Example S2 - db-RDA

Maxime Hervé, Florence Nicolè and Kim-Anh Lê Cao

16/10/2017

Contents

Packages to load

```
library(ade4)           # Function needed: is.euclid
library(vegan)          # Function needed: dbrda
library(RVAideMemoire) # Functions needed: MVA.synt, MVA.anova, pairwise.factorfit,
                        #                   MVA.plot
```

Data loading

The dataset is composed of 32 rows (32 samples, 8 per treatment/date) and 6560 columns (the date, the treatment and the abundance of 6558 ions (2507 negative, 4051 positive)).

```
tab <- read.table("Example S2.txt",header=TRUE)
```

Pre-treatment

Chemical data are fourth-root transformed (not log since zeroes occur):

```
Chemistry <- tab[,3:6560]^(1/4)
```

The data are then autoscaled:

```
Chemistry.scaled <- scale(Chemistry)
```

Then the distance matrix is computed, based on the Euclidian distance:

```
mat.dist.Chemistry <- dist(Chemistry.scaled)
```

Analysis

We first need to check wether the distance matrix has Euclidian properties:

```
is.euclid(mat.dist.Chemistry)
```

```
[1] TRUE
```

As it is the case, we perform the db-RDA. Since the interaction between date and treatment is relevant, it is included in the model:

```
dbRDA <- dbrda(mat.dist.Chemistry~Date*Treatment,data=tab)
```

How much total variance does the experimental design explain?

```
MVA.synt(dbRDA)
```

```
Criterion: total variance (%)
               Proportion Cumulative
   Constrained      22.08      22.08
 Unconstrained      77.92     100.00

Criterion: constrained variance (%)
 Axis Proportion Cumulative
    1      65.63      65.63
    2      23.15      88.78
    3      11.22     100.00

Criterion: unconstrained variance (%)
 Axis Proportion Cumulative
    1      17.32      17.32
    2       8.60      25.93
    3       5.23      31.16
    4       4.56      35.72
    5       4.36      40.08
```

The experimental design (the two factors and their interaction), taken together, explain 22 % of the total variance in the chemical data.

We test for the significance of this explained variance, *i.e.* that it is higher than under the null hypothesis of no effect of the experimental desig:

```
anova(dbRDA)
```

```
Permutation test for dbrda under reduced model
Permutation: free
Number of permutations: 999

Model: dbrda(formula = mat.dist.Chemistry ~ Date * Treatment, data = tab)
         Df Variance     F Pr(>F)
Model     3   1447.7 2.644  0.001 ***
Residual 28   5110.3
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

There is a significant global effect of the experimental design, then we test for the individual effects of the factors (and interaction):

```
MVA.anova(dbRDA)
```

```
Permutation test for dbrda under reduced model
Type II tests
Permutation: free
Number of permutations: 999

Model: dbrda(formula = mat.dist.Chemistry ~ Date * Treatment, data = tab)
               Df Variance      F Pr(>F)
Date            1    547.1 2.9974  0.001 ***
Treatment       1    521.0 2.8545  0.001 ***
Date:Treatment  1    379.7 2.0802  0.004 **
Residual       28   5110.3
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The interaction is significant, indicating that the effect of the treament is not the same depending on the date. See here for an explanation of what interactions mean.

We perform pairwise comparisons using r^2-based permutation tests to study the interaction:

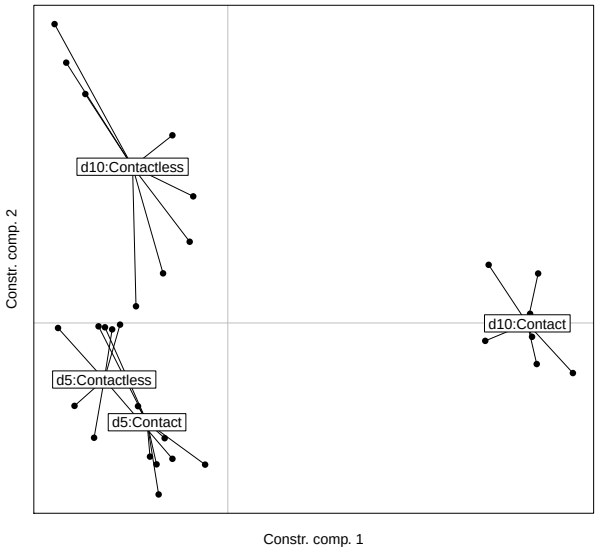
```
pairwise.factorfit(dbRDA,tab$Date:tab$Treatment)
```

```
    Pairwise comparisons using factor fitting to an ordination

data:  dbRDA by tab$Date:tab$Treatment
999 permutations

                d10:Contact d10:Contactless d5:Contact
d10:Contactless 0.0012      -               -
d5:Contact      0.0012      0.0012          -
d5:Contactless  0.0012      0.0012          0.1200

P value adjustment method: fdr
```

The two groups at day 5 are similar, while the two groups at day 10 are each significantly different from all other groups.

We draw the score plot of the constrained PCA to illustrate this interaction:

```
MVA.plot(dbRDA,fac=tab$Date:tab$Treatment,drawextaxes=FALSE)
```



From the MVA.synt() function used above, we know that the first component of this score plot explains 66 % of the constrained variance, the second component explains 23 %.

Information on the current R session

```
sessionInfo()
```

```
R version 3.4.0 (2017-04-21)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 7 x64 (build 7601) Service Pack 1

Matrix products: default

locale:
[1] LC_COLLATE=French_France.1252  LC_CTYPE=French_France.1252
[3] LC_MONETARY=French_France.1252 LC_NUMERIC=C
[5] LC_TIME=French_France.1252

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
[1] RVAideMemoire_0.9-68 vegan_2.4-4          lattice_0.20-35
[4] permute_0.9-4        ade4_1.7-8           knitr_1.17

loaded via a namespace (and not attached):
 [1] Rcpp_0.12.13       cluster_2.0.6      magrittr_1.5
 [4] splines_3.4.0      MASS_7.3-47        minqa_1.2.4
 [7] stringr_1.2.0      car_2.1-5          tools_3.4.0
[10] pbkrtest_0.4-7     nnet_7.3-12        parallel_3.4.0
[13] grid_3.4.0         nlme_3.1-131       mgcv_1.8-22
[16] quantreg_5.33      MatrixModels_0.4-1 htmltools_0.3.6
[19] lme4_1.1-14        yaml_2.1.14        rprojroot_1.2
```

```
[22] digest_0.6.12      Matrix_1.2-11      nloptr_1.0.4
[25] evaluate_0.10.1    rmarkdown_1.6      stringi_1.1.5
[28] compiler_3.4.0     backports_1.1.1    SparseM_1.77
```