dudi.pca(doubs$fish, scale = FALSE, scannf = FALSE)
biplot(pca1)
```



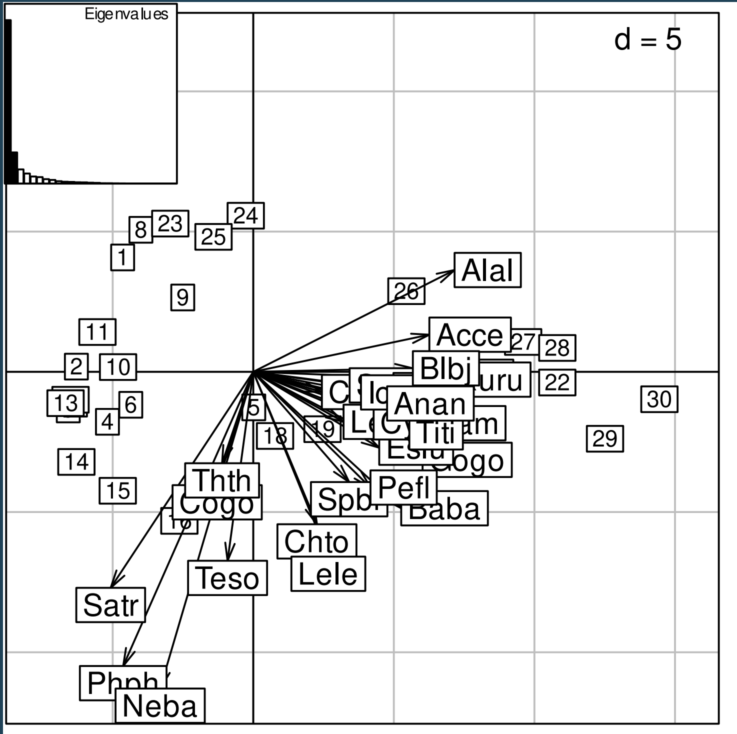
PCA vs CA

Compare the biplots of CA and PCA

```
biplot(coa1)
```



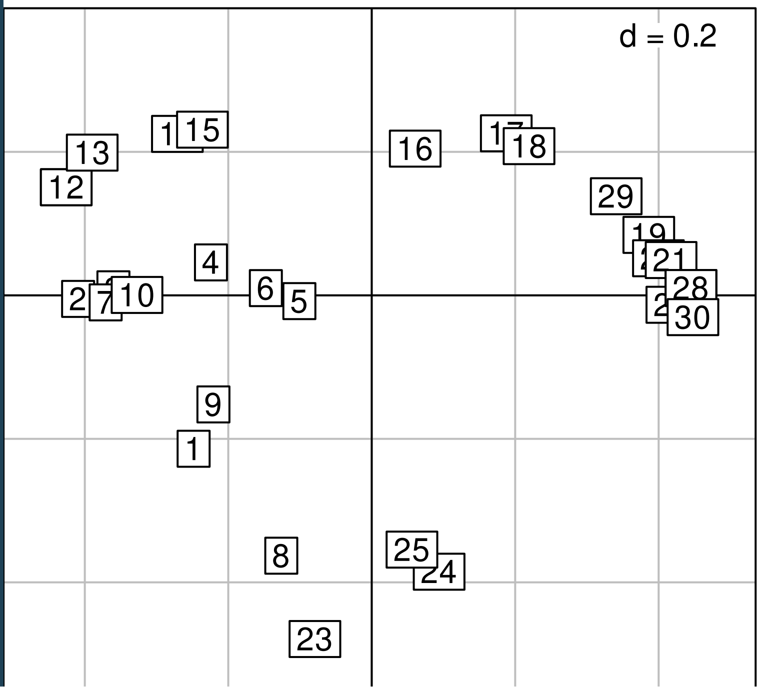

```
biplot(pca1)
```



Principal Coordinates Analysis

- Compute Jaccard distances between sites
- Perform principal coordinates analysis and display the ordination of sites

```
dis <- dist.binary(doubs$fish, method = 1)
pco1 <- dudi.pco(dis, scannf = FALSE)
s.label(pco1$li)
```



PCA and PCoA

- Compare the results of PCA and those of PCoA applied on Euclidean distance (function `dist`)

```
dis2 <- dist(doubs$fish)
pco2 <- dudi.pco(dis2, scannf = FALSE)
```

```
head(pca1$eig, 2)
```

```
## [1] 42.746273  8.158339
```

```
head(pca1$li)
```

```
##          Axis1      Axis2
## 1 -4.64572246  4.0779692
## 2 -6.29163098  0.1850865
## 3 -6.58040507 -1.3415646
## 4 -5.18637301 -1.7826102
## 5  0.01763648 -1.2682324
## 6 -4.35691748 -1.1761457
```

```
head(pco2$eig, 2)
```

```
## [1] 42.746273  8.158339
```

```
head(pco2$li)
```

```
##             A1         A2
## 1 -4.64572246  4.0779692
## 2 -6.29163098  0.1850865
## 3 -6.58040507 -1.3415646
## 4 -5.18637301 -1.7826102
## 5  0.01763648 -1.2682324
## 6 -4.35691748 -1.1761457
```