

Table of Contents

Constrained ordination	1
Monte Carlo permutation test	1
R functions	1
Examples of use	1
Significance of RDA on species composition of plants in Carpathian wetlands	1
Exercise 1	4

Constrained ordination

Monte Carlo permutation test

The function of permutation test is to evaluate whether the variation explained by the model (variables) is higher than would be variation explained by the same number of randomly generated variables.

R functions

- `anova.cca` (library `vegan`) - tests the significance of the variation in species composition explained by explanatory variables, using Monte Carlo permutation test. It can test the significance of the global model (default), only the first constrained axis (adding argument `first = TRUE`), or significance of all axes (`by = "axis"`), significance of individual terms (explanatory variables) added sequentially (`by = "terms"`) or significance of variation explained by individual explanatory variables after removing variation of all other variables in the model (`by = "margin"`).

Examples of use

Significance of RDA on species composition of plants in Carpathian wetlands

This example continues from the previous section with [example on using RDA](#).

After analysing RDA using all environmental variables as explanatory, the next question is whether the global model is significant:

```
anova ( rda.vasc )
```

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

Model: rda(formula = vasc.hell ~ Ca + Mg + Fe + K + Na + Si + S04 + P04 +
N03 + NH3 + Cl + Corg + pH + conduct + slope, data = chem)
      Df Variance      F Pr(>F)
Model   15  0.21277  2.1599  0.001 ***
Residual 54  0.35464
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Alternatively, we may be interested in significance of only the first constrained axis. This is achieved by adding argument `first = TRUE`:

```
anova (rda.vasc, first = TRUE)
```

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

Model: rda(formula = vasc.hell ~ Ca + Mg + Fe + K + Na + Si + S04 + P04 +
N03 + NH3 + Cl + Corg + pH + conduct + slope, data = chem)
      Df Variance      F Pr(>F)
RDA1   1  0.09793 14.912  0.001 ***
Residual 54  0.35464
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Or, we may calculate significance of each constrained axis independently. This can be done by adding argument by = "axis". Note that since there is 14 variables and hence 14 axes, the calculation takes rather long; you may speed it up using parallel calculation, if your computer has more than one core (use argument parallel with number of available cores, 4 in case of my computer):

```
anova (rda.vasc, by = 'axis', parallel = 4)
```

```
Permutation test for rda under reduced model
Marginal tests for axes
Permutation: free
Number of permutations: 999

Model: rda(formula = vasc.hell ~ Ca + Mg + Fe + K + Na + Si + S04 + P04 +
N03 + NH3 + Cl + Corg + pH + conduct + slope, data = chem)
      Df Variance      F Pr(>F)
RDA1   1  0.09793 14.9120  0.001 ***
RDA2   1  0.02237  3.4070  0.001 ***
RDA3   1  0.01546  2.3547  0.001 ***
RDA4   1  0.01110  1.6905  0.009 **
RDA5   1  0.01061  1.6152  0.012 *
RDA6   1  0.00930  1.4155  0.050 *
RDA7   1  0.00840  1.2798  0.110
RDA8   1  0.00637  0.9703  0.498
RDA9   1  0.00593  0.9023  0.644
RDA10  1  0.00582  0.8861  0.677
RDA11  1  0.00496  0.7549  0.877
RDA12  1  0.00423  0.6448  0.970
RDA13  1  0.00394  0.5994  0.992
RDA14  1  0.00331  0.5034  1.000
RDA15  1  0.00304  0.4626  1.000
Residual 54  0.35464
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Other testing option is to test each term (constraining variable) separately - this is done adding the argument by = "terms". It sequentially adds each variable one by one in order in which they enter

the model formula, and for each calculates partial explained variation and its significance with previous variables as covariables (i.e. for the first variable it is marginal variation explained by this variable and its significance, for the second variable it is partial variation explained by the second variable after removing the variation of the first variable as covariable, etc.). In the context of our data this option is not too useful, since variables in dataset are ordered without explicit meaning.

```
anova (rda.vasc, by = 'terms', parallel = 4)
```

```
Permutation test for rda under reduced model
Terms added sequentially (first to last)
Permutation: free
Number of permutations: 999
```

```
Model: rda(formula = vasc.hell ~ Ca + Mg + Fe + K + Na + Si + S04 + P04 +
N03 + NH3 + Cl + Corg + pH + conduct + slope, data = chem)
```

	Df	Variance	F	Pr(>F)	
Ca	1	0.07886	12.0079	0.001	***
Mg	1	0.01395	2.1242	0.009	**
Fe	1	0.00962	1.4643	0.082	.
K	1	0.00822	1.2511	0.155	
Na	1	0.01154	1.7577	0.031	*
Si	1	0.01387	2.1119	0.015	*
S04	1	0.00688	1.0476	0.348	
P04	1	0.00598	0.9111	0.569	
N03	1	0.00860	1.3102	0.124	
NH3	1	0.01239	1.8872	0.014	*
Cl	1	0.00601	0.9154	0.567	
Corg	1	0.00905	1.3778	0.100	.
pH	1	0.01151	1.7525	0.034	*
conduct	1	0.00959	1.4609	0.100	.
slope	1	0.00669	1.0186	0.387	
Residual	54	0.35464			

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The last testing option is with argument `by = "margin"`, testing variation explained by each explanatory variable with all the others used as covariables:

```
anova (rda.vasc, by = 'margin', parallel = 4)
```

```
Permutation test for rda under reduced model
Marginal effects of terms
Permutation: free
Number of permutations: 999
```

```
Model: rda(formula = vasc.hell ~ Ca + Mg + Fe + K + Na + Si + S04 + P04 +
N03 + NH3 + Cl + Corg + pH + conduct + slope, data = chem)
```

	Df	Variance	F	Pr(>F)	
Ca	1	0.01441	2.1947	0.008	**
Mg	1	0.00976	1.4857	0.075	.
Fe	1	0.00723	1.1006	0.263	

```

K      1  0.00690  1.0508  0.342
Na     1  0.00561  0.8539  0.631
Si     1  0.01221  1.8599  0.030 *
S04    1  0.00742  1.1306  0.260
P04    1  0.00470  0.7159  0.874
N03    1  0.00893  1.3600  0.114
NH3    1  0.01077  1.6404  0.054 .
Cl     1  0.00581  0.8849  0.628
Corg   1  0.00788  1.1997  0.185
pH     1  0.00858  1.3066  0.154
conduct 1  0.00922  1.4042  0.105
slope  1  0.00669  1.0186  0.373
Residual 54  0.35464
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Exercise 1

(This exercise is based on *Case study 4* from Šmilauer & Lepš 2014, and the coding was partly inspired by the [blog post of Gavin Simpson](#), analysing this data - I recommend it for checking on more details.)

Use vegetation data from an [experimental study on wet meadow](#), which tried to answer the question how is abundance and species composition of seedlings of herb species on wet meadow influenced by experimental removal of dominant clonal plant species (*Nardus stricta*), plant litter (remnants of plant biomass from the previous season in case that the meadow was not mown) and removal of litter+mosses (by raking). Before analysis, please get familiar with the data you will be using. Data are from completely randomized block design with four blocks, each with four plots (one per treatment, including control). The block design will need to be taken into consideration when testing the significance of explained variation using Monte Carlo permutation test.

1. Import datasets `seedlings.spe` and `seedlings.env` - the first containing species composition of seedlings per plot, the second contain assignment of the plot to the block and experimental treatment.
2. First, calculate DCA on compositional data (no transformation) to decide whether to use linear (RDA) or unimodal (CCA) constrained ordination method.
3. According to the result of previous analysis, use either RDA or CCA, where you will explain the species composition by treatment. Since the experiment design is a block design, you need to specify the variable `seedlings.env$blocks` as covariable in the model.
4. Draw the resulting ordination diagram.
5. Test the significance of the variation explained by treatment (use model based permutation test).
6. Test the significance of individual ordination axes. Which of them is the most important, and to which level of the treatment it corresponds?

```

1. read.delim;
2. check the length of the first axis in decorana results. If over 4, use CCA, if less then 3, use RDA, 3-4 -both are fine;
3. the formula will look like this: seedlings.spe ~ treatment + Condition (block),

```

```
data = seedlings.env;
```

4. ordiplot
5. anova without any other arguments; since the blocks are defined as covariable, he will test the significance of variation after removal of their variation;
6. anova with argument by = "axis".

Solution

From:

<https://anadat-r.davidzeleny.net/> - **Analysis of community ecology data in R**

Permanent link:

https://anadat-r.davidzeleny.net/doku.php/en:monte_carlo?rev=1467102545

Last update: **2017/10/11 20:36**