Table of Contents

Overview of analyses

The key unit in the analysis of community ecology data sets is **community sample**, representing presence/absence or quantity (count, cover or biomass) of each species in each sample. The way how to handle such samples is via **ecological resemblance**, which can be quantified e.g. by **compositional dissimilarity** between two such community samples. Compositional dissimilarity describes the imaginary distance between two community samples in the **multidimensional compositional space** - two samples with exactly the same species composition will occupy exactly the same spot in this space, and their distance will increase with increasing dissimilarity in terms of species composition. Ecological resemblance and multidimensional compositional space are two main concepts which you need to understand before you turn into learning multivariate methods. Most ordination and classification methods are based on some of the compositional dissimilarity measures, although in some of them the measure itself is not explicitly mentioned.

Ordination is the way how to make an order in the set of community samples and the way how to reduce multidimensional information stored in community data into few imaginable, interpretable and printable dimensions. There are many ordination methods, with different fields (botany, zoology, microbiology) preferring different ones. Ordinations are focused on finding interpretable trends in data, represented by changes in species composition with possible underlying changes in environmental gradients. We may use it either simply for a description of community pattern (which is usually the purpose of unconstrained = indirect ordination) or to explain and test changes in species composition by some (e.g. environmental) variables (constrained = direct ordination).

Nummerical classification stands opposite to ordination - while ordination seeks the main gradients in the continuum of community samples, classification tries to separate this continuum into a finite number of groups (clusters), each containing more or less similar samples.

Apart to *sample × species* matrix of species composition (**L** matrix), and optionally also *sample × environmental variable* matrix of environmental variables or other types of sample attributes (**R** matrix), in some cases we have also the third matrix, which contains species attributes like species traits or species indicator values (*species × traits matrix* or **Q** matrix). There are several methods of **analysis of species attributes**, including three-matrix methods (like the *fourth corner* or RLQ analysis), or other ways of relating species and sample attributes (e.g. by calculating a community-weighted mean of species attributes for individual samples and relating them to environmental variables by regression).

Diversity analysis is in certain sense also analysis of species composition matrix, whose originally two-dimensional information (samples × species) is reduced into one-dimensional variables (like numbers of species in samples - alpha diversity, differences in species composition among samples - beta diversity, or number of all species in the matrix - gamma diversity). But diversity is not only about numbers of species, but also about their relative abundances - we will briefly review also the concepts of true diversity, evenness and their representation by different diversity indices. Diversity is also influenced by sampled area (species-area curve) and sampling effort (bias due to undersampling).
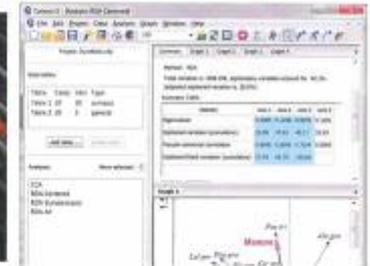
Software solutions

Figure 2: Are you not a vegan? Try something more palatable! An ad for CANOCO 5, 2015. I suspect the author is Petr Šmilauer. Just pity that CANOCO 5 after opening on your computer does not produce so lovely smell as those sausages in the advertisement. This website is focused on the use of R for all analyses. For most ordination methods we will use vegan package written by Jari Oksanen et al., with some functions also from ade4 by Stephen Dray & Anne B. Dufour and labdsv by Dave Roberts. You may want to, however, learn also some other software options. For most ordination methods, an excellent and user-friendly solution is CANOCO 5 (Windows only) developed by Cajo ter Braak & Petr Šmilauer, newly released update of the favorite CANOCO for Windows 4.5 (see ad on Fig. 2 and check here if you want to know how does CANOCO 5 compares with vegan from the view of CANOCO authors). Although I prefer to do ordination analyses in R, I sometimes opt for CANOCO 5 because of its convenience, and also because it does great ordination diagrams (R is still staying behind in this aspect). Cluster analysis (and also most of the ordination methods) are available in PC-ORD 6 software of Bruce McCune et al. (see here for comparison of CANOCO 5 to PC-ORD 6); I must admit I do not favor this software for somewhat not intuitive workflow, but it does offer rather wide range of multivariate methods (some more elaborated than in CANOCO 5, e.g. NMDS with more options and detail report). For analysis of diversity (e.g. diversity indices or rarefaction) consider using EstimateS 9 by Robert R. Colwell. These are just some examples with which I am familiar, you may find a range of other (free or paid) software.

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

Permanent link:
https://anadat-r.davidzeleny.net/doku.php/en:overview?rev=1489668686

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