Table of Contents

Ordination analysis

Monte Carlo permutation test (constrained ordination)

Theory R functions **Examples** Exercise 🔒

Example 1: Is the variance explained by environmental variables significant?

This example directly follows the Example 1 in RDA & tb-RDA section and Example 1 in Explained variance section, consider checking them first. We used Vltava river valley dataset, and two field-measured environmental variables, soil pH and soil depth (pH and SOILDPT), to explain variance in species composition. We found that they can explain 8.9% of overall variance; when compared to the variance which can be maximally explained by two explanatory variables (21.7% explained by two tb-PCA axes, see here) this sounds not bad (it is more than 40% of variance we can maximally explain in this dataset with two variables). But is the result significant? By significant, I mean: is the variance considerably higher than the variance explained (in average) by two random variables not related to species composition? This is the task for Monte Carlo permutation test (check Theory part to see how it works).

First, get the data and calculate tb-RDA on them (this is esentially repeating beginning of Example 1 in the section Explained variance):

```r
vltava.spe <- read.delim
('https://raw.githubusercontent.com/zdealveindy/anadat-r/master/data/vltava-
spe.txt', row.names = 1)
vltava.env <- read.delim
('https://raw.githubusercontent.com/zdealveindy/anadat-r/master/data/vltava-
env.txt')
spe <- vltava.spe
env <- vltava.env[, c('pH', 'SOILDPT')]
library (vegan)
spe.log <- log1p (spe)
spe.hell <- decostand (spe.log, 'hell')
tbRDA <- rda (spe.hell ~ pH + SOILDPT, data = env)
R2.obs <- RsquareAdj (tbRDA)$r.squared
R2.obs
```

```
[1] 0.08868879
```

In the next step, calculate variance explained by randomized env. variables

```r
env.rand <- env[sample (1:97),]  # the function "sample" will reshuffle the
rows with environmental variabels
tbRDA.rand <- rda (spe.hell ~ pH + SOILDPT, data = env.rand)
RsquareAdj (tbRDA.rand)$r.squared
```

This value represents the variance explained by two random explanatory variables. My result was 0.01854349, but this value will change in each run. We need to do enough repetitions (permutations) to get an idea about the distribution of this values (null model). We can use the "for" loop for it, or, as in this case, function "replicate", with two arguments: n = number of replicates, and expr = expression to be replicated (if more lines of script are involved, this expression needs to be enclosed in curly brackets {}):

```
n.perm <- 99  # set the number of permutations
R2.rand <- replicate (n = n.perm, expr = {
  env.rand <- env[sample (1:97),]
  tbRDA.rand <- rda (spe.hell ~ pH + SOILDPT, data = env.rand)
  RsquareAdj (tbRDA.rand)$r.squared
  })
<code>
The vector ''R2.rand'' contains 99 values of variance explained by random
variables. In the next step, we will merge them with the observed R2
(''R2.obs''), since this
```