

Table of Contents

Constrained ordination	1
Variation partitioning	1
R functions	2
Examples of use	2
Use of varpart function (using vegetation data from Carpathian wetlands)	2
plot.varpart (library vegan)	3

Constrained ordination

Variation partitioning

Note: variation partitioning is sometimes also called **commonality analysis** in reference to the *common* (shared) fraction of variation (Kerlinger & Pedhazur 1973). It is also a synonym to **variance partitioning**¹⁾.

In case we have two or more explanatory variables, one may be interested in variation in species composition explained by each of them. If some of these explanatory variables are correlated, one must expect that variation explained by the first or the other variable cannot be separated - it will be shared.

Invalid Link



Venn's diagram, showing variation explained by two environmental variables (or two sets of environmental variables) and coded names of these fractions. Created by function `showvarpart` (2) from library `vegan`.

The way how to approach this problem is variation partitioning, when variation explained by each variable (or set of variables) independently is partitioned into variation attributable purely to given environmental variable, and shared variation attributable to two or more variables.

Variation can be partitioned by individual variables (e.g. variation explained by soil pH vs variation explained by soil Ca) or by groups of variables (e.g. soil variables vs climatic variables).

Results can be visualized using Venn's diagram (see figure on the right). Meaning of the fractions in Venn's diagram is the following²⁾:

- [a] - variation explained by variable 1 (conditional (or partial) effect of variable 1, i.e. variation this variable would explain if putting variable 2 as covariable);
- [c] - variation explained by variable 2;
- [b] - shared variation explained by both variables (cannot be decided to which of them should be attributed, and is a result of correlation between both variables);
- [a+b] - variation explained by variable 1 (independent simple (or marginal) effect of variable 1, i.e. variation this variable would explain if it is as the only explanatory variable in the model);
- [b+c] - variation explained by variable 2;
- [d] - unexplained variation.

If the variation is partitioned among groups with the same number of variables (e.g. two soil and two climatic variables), then variation explained by each group is comparable without adjustment. However, if groups contain different numbers of variables, variation explained by not adjusted R^2 is not comparable, since R^2 tends to increase with the number of explanatory variables. Here, use of adjusted R^2 is recommended.

The library `vegan` offers function `varpart`, which can partition variation among up to four variables (or groups of variables). Note that `varpart` is based on redundancy analysis (`rda`) and uses adjusted R^2 to express explained variation. The reason for using only `rda` is that in R, there is still no function available to calculate adjusted R^2 for unimodal ordination methods (like `cca`).

R functions

- **varpart** (library `vegan`) - variation partitioning (using linear constrained ordination - `rda`) among up to four matrices of environmental variables. First argument (`Y`) is dependent (usually species composition) matrix (but could be also only one variable - in that case `varpart` is conducting linear regression). Next arguments (up to four) are (groups of) explanatory variables. Uses either formula interface (with `~`) or matrices.
- **plot** (library `vegan`) - draws Venn's diagram with fractions of explained variation. In default setting it doesn't show negative values of explained variation (argument `cutoff = 0`). Consult `?plot.varpart` for more details.
- **showvarparts** (library `vegan`) - draws schema of Venn's diagram with codes of individual fractions.

Examples of use

Use of `varpart` function (using vegetation data from Carpathian wetlands)

Example: how much variation in species data (`vasc.hell`) explains variables `Mg` and `Ca` (from `chem`)?

```
# Carpathian wetlands - import data
vasc <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/vasc_plants.txt',
row.names = 1)
chem <- read.delim
('http://www.davidzeleny.net/anadat-r/data-download/chemistry.txt',
row.names = 1)

# transform data using Hellinger transformation
vasc.hell <- decostand (vasc, 'hell')

# upload library vegan if not yet done
library (vegan)

# apply function varpart
vp1 <- varpart (vasc.hell, ~ Mg, ~ Ca, data = chem)
```

In this function, the first is coming the species matrix, then explanatory matrices (or variables) - if using formula interface, each has to start with tilda (`~`). If these variables are part of matrix of explanatory variables, you need to specify the environmental matrix in argument `data =`.

Result is:

```
vp1
```

```
Partition of variation in RDA
```

```
Call: varpart(Y = vasc.hell, X = ~Mg, ~Ca, data = chem)
```

Explanatory tables:

X1: ~Mg

X2: ~Ca

No. of explanatory tables: 2

Total variation (SS): 39.151

Variance: 0.56741

No. of observations: 70

Partition table:

	Df	R.squared	Adj.R.squared	Testable
[a+b] = X1	1	0.11987	0.10693	TRUE
[b+c] = X2	1	0.13898	0.12632	TRUE
[a+b+c] = X1+X2	2	0.16357	0.13860	TRUE

Individual fractions

[a] = X1 X2	1		0.01228	TRUE
[b]	0		0.09465	FALSE
[c] = X2 X1	1		0.03167	TRUE
[d] = Residuals			0.86140	FALSE

Use function 'rda' to test significance of fractions of interest

Alternative way how to use the function is not using formula interface:

```
varpart (vasc.hell, chem$Mg, chem$Ca)
```

or, in case you use not-transformed data and you want them to be transformed (using `decostand` function):

```
varpart (vasc, chem$Mg, chem$Ca, transfo = 'hell')
```

(argument `transfo` is passed into function `decostand` together with species data).

plot.varpart (library vegan)

I can also plot directly so called Venn's diagram:

```
plot (vp1)
```

Note - the plotting function is called `plot.varpart`, but if I use generic plot function on object `vp1`

(which is of class "varpart"), I don't have to specify



1)

As to the distinction between *variance* and *variation*, Legendre & Legendre (2012) note: "The term variation, a less technical and looser term than variance, is used because one is partitioning the total sum of squared deviations of y from its mean (total SS). In variation partitioning, there is no need to divide the total SS of y by its degrees of freedom to obtain the variance s^2_y ." The first edition of the book *Multivariate Analysis of Ecological Data using CANOCO* (Lepš & Šmilauer 2003) was using the term *variance partitioning*, while in the second edition (Šmilauer & Lepš 2014), authors adopted the term *variation partitioning*, noting: "It was called **variance** partitioning in the original paper, but we prefer, together with Legendre & Legendre 2012, the more appropriate name referring to *variation*, as we include also unimodal ordination methods in our considerations."

2)

Note that in CANOCO 5, the coding of the fraction follows different logic - [c] is shared variation, and [a] or [b] are partial fractions; meaning of [d] remains the same (unexplained variation).

From:

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

Permanent link:

<https://anadat-r.davidzeleny.net/doku.php/en:varpart?rev=1429531798>

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