Seriate operation; Seriate a shade plot dendrogram

Slightly more detail on how the seriate options are constructed, on both samples and species axes, can be found in Chapter 7 of CiMC but will be described here also, since the concept of a *seriation model matrix* has not yet been met (see Section 14). There are two distinct, but related, Seriate operations. If a cluster analysis is not being displayed, the •Seriate option will attempt to place the objects (samples or species) in such an order, numbered 1, 2, 3, …, that when the distance matrix among those pairs of integers is calculated (the seriation model matrix), and that distance matrix correlated element-by-element to the resemblance matrix for those objects (using a non-parametric Spearman correlation coefficient ρ), the resulting ρ is maximised. A perfect seriation here (ρ=1) would be when the order of dissimilarities among (say) the species exactly matches the seriation model matrix – the further species are away from each other in the ordered list, the greater their species dissimilarity. In practice, of course, ρ=1 values will not be attained, but the idea is to get as close to this situation as possible. However, to be sure that the optimum ρ has been found would require evaluation of all *p*! arrangements of *p* species. Even when the species list is reduced to 50, as here, this is still 50! computations – an impossibly large number – so an approximate search routine is implemented starting from different initial random species orderings. Every time you enter the **Special>Reorder** dialog for a shade plot, and exit it with **OK** – whether or not you have made changes to the requested sheets or constraints – the number of random restarts specified, for either (or both) selected Seriate option(s), will be re-run. A different solution may then be found, so that the species or sample ordering changes slightly. It is worth experimenting with larger numbers of restarts to try to optimise the search, but in the end it is simply a slightly re-ordered display and a sub-optimal solution is likely to be just as useful for interpretation as a marginally 'better' one, so this is something not to be too concerned about. Note that you can switch the axes direction(s) by **Graph>Flip X** or **Flip Y** – these are arbitrary but in some contexts you may prefer a diagonalised shade plot, such as the initial one produced above, to run from top left to bottom right, rather than bottom left to top right. If you do not want to lose a reordering of species that seems to be visually helpful, when then going on to run **Reorder** options on the samples, you can fix the species order with (Variables✓Keep variable order) – or *vice-versa*, fixing the sample order in its current state by (Samples✓Keep sample order) before making changes to the species order.

The second Seriate option is when a dendrogram has been selected to be displayed on that axis – this is specified by **Special>Reorder**>(Constraint•Sample dendrogram) on the shade plot *x* axis or (Constraint•Variable dendrogram) on its *y* axis, supplying the appropriate dendrogram/tree plot graph window. Order•Seriate is still a possible option but it is *constrained* to be consistent with a dendrogram rotation. Section 6 described how the axis order for cluster dendrograms was arbitrary to within all possible rotations of the structure, viewed as a 'mobile'. But here the (Order•Seriate) option searches through those possible rotations for one which again maximises the correlation ρ of the resemblances to the seriation model matrix. This is a greatly reduced subset of the possible set of orderings in the unconstrained case – though still needing an iterative process – and

is often more visually successful on a typically limited number of restarts. You can again **Flip X** or **Flip Y** on the resulting plot, but now also manually rotate the dendrogram after (or instead of) the Seriate option, exactly as you would for a standard cluster analysis, i.e. by clicking on the ‘bars’ of the mobile – these may be vertical or horizontal lines depending on whether the dendrogram is on the species or samples axis, respectively. After a careful manual rotation it may be particularly useful, again, to fix the current species ordering with (Variables✓Keep variable order) before tidying up the samples axis (or *vice-versa*). All such manual rotations are perfectly justifiable – in the end all we are doing is just looking at the data matrix! And a key point to remember is that multivariate analyses of samples (MDS, ANOSIM, RELATE etc.) do not care about the species order in the matrix – they will return the same results, whatever the order. The human eye, however, does find it helpful to put together species with similar responses across samples – this can make the patterns that a multivariate analysis is able to pick out automatically suddenly become visually apparent!

Revision #7
Created 7 August 2024 22:36:55 by Arden
Updated 4 February 2025 21:43:12 by Abby Miller