Abstracta Botanica 14:109–119, 1990

Rigid rotation of nonmetric multidimensional scaling axes to environmental congruence

N.C. Kenkel and C.E. Burchill

Department of Botany, University of Manitoba,

Winnipeg, Manitoba R3T 2N2 Canada

Abstract: One drawback of nonmetric multidimensional scaling (NMDS) ordination is that the axes have no meaning except insofar as they define a Cartesian coordinate system for the ordinated points (individuals). In this paper we present a straightforward strategy for examining vegetation-environment relationships that uses canonical correlation analysis to define a rigid rotation of the NMDS configuration which corresponds to the principal directions of trended environmental variation. It is assumed that information on both vegetation (e.g. species presence, abundance) and environmental factors has been collected for each relevé. The strategy can be summarized in four steps: (1) apply NMDS to a standard vegetation data set to obtain a two-dimensional ordination of relevés; (2) perform a canonical correlation analysis, with the two-dimensional NMDS relevé scores as the first variable set and the q environmental factors measured on each relevé as the second variable set; (3) since the canonical correlation axes define a rigid rotation of the NMDS configuration (to environmental congruence), relevé positions can be plotted in the rotated space after accounting for non-orthogonality of the canonical axes; (4) use the computed structure correlations to determine which of the environmental variables are most strongly correlated with trends in the NMDS ordination space. The utility of the method is illustrated using data from 135 relevés collected from inland boreal saline habitats near Overflowing River, Manitoba, Canada.

Introduction

Ordination is a well-established strategy for data reduction (Gnanadesikan 1977), and has long been used in ecology for both data presentation and factor revelation (Orlóci 1978). A large number of ordination procedures have been described, and the ecological literature contains a number of studies discussing the relative merits and faults of the various strategies. While some workers continue to champion a particular method or strategy, most recognize that no one ordination strategy will suit all needs; rather, the objectives of the analysis and the nature of the data collected must be considered in making intelligent decisions (Kenkel and Orlóci 1986, Minchin 1987).

There are many ways of classifying ordination strategies. One of the most useful is to distinguish between metric methods of multidimensional scaling (which use eigenanalysis to define ordination axes) and nonmetric multidimensional scaling procedures. Of the metric methods, principal components analysis (PCA) was the first to be used in ecology. Because PCA performs eigenanalysis on a correlation (covariance) matrix, it assumes an underlying linear data structure. As a consequence, nonlinear trends are distorted into higher dimensions, rendering interpretation of ordination scattergrams difficult. Principal co-ordinates analysis (PCoA) is a more flexible metric model, as it can perform an eigenanalysis of virtually any dissimilarity matrix between individuals (although, strictly speaking, only metric coefficients should be used to prevent negative eigenvalues; see Digby and Kempton 1987). Despite its flexibility and apparent utility, PCoA has not been commonly used. Instead, correspondence analysis (CA) and its variants are the 'methods of choice' in plant ecology. CA can be 'derived' in a number of ways (Orlóci 1978, ter Braak 1987); in the eigenanalysis derivation, CA involves a linear additive partitioning of the total chi-square. Empirical studies indicate that CA is often more effective than PCA in summarizing nonlinear data structures. However, the second CA axis is often simply a quadratic function of the first, containing little information and rendering interpretation difficult. To correct for this and other perceived faults, a method of 'detrending' correspondence analysis was developed (DCA, Hill and Gauch 1980). The detrending procedure has been strongly criticized on both theoretical and empirical grounds (Wartenberg et al. 1987), although Peet et al. (1988) defended the procedure based on empirical evidence. Jackson and Somers (1990) have recently presented data indicating that detrending is arbitrary and produces unstable and unreliable results. Detrending by polynomials (ter Braak 1987) may overcome some of these problems, but more empirical tests are required. DCA should therefore be used with some caution; if it is applied, users are advised to also perform CA in order to objectively assess the results of the detrending procedure (Kenkel and Orlóci 1986).

Nonmetric Multidimensional Scaling

This ordination method attempts to find the best possible rank-order approximation between dissimilarities in resemblance space and interpoint distances in ordination space. It differs from the metric eigenanalysis strategies described above in a number of respects: (a) since no eigenanalysis is performed, no assumptions are made regarding the underlying structure of the data, whether linear, 'Gaussian' or otherwise. (b) nonlinear data structures are readily accounted for since the method can accommodate virtually any measure of association between relevés. (c) the method requires the user to specify the dimensionality of the final solution; this differs from the metric methods, which 'discard' higher axes to obtain the dimensionality desired. (d) NMDS axes merely define an arbitrary Cartesian coordinate system for the relevés in ordination space; by contrast, the axes of PCA and other eigenanalysis methods are 'extracted' in order of decreasing importance, and represent direction of trended variation in vegetation space. (e) NMDS finds an ordination solution by an iterative procedure of successive approximations. The solution is therefore not unique, though in practice ordination solutions will be very similar if the data are reasonably well structured.

Ordination and the Examination of Vegetation-Environment Relationships

Data reduction is normally only the first step in a complete analysis of ecological information. As a second step, vegetation - environment relationships are examined by asking the following questions: (a) how strong is the overall relationship between vegetation trends and the measured environmental factors, and (b) which environmental factors are most highly correlated with the vegetation trends, and which are of lesser importance? Various procedures have been developed to answer these questions. The most common approach involves computation of simple or multiple correlations between axis scores and a given environmental variable. These and other 'indirect' methods for examining vegetation - environment relationships are outlined in ter Braak (1987, p. 132). Sophisticated 'direct' strategies incorporating canonical analysis (Gittins 1985) are a more recent development. An example is canonical correspondence analysis (CCA), a 'restricted correspondence analysis' approach in which "... the site scores are restricted to be a linear combination of measured environmental variables" (ter Braak 1987). This is a potentially powerful approach, but because it makes a number of assumptions regarding data structure it may not be robust. The relative advantages and disadvantages of the 'indirect' and 'direct' approaches are discussed in detail by ter Braak and Prentice (1988).

Strategies for examining vegetation - environmental relationships in nonmetric multidimensional scaling have also been developed. Multiple correlations and graphical presentations can be used together to reveal environmental trends in the scattergram; see Kenkel (1987) for an example. Such a univariate strategy is very useful, but it ignores the multivariate nature of the environmental information. Furthermore, the problem of presentation of the ordination results is unresolved since the NMDS axes remain 'arbitrary'.

Axis Rotation to Environmental Congruence

We suggest the following 'indirect' approach for examining vegetation - environment relationships in NMDS ordination. Application of the method achieves two goals: it takes a multivariate approach to the examination of correspondence between vegetation and environmental factors, and defines a rigid rotation of the ordination configuration to environmental congruence. The method is outlined in the following steps:

1. Select an appropriate resemblance measure for use as input to NMDS. The measure chosen should follow from considerations of the data structure, for it is at this stage that inherent nonlinearities in data are accounted for (Kenkel

and Orlóci 1986, Bradfield and Kenkel 1987). Compute the resemblance matrix between relevés and use this to obtain a two-dimensional NMDS ordination.

2. Apply canonical correlation analysis, using the relevés 'scores' of the NMDS configuration as the first variable set ($p = 2$ variables) and the q environmental variables as the second set (where $q \geq 2$). This analysis will produce two uncorrelated canonical axes representing the major directions of trended environmental variation in the NMDS configuration. Note that because only two canonical axes are produced (since $q \geq p = 2$), they define a rigid rotation of the NMDS ordination. Canonical analysis also produces scores for the relevés in the rotated space, and axis multiple correlations which measure linear correspondence between vegetational variation and environmental factors.

3. The next step is to represent the relevé configuration in the rotated canonical space. However, plotting canonical scores in the usual way (with axes 1 and 2 at 90°, defining a Cartesian coordinate system) will distort the original NMDS configuration, for the simple reason that canonical axes, while linearly uncorrelated, are not orthogonal (Gittins 1985, p. 150). Distortion is avoided by presenting the canonical axes at their true angle of intersection, which is calculated as the vector product:

$$\mathbf{V}_1\mathbf{V}_2'$$

where $\mathbf{V}_1$ contains the canonical weights of the two NMDS axes on the first canonical axis, and $\mathbf{V}_2'$ their weights on the second canonical axis. The scalar obtained is the cosine of the angle between canonical axes 1 and 2 (in radians). A simple trigonometric function is then used to represent the original NMDS configuration, undistorted but rotated to environmental congruence.

4. In the final step, examination of the canonical structure correlations determines which of the environmental variables are most highly associated with the two canonical (environmental) axes. It may also be useful to produce an ordination biplot (see ter Braak 1987, for details).

An Example

The method is illustrated using vegetation and environmental data collected from inland boreal saline habitats near the shores of Dawson Bay, Lake Winnipegosis, Manitoba, Canada. Saline (predominantly sodium chloride) seeps in the area have led to the development of strong salinity and vegetation gradients. A detailed description of the study area, and a more complete analysis of the data, can be found in Burchill and Kenkel (1990).

Table 1. Structure correlations of the eight environmental variables on the two canonical correlation axes. All variables are soil factors except for relative elevation.

	Canonical Axis	
	I	II
Total Salts	-0.813	0.415
pH	-0.756	-0.101
Relative Elevation	0.643	-0.026
Bulk Density	-0.488	0.456
Potassium	-0.440	0.328
Phosphorus	0.220	-0.132
Nitrogen	-0.170	-0.180
Percent Organic Matter	0.145	-0.301

Percent cover of all species was recorded in each of 135 relevés (quadrats measuring 0.5 x 0.5 m). In addition, eight environmental variables (seven soil factors, plus 'relative' elevation of the relevé above the unvegetated salt pan) were measured (Table 1). Relevés were ordinated (based on the vegetation data) using nonmetric multidimensional scaling, specifying a two-dimensional solution and using as input a chord distance matrix after flexible shortest path adjustment (Bradfield and Kenkel 1987). Shortest path adjustment was used since the data structure was highly nonlinear. The two variable sets (NMDS ordination scores, $p = 2$, and environmental variables, $q = 8$) were then subjected to canonical correlation analysis and the results interpreted as outlined in the previous section.

The rotated two-dimensional NMDS ordination is presented in Fig. 1. Multiple correlations for the two canonical axes are $R_1^2 = 0.674$ and $R_2^2 = 0.275$, indicating that the first axis summarizes most of the trended vegetation - environment variation. Canonical weights for the original NMDS 'axes' on the two canonical axes are:

$$\mathbf{V}_1 = [0.4261 \quad -0.6575] \qquad \mathbf{V}_2 = [1.3124 \quad 1.2131]$$

the scalar product $\mathbf{V}_1\mathbf{V}_2' = -0.2384$, and $\cos^{-1}(-0.2384) = 1.8115$ radians, or 103.79^o. Clearly the two canonical axes are not orthogonal.

Structure correlations for the environmental variables on the two canonical axes are presented in Table 1. Total salts, relative elevation, and soil pH are most highly correlated with the first canonical axis. This axis is interpreted as a salinity gradient from left (salinity highest in the Salt Pan and *Puccinellia* vegetation types) to right (salinity lowest in the *Calamagrostis* and *Rosa* vegetation types).

For illustrative purposes, the same scattergram was produced under the assumption that the canonical axes are orthogonal (Fig. 2). Distortion of the original NMDS configuration (see Fig. 1) is apparent, particularly for points with large negative or positive scores on the second axis. For example, relevés belonging to the Salt Pan and *Triglochin* vegetation types appear closer to the *Hordeum* and *Spartina* types than they actually are.

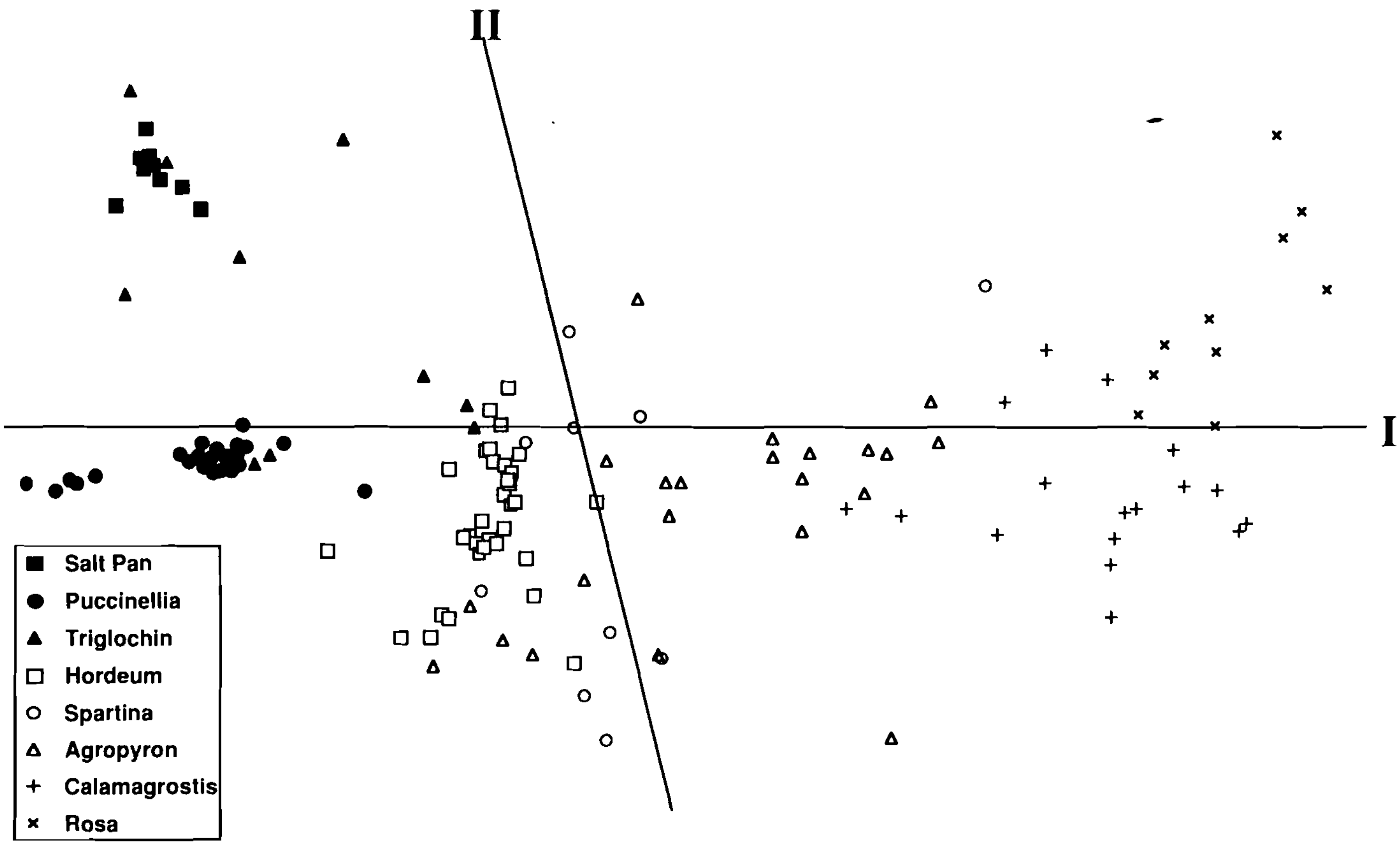

Figure 1. Rotated nonmetric multidimensional scaling ordination configuration, in two dimensions, of the 135 salt pan relevés. Axes I and II are the rotated canonical axes (non-orthogonal, angle of 103.79°). For presentation purposes, the affinity of each relevé to one of eight vegetation types is also indicated (see Burchill and Kenkel 1990 for details).

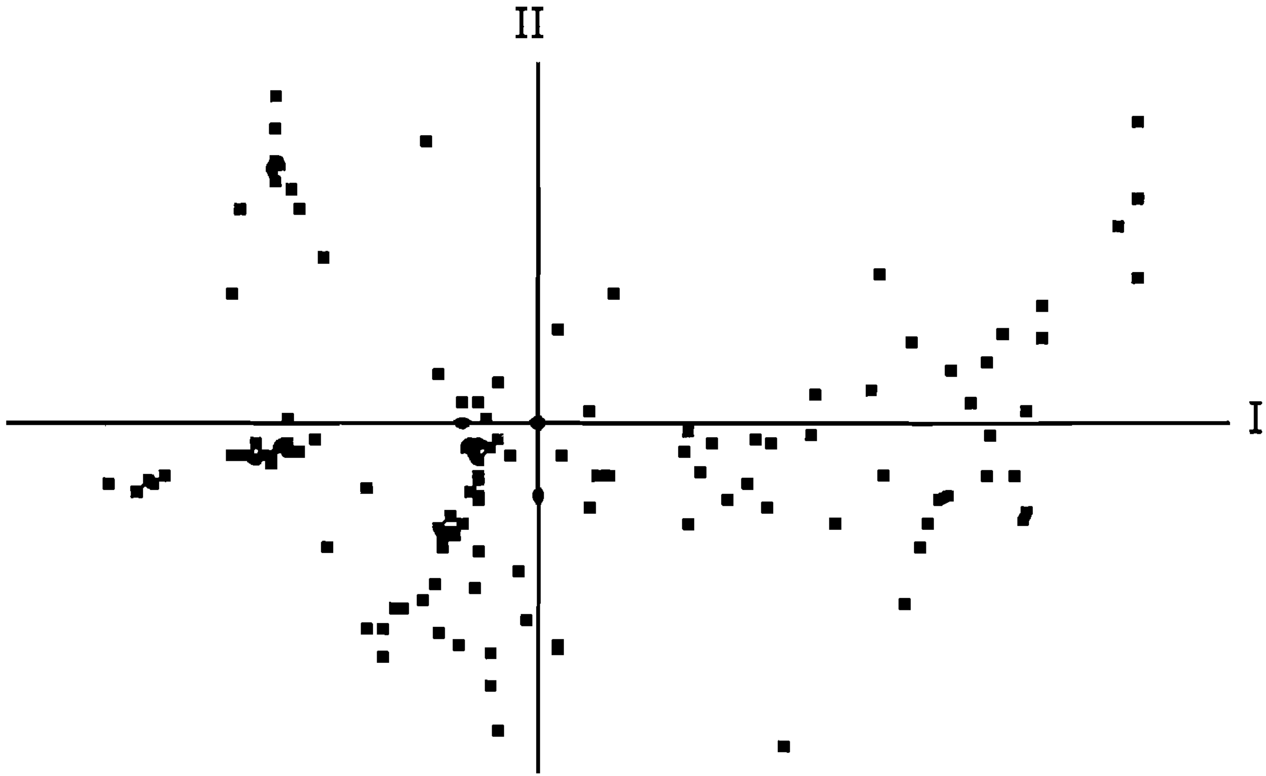

Figure 2. Ordination configuration of the 135 relevés (as in Figure 1), but with axes I and II represented orthogonally (90°). Note the distortion in the configuration.

Discussion

The method described in this paper accomplishes two goals: it determines which environmental factors are most highly correlated with vegetational variation (as revealed by the NMDS configuration), and it offers an objective method for fitting ecologically meaningful axes to the configuration. Because these axes result from a rigid canonical rotation, the ordination can be represented without distortion if non-orthogonality of the canonical axes is accounted for.

There is accumulating evidence to suggest that NMDS is a comparatively robust ordination technique, and one which may be particularly useful in the summarization of highly nonlinear data structures (Kenkel and Orlóci 1986, Bradfield and Kenkel 1987, Minchin 1987). Whenever NMDS is used to examine vegetation - environment relationships, we suggest that the method outlined in this paper be employed to objectively define ordination axes which describe major trends in environmental variation.

Acknowledgments

This research was supported by Natural Sciences and Engineering Research Council of Canada operating grant A3140 to N.C. Kenkel.

References

Bradfield, G.E. and N.C. Kenkel. 1987. Nonlinear ordination using flexible shortest path adjustment of ecological distances. Ecology 68: 750-753.

Burchill, C.E. and N.C. Kenkel. 1990. Vegetation-environment relationships of an inland boreal salt pan. Can. J. Bot. (in press).

Digby, P.G.N. and R.A. Kempton. 1987. Multivariate analysis of ecological communities. Chapman and Hall, London.

Gittins, R. 1985. Canonical analysis with applications in ecology. Springer-Verlag, Berlin.

Gnanadesikan, R. 1977. Methods for statistical analysis of multivariate observations. Wiley, New York.

Hill, M.O. 1974. Correspondence analysis: a neglected multivariate method. Appl. Stat. 23: 340-354.

Hill, M.O. and H.G. Gauch. 1980. Detrended correspondence analysis: an improved ordination technique. Vegetatio 42: 47-58.

Jackson, D.A. and K.M. Somers. 1990. Putting things in order: the ups and downs of detrended correspondence analysis. Am. Nat. (in press).

Kenkel, N.C. 1987. Trends and interrelationships in boreal wetland vegetation. Can. J. Bot. 65: 12-22.

Kenkel, N.C. and L. Orlóci. 1986. Applying metric and nonmetric multidimensional scaling to ecological studies: some new results. Ecology 67: 919-928.

Minchin, P.R. 1987. An evaluation of the relative robustness of techniques for ecological ordination. Vegetatio 56: 1167-1179.

Peet, R.K., R.G. Knox, J.S. Case and R.B. Allen. 1988. Putting things in order: the advantages of detrended correspondence analysis. Am. Nat. 131: 924-934.

ter Braak, C.J.F. 1987. Ordination. In: Jongman, R.H.G., C.J.F. ter Braak and O.F.R. van Tongeren (eds.). Data analysis in community and landscape ecology, pp. 91-173. Pudoc, Wageningen.

ter Braak, C.J.F. and I.C. Prentice. 1988. A theory of gradient analysis. Adv. Ecol. Res. 18: 271-317.

Wartenberg, D., S. Ferson and F.J. Rohlf. 1987. Putting things in order: a critique of detrended correspondence analysis. Am. Nat. 129: 434-448.