Numerical Ecology with R, second edition, 2018

This file provides corrections of small mistakes found in the printed edition (Errata), as well as improved explanations and updated, additional information (Addenda and updates).

Last update: 19 July 2020

Errata

Corrected words or other elements are in red.

P. 56, Table 3.1

Bottom right, last entry. One should read:

Binary variables:
Simple matching coefficient
`dist.binary(., method=2)`

Explanation: function **`dist.binary()`** accepts 0 and nonnegative values only (i.e., no standardized variables). All strictly positive values are converted to 1, and the coefficient `method = 2` computes the (square root of the one-complement of the) ratio between matching pairs (i.e. double zeros and double 1s) and the total number of variables.

P. 398, R code

Add the red code line below:

```
fish.jac.neigh <- diag(fish.jac[-1, ]) # Jaccard DJ index
absc <- c(2:7, 9:30) # Abscissa
```

Explanation: the line is present in the script but has been inadvertently dropped from the book version.

Addenda and updates

The following entries are proposed to follow changes in recent versions of R and packages, to improve explanations or to add some recent pieces of information.

P. 27, text and code at bottom of page

Replace **`vegtrans()`** by **`abundtrans()`**.

Explanation: in package **`labdsv{}`** the name of function **`vegtrans()`** has been replaced by **`abundtrans()`**. The function is the same.

P. 146, text

4.15.2 Noise clustering using the `vegclust()` function

Original paragraph in the published book:

A recent package called **`vegclust`**, developed by Miquel De Cáceres, provides a large [……] in the "Noise" cluster.

Replace by the following, more detailed paragraph:

A recent package called **`vegclust`**, developed by Miquel De Cáceres, provides a large range of options to perform non-hierarchical or hierarchical fuzzy clustering of community data under different models (De Cáceres et al. 2010). An interesting one is called "noise clustering" (Davé and Krishnapuram 1997). This method is an attempt to make fuzzy clustering more robust to outliers. Outliers are defined as follows: once "true cluster" centroids have been defined, capture into a fictitious "Noise" cluster the objects that lie farther than a distance δ from the "true cluster" centroids (De Cáceres et al. 2010). The choice of the value of δ is critical: too small a δ value results in an overly large number of outliers, i.e., a large membership in the "Noise" cluster. Note also that if a cluster has a larger intragroup dispersion than the others, this increases the likelihood that some of its legitimate members be considered as outliers.

P. 169, first block of text (below the **R** code)

Original text:

`envfit()` also proposes permutation tests to assess the significance of the R^2 of each explanatory variable regressed on the two axes of the biplot. But this is not, by far, the best way to test the effect of explanatory variables on a table of response variables. We will explore this topic in Chap. 6.

Replace by the expanded text below:

`envfit()` also proposes permutation tests to assess the significance of the R^2 of each explanatory variable regressed on the two axes of the biplot. The R^2 statistics (noted `r2` in the **`envfit()`** output) are produced for quantitative explanatory variables and for factors. They measure the fit of the data to the explanatory variables. With the default option `choices = c(1, 2)`, only the first two axes of the ordination are considered and the R^2

measures the fit of the data ordinated in two dimensions to each explanatory variable. If the calculation is made to involve all dimensions of a PCA ordination (this can be obtained by changing the values in argument `choices`), the R^2 statistic measures the fit of the full-dimensional data to the explanatory variables. If the ordination was produced by PCoA (Sect. 5.5) or NMDS (Sect. 5.6) of a dissimilarity matrix, the fit is between the response variables and the data transformed by the dissimilarity index used in the ordination. If the ordination is a PCA and the envfit analysis involves all PCA axes, the R^2 is identical to that produced by **`adonis2()`** (Chap. 6). Note, however that function **`envfit()`** has not been designed to replace this other function, which was designed for multivariate analysis of variance by RDA (Sect. 6.3.2.9); its role is to draw explanatory variables onto simple ordination plots.

P. 234, last text paragraph

Original text:

The three RDAs can be tested as usual, and fractions [a] and [c] can be computed and tested by means of partial RDA. Fraction [b], however, is not an adjusted component of variance and cannot be estimated and tested by regression methods. It has zero degree of freedom. […]

Expanded text:

The three RDAs can be tested as usual, and fractions [a] and [c] can be computed and tested by means of partial RDA. Fraction [b], however, is not an adjusted component of variance and cannot be estimated and tested by regression methods. It has zero degree of freedom. However, an elegant workaround has been devised by Bauman et al. (2018) in the special case of the shared space-environment relationship (i.e., the [b] fraction of two explanatory matrices, one of them modeling spatial structures using methods such as MEM variables [Sect. 7.4]), by means of special permutation procedures based on the spatial layout of the sampling units (torus translations and, in the case of irregular sampling, Moran spectral randomization, Wagner and Dray 2015).

Additional references:

Bauman, D., Vleminckx, J., Hardy, O. J., Drouet, T.: Testing and interpreting the shared space-environment fraction in variation partitioning analyses of ecological data. Oikos **128**, 274-285 (2018)

Wagner, H.H., Dray, S.: Generating spatially constrained null models for irregularly spaced data using Moran spectral randomization methods. Methods in Ecology and Evolution **6**, 1169-1178 (2015)

P. 234, R code

Depending on the version of package {spdep}, due to a reversal in the recording of the two dimensions of the grid (despite the absence of change in the names of the arguments), the line of code involving function **`cell2nb`** must be adapted:

```r
if(packageVersion("spdep") < 0.8 {
nb <- cell2nb(4, 10, "queen")
} else {
nb <- cell2nb(nrow = 10, ncol = 4, type = "queen")
# or, equivalent
# nb <- cell2nb(4, 10, "queen", legacy = TRUE)
}
```