5.1 Other ordination methods

Principal Co-ordinates Analysis

The two main weaknesses of PCA, identified at the end of Chapter 4, are its inflexibility of dissimilarity measure and its poor distance-preservation. The first problem is addressed in an important paper by Gower (1966) , describing an extension to PCA termed *Principal Co-ordinates Analysis* (PCO), also sometimes referred to as *classical scaling*. This allows a wider definition of distance than simple Euclidean distance in the species space (the basis of PCA), but was initially restricted to a specific class of resemblance measures for which the samples could be represented by points in some reconfigured high-dimensional (real) space, in which Euclidean distance between two points is just the (non-Euclidean) resemblance between those samples. Effectively none of the most useful biological resemblance coefficients fall into this class – the high-d space representing those dissimilarities has both real and imaginary axes – but it has become clearer in the intervening decades that much useful inference can still be performed in this complex space, e.g. McArdle & Anderson (2001) , Anderson (2001a) , Anderson (2001b) . (This is essentially the space in which the PERMANOVA+ add-on routines to the PRIMER software carry out their core analyses). PCO can thus be applied completely generally to any resemblance measure but the final step is again a projection onto a low-dimensional ordination space (e.g. a 2-dimensional plane), as in ordinary PCA. It follows that PCA is just a special case of PCO, when the original dissimilarity is just Euclidean distance, but note that PCO is still subject to the second criticism of PCA: its lack of emphasis on distance-preservation when the information is difficult to represent in a low number of dimensions.

Detrended Correspondence Analysis

Correspondence analyses are a class of ordination methods originally featuring strongly in French data-analysis literature (for an early review in English see Greenacre (1984)). Key papers in ecology are Hill (1973a) and Hill & Gauch (1980) , who introduced *detrended correspondence analysis* (DECORANA). The methods start from the data matrix, rather than a resemblance measure, so are rather inflexible in their definition of sample dissimilarity; in effect, multinomial assumptions generate an implicit dissimilarity measure of *chi-squared distance* (Chapter 16). Correspondence analysis (CA) has its genesis in a particular model of unimodal species response to underlying (unmeasured) environmental gradients. Description is outside the scope of this manual but good accounts of CA can be found in the works of Cajo ter Braak (e.g. in Jongman, ter Braak & Tongeren (1987)), who has contributed a great deal in this area, not least CCA, Canonical Correspondence Analysis (ter Braak (1986)).¶

The DECORANA version of CA, widely used in earlier decades, has a primary motivation of straightening out an *arch effect* in a CA ordination, which is expected on theoretical grounds if species abundances have unimodal (Guassian) responses along a single strong environmental

gradient. Where such models are not appropriate, it is unclear what artefacts the algorithms may introduce into the final picture. In the Hill & Gauch (1980) procedure, the detrending is essentially carried out by first splitting the ordination space into segments, stretching or shrinking the scale in each segment and then realigning the segments to remove wide-scale curvature. For some people, this is uncomfortably close to attacking the data with scissors and glue and, though the method is not as subjective as this would imply, some arbitrary decisions about where and how the segmentation and rescaling are defined are hidden from the user in the software code. Thus Pielou (1984) and others criticized DECORANA for its 'overzealous' manipulation of the data. It is also unfortunate that the multivariate methods which were historically applied in ecology were often either poorly suited to the data or were based on conceptually complex algorithms (e.g. DECORANA and TWINSPAN, Hill (1979a) and Hill (1979b)), erecting a communication barrier between data analyst and ecologist.

The ordination technique which is adopted in this manual's strategy, *non-metric MDS*, is itself a complex numerical algorithm but it will be argued that it is *conceptually* simple. It makes few (if any) model assumptions about the form of the data, and the link between the final picture and the user's original data is relatively transparent and easy to explain. Importantly, it addresses both the major criticisms of PCA made earlier: it has *great flexibility* both in the definition and conversion of dissimilarity to distance and its rationale *is* the preservation of these relationships in the low-dimensional ordination space.

¶ *A convenient way of carrying out CA-related routines is to use the excellent CANOCO package, ter Braak & Smilauer (2002) .*

Revision #9
Created 21 February 2022 11:19:26 by Arden
Updated 27 June 2023 23:46:32 by Arden