Winter 2013

Foote

5. Introduction to Ordination

I. General Features of Ordination (in contrast to clustering)

A. allow better representation of larger distances
B. objects arranged on continuous axes
C. no forced hierarchical structure
D. *shapes* of clusters can be seen
E. gradients can be seen

II. Polar Ordination

A. Basic Procedure

0. Let d_{ij} be the original measured distance between entities i and j.
1. Select two entities as poles.
2. Calculate coordinate along the axis formed from the two poles. For entity i ordinated along axis formed by entities 1 and 2:

$$D_{1i} = \frac{d_{12}^2 + d_{1i}^2 - d_{2i}^2}{2d_{12}},$$

   which implies that $D_{11} = 0$ and $D_{12} = d_{12}$.
3. Calculate residual distances (either with respect to this ordination axis or with respect to entire ordination, depending on method).
4. Select poles for second axis.
5. Repeat steps (2)-(4).

B. Approaches for Selecting Poles

1. Classic Bray-Curtis Ordination
   a. First axis poles are maximally dissimilar pair
   b. Point with maximal residual distance on axis 1 is first pole of axis 2, and point farthest from this is second pole on axis 2. I.e., residuals are calculated with respect to single axis:

$$d_{r_i}^2 = d_{1i}^2 - D_{1i}^2$$

   c. Likewise for axes 3, 4, etc., with residual distance measured with respect to the most recent axis.

- Can be especially sensitive to outliers.

2. Sepkoski's Approach
   a. At each round, start with complete ordination distances (which are zero at the start of first round), not just those relative to the previous axis: $d_{O_{ij}}$ is the distance between two entities implied by the ordination.
   b. Calculate residual between true and ordination distances.

$$d_{r_{ij}}^2 = d_{ij}^2 - d_{O_{ij}}^2$$

   c. Point with maximal variance in residual distance is pole 1; point farthest from this is pole 2.

3. Beals's Approach
   a. Like Sepkoski, each round is the same and starts with residual distances relative to the entire ordination, not just the previous axis.

   b. Take every possible pair of points as candidate poles. Choose as poles the pair that leads to minimal residual distance.
- Can be modified to disallow poles that would lead to negative residual distances.
- Can be computationally expensive.

C. Properties of polar ordination (depending on the particular algorithm)

1. Axes needn't be orthogonal.
2. The same entity can serve as pole for more than one axis.
3. Residual distances can be negative (i.e. two entities can actually be more similar than implied by ordination scores).

- For these and other reasons, polar ordination is not used much anymore.

III. Nonmetric Multidimensional Scaling (NM-MDS)

A. Basic Procedure

1. Decide on number of axes (k) in advance.
2. Use iterative algorithm to find k-dimensional ordination that maximizes *rank order* correlation between original distances and distances implied by ordination.
3. Maximize this correlation by minimizing *stress*

$$stress = \sqrt{\frac{\sum (\hat{d}_{ij} - d_{MDS_{ij}})^2}{\sum d_{MDS_{ij}}^2}}$$

   where $\hat{d}$ is the value of d_{MDS} that would be required to yield a perfectly monotonic function.

B. Properties

1. Designed explicitly to minimize distortion between true distances and ordination distances. Thus it differs from, say, cluster analysis, in which small distances are well represented and large distances are distorted, and principal component analysis and other methods, in which larger distances are better represented.
2. Algorithm works largely by estimating gradient of stress and following downhill.

- Computational time goes up with square of dimension of matrix, so can be slow for large datasets.

3. Units of ordination axes arbitrary.
4. Axes not constrained to be orthogonal.
5. Useful with nonmetric distance measures, where procedures like principal component analysis are not valid.
6. Algorithm can get stuck on local optimum, so it's best to start with an approximate configuration of points. (This is what is done by the R function `isoMDS()` in the `MASS` package.)

IV. Interpretation of Ordination Axes

- General approach is to study patterns of correlation/association between original columns/rows and ordination scores.
- More on this when we get to principal component analysis.