ON THE INTERPRETABILITY OF ORDINATION DIAGRAMS

Eddy VAN DER MAAREL

Division of Geobotany, Toernooiveld, 65 25 ED Nijmegen, The Netherlands

Keywords:
Distortion, Gradient, Horseshoe effect, Ordination

In my review of multivariate methods in phytosociology (van der Maarel 1979) I briefly commented on the interpretability of Type A distortions (see Orlóci 1974b, 1978), i.e. 'horseshoe', ('arch') effects in ordination diagrams of axes 1 and 2. Such effects are now generally recognized as the result of curvilinearities intrinsic to community data i.e. (1) the bellshaped species response curves along environmental gradients and (2) the non-linear decrease of sample similarity with increasing sample separation (Gauch, Whittaker & Wentworth 1977, see also Swan 1970, Noy-Meir & Austin 1970, Gauch & Whittaker 1972, Austin 1976, Noy-Meir & Whittaker 1978 and Whittaker & Gauch 1978 to mention a few papers I consulted again for this purpose). Hill & Gauch (1980) even say more straightforwardly: 'the arch effect is simply a mathematical artifact, corresponding to no real structure in the data', when discussing faults of reciprocal averaging.

The demonstration of this drawback with simulated coenoclines of varying β-diversity, based on series of overlapping Gaussian species response curves is convincing. My concern has been for some time that in many ordination diagrams produced in phytosociological studies of our group or with the same methods as we use (mostly ORDINA, a PCA on a Euclidean distance matrix, Roskam 1971 and also RA, Hill 1973) we found either no horseshoe at all, or we could easily interpret the arch appearing in the diagram of axes 1 and 2. This makes me somewhat doubtful about the realism of simulated coenoclines as reflections of real environmental gradients. Such coenoclines could be unrealistic in two ways: (1) the Gaussian species response model is not as universal as it is supposed to be by users of this type of simulated data. (2) the length of the coenocline in terms of the simulated β-diversity, or maybe rather the relation between β-and α-diversity (local concentration of species in a narrow part of the gradient) does not reflect the amplitude of environmental gradients or variation in community composition phytosociologists usually describe.

Before discussing this further I confirm that part of the relatively positive results we obtained may be due to the choice of resemblance function and that of the data transformation involved. As to the resemblance function: we normally use Euclidean distance. Although this measure was found to produce more deviation from expected ordination patterns of simulated data (e.g. Whittaker & Gauch 1978) it has some advantages, certainly in combination with some sort of transformation.

First it is a fully metric resemblance function which is moreover directly related to the geometric sample space ('phytosociological space', van der Maarel 1979). Second it describes sample separation along a coenocline better than similarity measures (van der Maarel & Leertouwer 1967): with similarity measures ranging from zero to one, the resemblance profile along the coenocline decreases from one to zero rather rapidly and remains zero. With Euclidean distance the profile shows an increase which can be followed by a decrease in case of species poor extremes of the cline. In any case there is no problem with high numbers of zero values in the resemblance matrix.

As to data transformation I may add that we usually apply an ordinal transformation to Braun-Blanquet cover-abundance values (van der Maarel 1979a) which we found of intermediate impact, i.e. between dominance weighting and presence weighting (cf. Gauch et al. 1977, van der Maarel 1979) and thus comparable to standardization by species (cf. Noy-Meir, Walker & Williams 1975).

This would mean that our ordinations are relatively 'favoured' as are PCA's with standardization by species (cf. Whittaker & Gauch 1978). I agree with the latter authors that relativation by relevé totals would be a useful transformation before calculating a resemblance function, especially if this would be Euclidean distance (van der Maarel 1979), but we did not try this out yet.

Let me now discuss the reality of simulated coenoclines. First I may confirm the findings of Austin (1976) that Gaussian response curves are both theoretically questionable and less frequently found in reality. Austin (1980) discusses various other models, in which the intervals along the species response is measured as a relative performance, and the intervals along the environmental gradient are taken logarithmically. He shows how the ecological displacement model of Ellenberg (see Mueller-Dombois & Ellenberg 1974) and the environmental stress model of Grime (1973, 1977) can predict bimodal response curves and wedge-type curves. I also suggested the latter type in relation to environmental dynamics (van der Maarel 1976).

If we look at phytosociological tables arranged with the help of ordination techniques (e.g. van der Maarel, Janssen & Louppen 1978, and also in Gauch et al. 1977) we observe that, indeed, the wedge-type is frequently occurring in various data-sets, whereas the Gaussian-type is almost an exception.

So in conclusion to these considerations we may state that in phytosociological data-sets as used in European ordination studies the supposed general basis of type A distortion in ordination diagrams in the form of non-linear Gaussian-type species behaviour along environmental – or phytosociological gradients does not exist.

Let us now look at the arch effect in some detail. First I confirm that in some of the ordinations I referred to above (van der Maarel et al. 1978) the diagram of axes 1 and 2 definitely showed a horseshoe shape despite the non-Gaussian behaviour of most species involved. Here Hill & Gauch's (1980) remark on the independence of axes may be relevant. However, in some clear cases of arch shape configurations representing long coenoclines the result was still interpretable, but in a special way.

The first example I examined was that of Werger, Wild & Drummond (1978). This study deals with plant community types of the Great Dyke, Rhodesia, forming a long coenocline, but with discontinuities between types according to sharp environmental boundaries due to contrasts between different bed rock types. Both PCA and RA produced an arch shape configuration in the diagram of axes 1 and 2. The underlying environmental gradient appeared to be very complex and in fact consists of three parts: (a) a part with relatively favourable soils as to toxicity (high Mg content) but with decreasing moisture content, (b) a part with a rapid increase in unfavourable heavy metal conditions, combined with a decrease in moisture conditions, and (c) a part on serpentine soils with increasing soil moisture conditions. The ordination diagram clearly shows the relative position of these three parts.

The second example is taken from Jensén's (1978) study of a Swedish lake shore vegetation. A PCA was applied to cover data of *helophyte* species in 38 subsequent 1 sq. m. plots lying in a belt transect. A clear horseshoe pattern arises with some involution, this depending on the transformation used. The sequence of relevés according to the actual position along the belt transect is acceptably reflected in cases of transformation which reduce the influence of dominant species. The transect itself is described by Jensén and appears to consist of three parts: (a) a part above low water level in which the shrub *Myrica gale* appears and which is probably characterized by an increase in moisture content of the soil; (b) a part in which the period of inundation increases from short to almost the year round; and (c) a portion which is permanently inundated with an increase in average water depth.

I think we have a rather clear case of what may be much more generally occurring in nature than so far suggested: the singular topographical gradient consists of various environmental gradients which differ as to the master factor. PCA ordination does reveal this differentiation and any ordination which might be better from other, theoretical, points-of-view but does not produce this result in two dimensions, may be called less effective!

The practical conclusion I would draw here, is to split the data set into subsets and apply an ordination to each of them. This was done in a similar study of a topographical limestone gradient in grassland (Louppen in prep., see Louppen, Werger & Eppink 1978) and – as Austin (pers. comm.) presumed – the ordination diagrams of parts of the entire sequence again produced horseshoe configurations, be it less pronounced.

The latter remark brings me to a differentiated conclusion: ordination diagrams should not be disregarded simply because they seem to produce ineffective results. As long as an ordination diagram is interpretable it should be used, and of course the ordination technique being used, be welcomed. In this respect the user should be careful about the real dimensionality of the coenocline under study. What looks like a one dimensional variation in the phytosociological space (belt transect) may well appear as a more – dimensional partly dependent variation in the ecological space. On the other hand: mathematical artifacts due to linear constraints in a method applied to non-linearly varying variables should be overcome. Therefore we are anxious to see results of the new detrended correspondence analysis of Hill (1979, Hill & Gauch 1980) when applied to phytosociological data sets of the kind discussed here.

References

Austin, M.P. 1976. On non-linear species response models in ordination. Vegetatio 33: 33–41.

Austin, M.P. 1980. Searching for a model for use in vegetation analysis. In: E. van der Maarel (ed.) Advances in vegetation science: Classification and ordination. Vegetatio 42: 11–21.

Gauch, H.G. & R.H. Whittaker. 1972. Coenocline simulation. Ecology 53: 446–451.

Gauch Jr. H.J., R.H. Whittaker & T.R. Wentworth, 1977. A comparative study of reciprocal averaging and other ordination techniques. J. Ecol. 65: 157–174.

Grime, J.P. 1973. Control of species density in herbaceous vegetation. J. Envir. Manage. 1: 151–167.

Grime, J.P. 1977. Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. Amer. Natur. 111: 1169–1194.

Hill, M.O. 1973. Reciprocal averaging: An eigenvector method of ordination. J. Ecol. 61: 237–249.

Hill, M.O. 1979. Decorana. A Fortran program for detrended correspondence analysis and reciprocal averaging. Ecology and Systematics, Cornell University, Ithaca. 52 pp.

Hill, M.O. & H.G. Gauch Jr. 1980. Detrended correspondence analysis, an improved ordination technique. In: E. van der Maarel (ed.). Advances in vegetation science: Classification and ordination. Vegetatio 42: 47–58.

Jensén, S. 1978. Influences of transformation of cover values on classification and ordination of lake vegetation. Vegetatio 37: 19–31.

Louppen, J.W.M., M.J.A. Werger & J.H.M. Eppink. 1978. Vegetation boundaries and species performances along an environmental gradient. Abstracts 2nd. Int. Congr. Ecology, Jerusalem, p. 217.

Maarel, E. van der. 1972. Ordination of plant communities on the basis of their plant genus, family and order relationships. In: E. van der Maarel & R. Tüxen (ed.). Grundfragen und Methoden der Pflanzensoziologie. Ber. Int. Symp. Rinteln 1970, Junk, Den Haag. p. 183–192.

Maarel, E. van der. 1976. On the establishment of plant community boundaries. Ber. Deutsch. Bot. Ges. 89: 415–443.

Maarel, E. van der. 1979. Multivariate methods in phytosociology, with reference to the Netherlands. In M.J.A. Werger (ed.) The study of vegetation, p. 161–225.

Maarel, E. van der. 1979a. Transformation of cover-abundance values in phytosociology and its effects on community similarity. Vegetatio 39: 97–114.

Maarel, E. van der, J.G.M. Janssen & J.M.W. Louppen. 1978. TABORD, a program for structuring phytosociological tables. Vegetatio 38: 143–156.

Maarel, E. van der & J. Leertouwer. 1967. Variation in vegetation and species diversity along a local environmental gradient. Acta Bot. Neerl. 16: 211–221.

Mueller-Dombois, D. & H. Ellenberg. 1974. Aims and methods of vegetation ecology. Wiley, New York, XI + 547 pp.

Noy-Meir, I. & P. Austin. 1970. Principal component ordination and simulated vegetation data. Ecology 51: 551–552.

Noy-Meir, I.D. Walker & W.T. Williams. 1975. Data transformations in ecological ordination. II. On the meaning of data standardization. J. Ecol. 63: 779–800.

Noy-Meir, I. & R.H. Whittaker. 1978. Recent developments in continuous multivariate techniques. In R.H. Whittaker (ed.) Ordination of plant communities, p. 337–378.

Orlóci, L. 1974. On information flow in ordination. Vegetatio 29: 11–16.

Orlóci, L. 1978. Multivariate analysis in vegetation research. 2nd. ed. Junk, The Hague, Boston, IX + 451 pp.

Roskam, E. 1971. Programme ORDINA: Multidimensional ordination of observation vectors. Programme Bull. 16, Psychology Lab. Nijmegen, 8 pp.

Swan, J.M.A. 1970. An examination of some ordination problems by use of simulated vegetational data. Ecology 51: 89–102.

Werger, M.J.A., H. Wild & B.R. Drummond. 1978. Vegetation structure and substrate of the Northern part of the Great Dyke, Rhodesia: Gradient analysis and dominance-diversity relationships. Vegetatio 37: 151–161.

Whittaker, R.H. & H.G. Gauch Jr. 1978. Evaluation of ordination techniques, 2nd ed. In R.H. Whittaker (ed.) Ordination of plant communities, p. 277–336. Junk, The Hague, 388 pp.