Explained variation in constrained ordination

Theory R functions **Examples**

Example 1: Is the explained variance high enough to be considered as interesting?

This is a direct continuation of Example 1 in tb-RDA section. We used Vltava river valley dataset to calculate tb-RDA (on log-transformed and then Hellinger transformed species composition data) with two measured variables (pH and SOILDPT) as explanatory. They explained 8.9% of the variance, and we may ask: if variables explain this amount of variance (which may seem rather low, less than 10%), is it enough to report this result?

To decide, we need to do two more things. One is to test whether the analysis is significant, meaning that this amount of variance is high enough to interpret it (this is the topic for Example 1 in Permutation test). The other thing is to compare the explained variation (R^2 = 8.9%) with the variation that would be explained by the two best variables, i.e., the variables that represent the strongest gradient along which species composition changes. As you can see in the theoretical part, we can create such a variable for our dataset by applying unconstrained ordination on our data and extracting the first few unconstrained ordination axes (i.e., sample scores along them). These axes represent the main directions along which species composition changes most. Then we can use these axes as explanatory in the constrained version of the same ordination (here tb-RDA) to see how much variance they explain. The number of axes I should take depends on the number of real variables I want to compare variances on (in this case, two).

Let's get data in and calculate tb-RDA on two explanatory variables (this is essentially repeating steps in Example 1 from tb-RDA:

```
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)  # species data are in percentage scale which is strongly
rightskewed, better to transform them
spe.hell <- decostand (spe.log, 'hell')  # we are planning to do tb-RDA, this is
Hellinger pre-transformation
tbRDA <- rda (spe.hell ~ pH + SOILDPT, data = env)  # calculate tb-RDA with two
explanatory variables
```

I can use the function `RsquareAd` from `vegan` to extract the explained variance directly from the ordination object:

```
R2.obs <- RsquareAdj (tbRDA)$r.squared
R2.obs
```

```
[1] 0.08868879
```

Result is 0.089 = 8.9%. Note that the function `RsquareAdj` returns a list with two components, `r.squared` and `adj.r.squared`, the first referring to the R^2 we are interested now, and the second to adjusted R^2, corrected for the number of samples and number of explanatory variables (we will use this later).

We need to calculate the unconstrained version of tb-RDA on the same data and extract the first two ordination axes. The unconstrained version of tb-RDA is tb-PCA (i.e., PCA calculated on pre-transform data, where the pre-transformation should be the same in both analyses):

```
tbPCA <- rda (spe.hell)
tbPCA
```

```
Call: rda(X = spe.hell)

              Inertia Rank
Total          0.7048
Unconstrained  0.7048   96
Inertia is variance

Eigenvalues for unconstrained axes:
    PC1     PC2     PC3     PC4     PC5     PC6     PC7     PC8
0.09197 0.06075 0.04684 0.03537 0.02650 0.02361 0.02093 0.02035
(Showed only 8 of all 96 unconstrained eigenvalues)
```

For comparison, let's also include the summary of the tb-RDA method:

```
tbRDA
```

```
Call: rda(formula = spe.hell ~ pH + SOILDPT, data = env)

              Inertia Proportion Rank
Total         0.70476    1.00000
Constrained   0.06250    0.08869    2
Unconstrained 0.64226    0.91131   94
Inertia is variance

Eigenvalues for constrained axes:
   RDA1    RDA2
0.04023 0.02227

Eigenvalues for unconstrained axes:
    PC1     PC2     PC3     PC4     PC5     PC6     PC7     PC8
0.07321 0.04857 0.04074 0.03144 0.02604 0.02152 0.01917 0.01715
(Showed only 8 of all 94 unconstrained eigenvalues)
```

You can see that total inertia (variance of the whole dataset) is identical in both analyses (0.70476). This is the variance we aim to explain.

Extract the two PCA axes from the `tbPCA` object and use them as explanatory variables in tbRDA on the same data to see how much variance they explain:

```
PCA12 <- scores (tbPCA, display = 'sites', choices = 1:2)
tbRDA_PCA12 <- rda (spe.hell ~ PCA12)
RsquareAdj (tbRDA_PCA12)$r.squared
```

```
[1] 0.2167075
```

The first two tb-PCA axes, if used as explanatory variables in the tb-RDA on the same dataset, explain 21.7% of the variance. In fact, we don't need to really do the whole procedure (extract tb-PCA axes and use them as explanatory in tb-RDA), we can simply check the variance of *n* axes in unconstrained ordination (*n* = number of explanatory variables) and compare it with the real variance explained by environmental variables (eigenvalue

of PCA1 = 0.09197, eigenvalue of PCA1 = 0.06075, total variance (inertia) = 0.70476, recalculated into percentage: (0.09197+0.06075)/0.70476 = 21.7%).

The comparison here is: 8.9% explained by real variables (pH and SOILDPT) vs 21.7%, which would be explained by the two best, not correlated variables (if we had them). We see that measured variables explain something over 40% of the variation they could (8.9/21.7 = 0.41), which is not bad. Remember important difference: PCA axes are (from definition) not correlated, while our real variables often will be (as in this case: `cor.test (~ pH + SOILDPT, data = env)` shows that Pearson's correlation coefficient between pH and SOILDPT is r = 0.273, and this correlation is significant at P = 0.0069).

The next step is to test whether the variation explained by our variables is significant - this is a topic for Example 1 in Permutation test section.

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

Permanent link:
https://anadat-r.davidzeleny.net/doku.php/en:expl_var_examples

Last update: **2026/04/06 19:06**