Reciprocal Averaging: An Eigenvector Method of Ordination
Author(s): M. O. Hill
Source: *Journal of Ecology*, Vol. 61, No. 1 (Mar., 1973), pp. 237-249
Published by: British Ecological Society
Stable URL: https://www.jstor.org/stable/2258931
Accessed: 28-09-2018 15:20 UTC

# RECIPROCAL AVERAGING: AN EIGENVECTOR METHOD OF ORDINATION

By M. O. HILL*

*School of Plant Biology, University College of North Wales, Bangor*

## INTRODUCTION

Principal components analysis now has a long history of use in phytosociological surveys. It was originally introduced to analyse quantitative cover data (Goodall 1954) but has subsequently been found to give satisfactory results with presence–absence data. The reasons why the method is successful with such data have not been made clear. Both Gower (1966) and Orloci (1966) work from the assumption that an ordination should reflect as accurately as possible a distance-matrix between the stands. Gower's approach is more general in that he gives a method which can start from any suitable measure of distance, whereas Orloci considers only the Euclidean distance. Geometrically, presence–absence data can be regarded as a set of points situated at the vertices of a hypercube, and there is no *a priori* reason why a satisfactory ordination need be based on the distances between the vertices. The method described here does not explicitly use a distance-matrix, and its rationale is developed from a different standpoint.

A development which has run parallel to that of principal components has been that of gradient analysis pioneered by Whittaker in the United States, and also the rather similar methods of the Wisconsin school (for a review of these and Whittaker's techniques see Greig-Smith 1964, pp. 192–6). Whittaker (1967), in his comprehensive review of ordination methods, maintains that the mathematically simpler methods of ordination give better results. In particular, he favours 'direct' methods of ordination as opposed to 'indirect'. Since his use of terminology is confusing—particularly his use of 'gradient analysis' to mean any method of ordination—it is not adopted here. By 'indirect gradient analysis' he means ordination based upon floristics alone, while by 'direct gradient analysis' he means ordination based on some evident and apparently important physical gradient. In what follows 'direct gradient analysis' will be called simply 'gradient analysis'. The method of gradient analysis is to take some well-marked gradient—e.g. altitude in the Great Smoky Mountains (Whittaker 1956)—and to assign scores to the species according to their altitudinal preferences. Sites are then ordinated by taking averages of the scores of the species which occur in them.

As Whittaker points out, the great advantage of this method is that there is never any difficulty with interpretation. The site score is roughly an index of altitude, or of water-relations, or of some other well-understood physical variable—whereas the axes derived from 'indirect' (floristic) methods such as principal components may be hard to interpret or are sometimes uninterpretable. This is an important point: in cases where gradient analysis can be used it may well be the most satisfactory way of ordinating the vegetation. However, the physical interpretation of a floristic gradient is commonly not apparent;

* Present address: The Nature Conservancy, Penrhos Road, Bangor, Caernarvonshire.

indeed the gradient itself may not be obvious. A floristic ordination is then the only kind of ordination possible.

A combination of Whittaker's gradient analysis and the method of successive approximation leads to a technique of floristic ordination rather similar to principal components. This is described here as 'reciprocal averaging'. Whittaker's (1967, p. 217) fundamental objection to floristic ordinations is that 'when an ordination based on species distributions is used to study species distributions the approach is circular'. But Poore (1962) points out that much of scientific research can be regarded as successive approximation, whereby a rough conceptualization leads to new concepts and a revision of the old. In phytosociological terms this means that describing the distributions of individual species in terms of a floristic ordination involves a circularity which need not be vicious.

## THE METHOD

Gradient analysis can be combined with a scheme of successive approximation as follows. From a rough floristic gradient, e.g. a gradient in water-relations, the species are divided into 'wet' and 'dry' species. An initial gradient analysis of sites can be obtained by scoring 0 for each 'wet' species and 100 for each 'dry' species, and using average species scores to define stand scores. A stand containing only 'dry' species will score 100 and a stand containing only 'wet' species will score 0. A stand containing half and half will score 50, and so on. Now provided that the stand scores are at this stage a reasonable indication of water-relations, they can be used to derive a new and improved calibration of the species. The new species-scores are the averages of the scores of the stands in which they occur. In the example above, species which occur mainly in wet stands will have a low score, whereas species which occur mainly in dry stands will have a high score. Intermediate species and ubiquitous species will have an intermediate score. If the new species scores are rescaled to run from 0 to 100, the process can be repeated and the stands re-calibrated. A worked example is given in Appendix 2.

By simple matrix algebra, it can be shown that this process of repeated cross-calibration gives a unique one-dimensional ordination both of the species and of the stands. In other words, after enough iterations, the scores will stabilize and they may then be taken as the final ordination. The process may be called 'reciprocal averaging' because the species-scores are averages of the stand-scores and reciprocally the stand-scores are averages of the species-scores. It is important to note that in the iterative process the final scores do not depend on the initial scores, although the number of iterations required to reach them does. A good initial guess will reduce the number of iterations needed.

The foregoing definition of the process of reciprocal averaging defines an algorithm which will generate a unique one-dimensional ordination both of the species and of the stands. The algorithm makes the rationale of the method obvious, but does not convey its close relation to principal components analysis. A mathematical proof of the close similarity of the two methods is given in Appendix 1. It is worth noting, however, that the method can readily be extended to cover quantitative as well as presence–absence data. The mathematical treatment makes this obvious.

Now although the two-way successive calibration will converge to a unique solution, there are also other 'solutions' which have the reciprocal averaging property—species score = average stand score for those stands in which the species occurs (but scaling so that the total range is from 0 to 100); and stand score = average species score for those species which occur in the stand (but rescaling so that the total range is still from 0 to 100).

These other 'solutions' are the analogues of axes other than the first principal axis in principal components. Their existence cannot be demonstrated without the use of matrix algebra, so the proof is relegated to Appendix 1; but the consequences for the user are quite simple. Thus, by use of reciprocal averaging a stand ordination is reached very comparable to what would have been obtained from principal components analysis; and to go with this will be a species ordination which is markedly more meaningful than the principal components species loadings. According to the rationale of reciprocal averaging, species ordinations and stand ordinations come in dual pairs, neither of which has logical pre-eminence.

The symmetry of the method is attractive but does not give it any particular advantage over a conventional principal components analysis. What does confer the advantage is that it uses a good species ordination to derive the stand ordination, and *vice versa*. This is not true, for example, of unstandardized principal components analysis, alias Orloci's (1966) analysis of the 'weighted similarity coefficient'. For if there are two species, one of which has a similar distribution to that of another but is rarer, they will with reciprocal averaging have much the same species score; but in an unstandardized principal components analysis the rarer will usually have a smaller axis loading. In other words, the species scores derived by reciprocal averaging are corrected for species abundance, but those obtained by unstandardized principal components analysis are not. If, however, the data are standardized so that each species has unit variance, principal components may give an ordination which closely resembles that obtained by reciprocal averaging of the unstandardized data. A proof is given in Appendix 1. The main advantage of reciprocal averaging, then, is that it gives good species ordinations to go with the stand ordinations, not that it gives markedly better stand ordinations.

The method is not new, having been developed under the title 'Analyse Factorielle des Correspondances' by J. P. Benzécri and his co-workers in France (Benzécri 1969; Escofier-Cordier 1969). The translation 'Factorial Analysis of Correspondences' is so vague that an alternative name is desirable; 'reciprocal averaging' is succinct and appropriate. Moreover, the rationale which makes the method particularly suitable for use in phytosociological ordinations is different from the geometric rationale developed by Benzécri. Phytosociologically the method arises as a natural extension of gradient analysis, whereas Benzécri's development is more appropriate to the case where the data matrix is a table of the numbers of joint occurrences (correspondences) of individuals of a like kind. For example, he considers a case where the data matrix is a set of numbers $a_{ij}$ which are counts of the times that a word $i$ is rhymed with another word $j$ in a poem.

More doubtful is whether one should distinguish reciprocal averaging by a name at all. Thus Noy-Meir's (1970) lucid comparison of the many varieties of principal components analysis alludes to Benzécri's method. According to Noy-Meir's broad definition, reciprocal averaging is itself merely a variant of principal components analysis, with a particular standardization, no centring and a particular post-normalization. But such a characterization does not emphasize the duality of the species and the stand ordinations; as this property is the method's chief strength and peculiarity, the name 'reciprocal averaging' is worth retaining.

## RESULTS AND DISCUSSION

An ALGOL program has been written to perform the calculations of reciprocal averaging, and these present no difficulty. The resulting ordinations have in practice been at least as

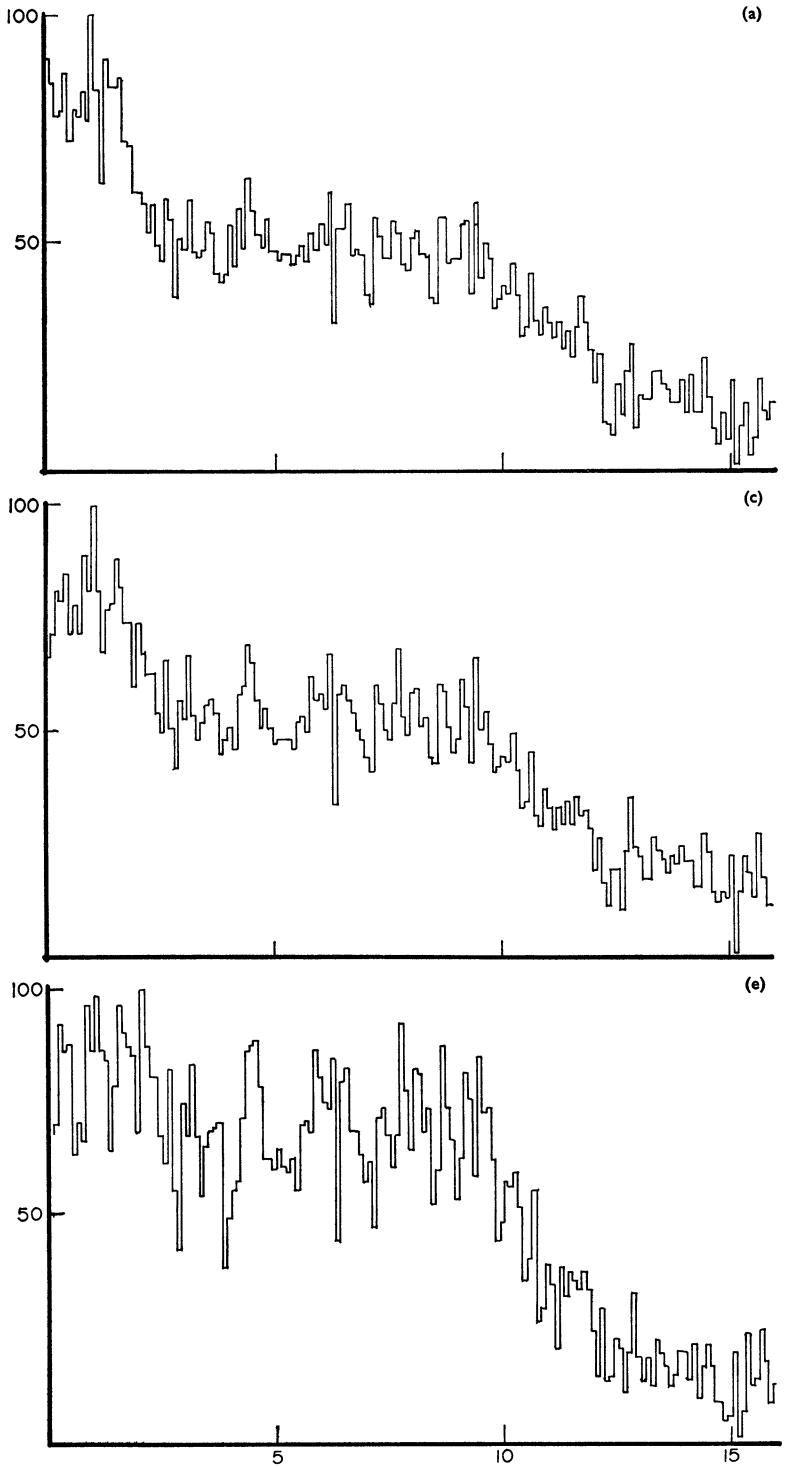
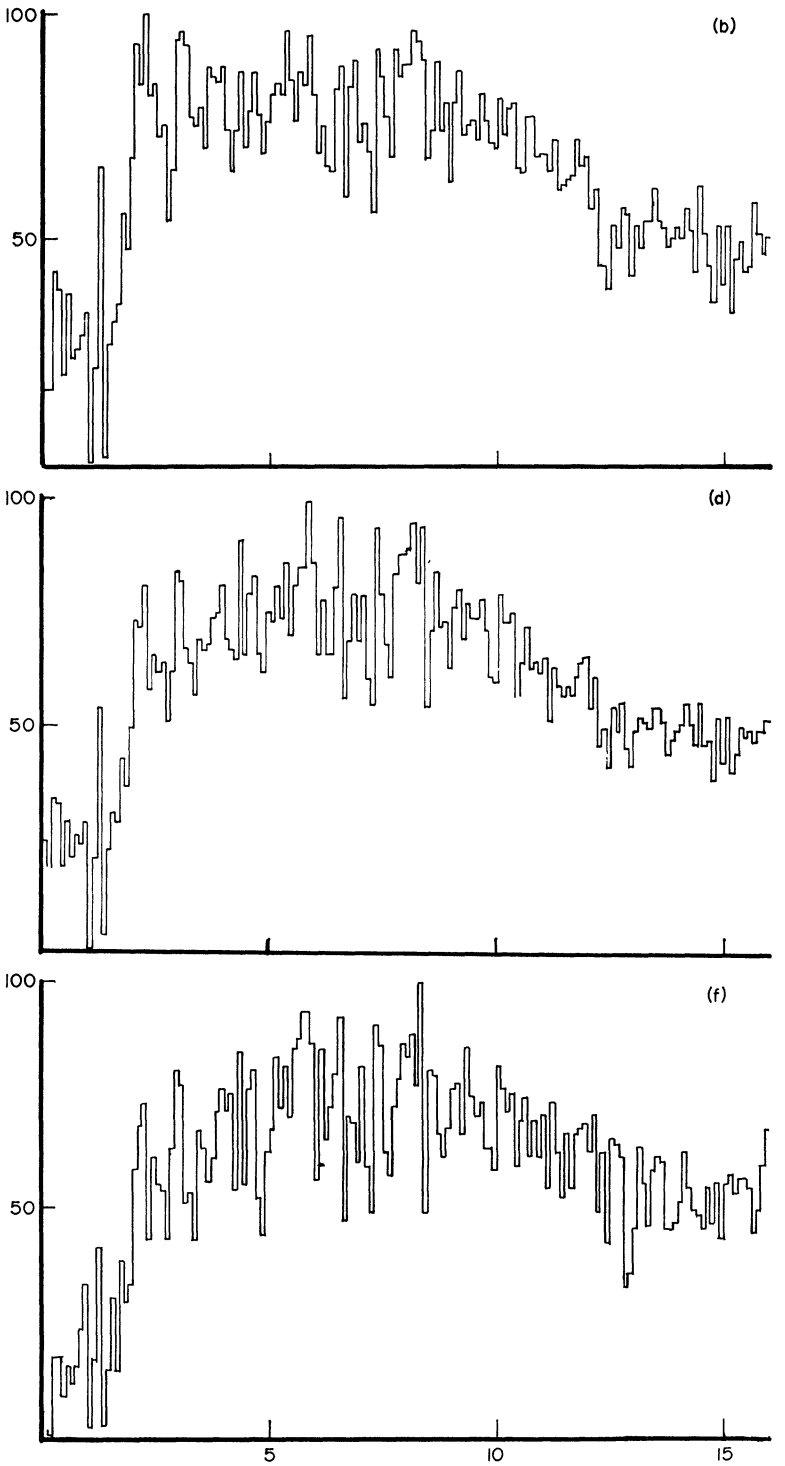



FIG. 1. Transect profiles of the first (a, c, e) and second (b, d, f) axes derived by ordinating presence–absence data from 10 cm × 10 cm quadrats along a 16-m belt transect in grassland. The *x*-axes represent position along the transect; the *y*-axes represent the values of the




ordination scores, scaled arbitrarily to run from 0 to 100. (a) and (b), Reciprocal averaging; (c) and (d) principal components using standardized data; (e) and (f), principal components using unstandardized data.

satisfactory as those obtained by principal components analysis, and where different they have been preferable. The main theoretical difference is that where there is a long floristic gradient it will always be presented linearly on the first axis of an ordination by reciprocal averaging. This is in contrast to the results of Swan (1970) and Noy-Meir & Austin (1970) which show that with principal components analysis, where there is a long and strong floristic gradient, stands which are extreme on the first axis of the ordination need not be extreme on the floristic gradient, and *vice versa*.

Floristic gradients of the type considered by Swan and Noy-Meir & Austin are not particularly common in practice, and a less extreme example is given here to exhibit the differences which may result from a comparison of ordination methods. The methods

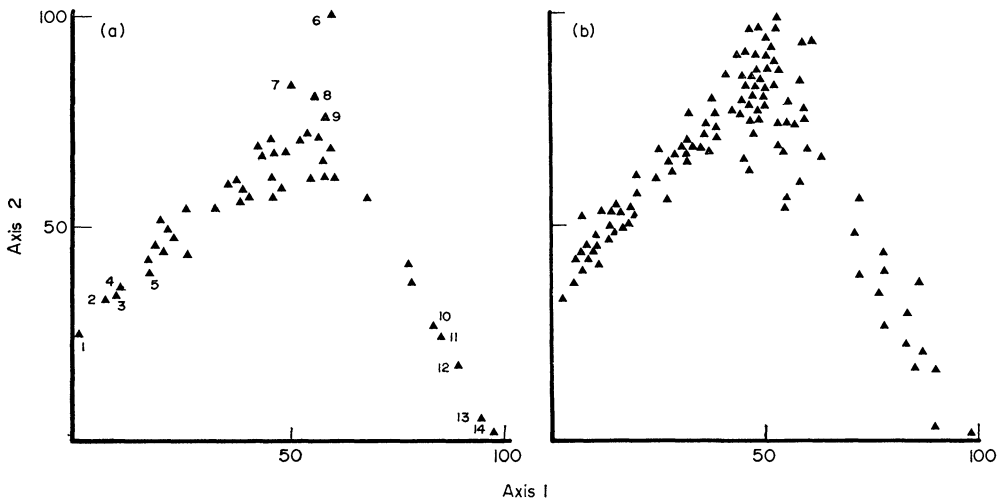



FIG. 2. Scatter diagrams of the first two axes of (a) the species and (b) the stand ordinations derived by reciprocal averaging from the same data as in Fig. 1. The numbered species (a) are: 1, *Glaux maritima*; 2, *Lythrum salicaria*; 3, *Agrostis stolonifera*; 4, *Cardamine pratensis*; 5, *Lychnis flos-cuculi*; 6, *Carex pilulifera* (occurred in only one stand); 7, *Juncus articulatus*; 8, *Achillea ptarmica*; 9, *Trifolium pratense*; 10, *Luzula multiflora*, 11, *Succisa pratensis*; 12, *Sieglingia decumbens*; 13, *Polygala serpyllifolia*; 14, *Euphrasia brevipila*. Nomenclature follows Clapham, Tutin & Warburg (1962).

under consideration are: (1) reciprocal averaging; (2) principal components using the correlation matrix (i.e. on data standardized by species to unit variance); and (3) principal components using the covariance matrix (i.e. on unstandardized data).

The data are presences and absences of species in 10-cm square quadrats along a transect in a sandy pasture in North Wales.

The results are presented in Fig. 1 as transect profiles. One thing which stands out immediately is that unstandardized principal components (method 3) divides the variation along the transect into two parts, one part being represented by differences in the scores on the first axis and the other by differences in the scores on the second. Methods 1 and 2 give a more satisfactory first axis, which does display the whole range of the floristic gradient. The ordinations generated by these two methods are very similar; not only are the first axes almost the same but the second axes are in this case essentially a quadratic function of the first. This phenomenon was noted by Goodall (1954) and is clearly visible in the ordination diagrams of many subsequent authors (e.g. Harberd 1962). It means

that the 'independence' of the second axis from the first is quite spurious. The second axis is (with principal components) orthogonal to the first, but it is independent only when the underlying structure of the data is multivariate normal. With presence–absence data such a condition can never be realized, and with reciprocal averaging the axes are not even orthogonal.

Scatter diagrams (Fig. 2) are included to show that with reciprocal averaging the quadratic relation between the first and the second axes applies both to the species and to the stand ordinations. Species which are high on the second axis tend to occur in the middle part of the transect, and those which are low tend to occur at the extremities. Fig. 2 shows another important fact, that the 'quadratic' relation between the first two axes is only approximate quadratic, and might as well be described as two straight lines. The explanation is simple. The second axis reflects presences and absences of extreme species, while the first is a more general average of the species-complement. That the general average should have a locally linear relation to an average based on the occurrences of extreme species is no wonder.

The phenomenon of the higher axes being dependent on the lower is an important feature of floristic ordinations. The higher axes should not be supposed therefore to give no extra information. In the example considered here the relation between the first two axes is indeed so close that any extra information is scanty. But in other cases the relation, although close, leaves room for such considerations as that a given species has an unexpectedly high second-axis score in relation to its first-axis score; and this may require interpretation in terms of some genuinely independent but rather weak second axis of variation. Very often in such cases, the third axis turns out to be more informative than the second and, although partially dependent on the first two axes taken together, may be nearly independent of the first axis taken alone.

## CONCLUSIONS

In most practical applications, reciprocal averaging gives stand ordinations which are similar to those derived by principal components analysis of standardized data. As a general method for use in phytosociological contexts it is preferable because it generates good simultaneous species ordinations. The rationale of the method is close to that of gradient analysis, so that it is more suitable than principal components analysis for displaying strong floristic gradients.

## ACKNOWLEDGMENTS

I thank various people for help with this work. My largest debt is to my supervisor, Professor P. Greig-Smith, whose encouragement and criticism have been invaluable. Dr J. C. Gower of Rothamsted Experimental Station made important criticisms of an earlier draft, and drew my attention to the work of Benzécri. Dr J. P. Nakache of Paris very kindly sent me a draft of his monograph '*Analyse Factorielle des Correspondances*'. T. P. T. Williams helped with the computer programming, as did M. D. Swaine, who also discussed the method in detail and made many useful suggestions.

## SUMMARY

Principal components analysis has often been used to analyse multivariate data. An

alternative and rather similar method, developed in France under the title 'Analyse Factorielle des Correspondances' and here termed reciprocal averaging, is advocated as being more satisfactory in ecological contexts. Its particular strength is that it simultaneously ordinates both the species and the stands. The derived stand ordinations are often similar to those obtained by principal components analysis of standardized data. The theory of the method is presented in a mathematical appendix and its relation to principal components analysis is discussed. An algorithm suitable for deriving reciprocal averaging ordinations is briefly described.

## REFERENCES

**Benzécri, J. P. (1969).** Statistical analysis as a tool to make patterns emerge from data. *Methodologies of Pattern Recognition* (Ed. by S. Watanabe), pp. 35–60. Academic Press, New York.

**Clapham, A. R., Tutin, T. G. & Warburg, E. F. (1962).** *Flora of the British Isles*, 2nd edn. Cambridge University Press, London.

**Clint, M. & Jennings, A. (1970).** The evaluation of eigenvalues and eigenvectors of real symmetric matrices by simultaneous iteration. *Comput. J.* **13,** 76–80.

**Escofier-Cordier, B. (1969).** L'analyse factorielle des correspondances. *Cah. Bur. univ. Rech. opér. Univ. Paris*, **13.**

**Goodall, D. W. (1954).** Objective methods for the classification of vegetation. III. An essay in the use of factor analysis. *Aust. J. Bot.* **2,** 304–24.

**Gower, J. C. (1966).** Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, **53,** 325–38.

**Greig-Smith, P. (1964).** *Quantitative Plant Ecology*, 2nd edn. Butterworth, London.

**Harberd, D. J. (1962).** Application of a multivariate technique to ecological survey. *J. Ecol.* **50,** 1–17.

**Jennings, A. (1967).** A direct iteration method for obtaining the latent roots and vectors of a symmetric matrix. *Proc. Camb. phil. Soc. math. phys. Sci.* **63,** 755–65.

**Noy-Meir, I. (1970).** *Component analysis of semi-arid vegetation in southeastern Australia.* Ph.D. thesis, Australian National University.

**Noy-Meir, I. & Austin, M. P. (1970).** Principal component ordination and simulated vegetational data. *Ecology*, **51,** 551–2.

**Orloci, L. (1966).** Geometric models in ecology. I. The theory and application of some ordination methods. *J. Ecol.* **54,** 193–215.

**Poore, M. E. D. (1962).** The method of successive approximation in descriptive ecology. *Adv. ecol. Res.* **1,** 35–68.

**Swan, J. M. A. (1970).** An examination of some ordination problems by use of simulated vegatational data. *Ecology*, **51,** 89–102.

**Whittaker, R. H. (1956).** Vegetation of the Great Smoky Mountains. *Ecol. Monogr.* **26,** 1–80.

**Whittaker, R. H. (1967).** Gradient analysis of vegetation. *Biol. Rev.* **42,** 207–64.

(*Received* 10 *May* 1972)

## NOTE ADDED SEPTEMBER 1972

Dr M. B. Dale has drawn my attention to a recent paper on the above topic (Hatheway 1971). Hatheway's treatment is similar to the foregoing. He does not suggest a name for the method, so 'Reciprocal Averaging' remains the correct designation.

## REFERENCE

**Hatheway, W. H. (1971).** Contingency-table analysis of rain forest vegetation. *Statistical Ecology*, Vol. 3. *Many Species Populations, Ecosystems, and Systems Analysis* (Ed. by G. P. Patil, E. C. Pielou & W. E. Waters), pp. 271–313. Pennsylvania State University Press, University Park, Pennsylvania.

## APPENDIX 1

A mathematical formulation of reciprocal averaging is presented here together with an explanation of how it relates to principal components. The algorithm used to perform the calculations is described. A worked example of how to do the calculations by hand is given in Appendix 2.

Presence–absence data may be represented by an incidence matrix

$$\mathbf{A} = [a_{ij}] \quad (i = 1, \ldots, m; j = 1, \ldots, n)$$

of 0s and 1s. Rows represent species and columns represent stands.
Let

$$r_i = \sum_j a_{ij}$$

$$c_j = \sum_i a_{ij}$$

be the row and column totals respectively. The reciprocal averaging procedure can then be represented as

$$x_i = \sum_j a_{ij} y_j / r_i \quad (i = 1, \ldots, m); \tag{1}$$

$$y_j = \sum_i a_{ij} x_i / c_j \quad (j = 1, \ldots, n). \tag{2}$$

In matrix notation this is

$$\mathbf{x} = \mathbf{R}^{-1}\mathbf{A}\mathbf{y};$$
$$\mathbf{y} = \mathbf{C}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{x};$$

where $\mathbf{x}$ and $\mathbf{y}$ are the column vectors formed from the $x_i$s and the $y_j$s respectively, and

$$\mathbf{R} = \mathbf{diag}\,(r_i);$$
$$\mathbf{C} = \mathbf{diag}\,(c_j).$$

Looked at from the point of view of $x$, the iteration can now be represented as

$$\mathbf{x}' = \mathbf{R}^{-1}\mathbf{A}\mathbf{C}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{x} \tag{3}$$

where $\mathbf{x}'$ is the new value of the $x_i$s and $\mathbf{x}$ is the old value. Eigenvectors of (3) are 'solutions' of the reciprocal averaging problem. There is a slight divergence between this formulation and that given informally in the text, in that here there is no rescaling. This means that the vector

$$\mathbf{u} = [1,1, \ldots, 1]^{\mathrm{T}},$$

corresponding to the eigenvalue 1 is a trivial 'solution' of the problem. 1 is the largest eigenvalue of (3), since it is not possible to exceed the limits of the original $x$s by a two-way averaging procedure. Computationally, the simplest way to find the eigenvectors of eqn (3) is to transform to the variable

$$\boldsymbol{\xi} = \mathbf{R}^{\frac{1}{2}}\mathbf{x},$$

where

$$\mathbf{R}^{\frac{1}{2}} = \mathbf{diag}\,(\sqrt{r_i}).$$

Eqn (3) then becomes

$$\boldsymbol{\xi}' = \mathbf{R}^{-\frac{1}{2}}\mathbf{A}\mathbf{C}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-\frac{1}{2}}\boldsymbol{\xi} \tag{4}$$

the matrix preceding $\boldsymbol{\xi}$ on the right hand side of (4) is now symmetric; and being of the form $\mathbf{B}\mathbf{B}^{\mathrm{T}}$, where

$$\mathbf{B} = \mathbf{R}^{-\frac{1}{2}}\mathbf{A}\mathbf{C}^{-\frac{1}{2}},$$

it has a complete set of eigenvectors, and all the eigenvalues are non-negative.

To see how principal components relates to reciprocal averaging, it is necessary to represent it as a reciprocal scoring system analogous to eqns (1) and (2). Unstandardized principal components may be represented as

$$x_i = \sum_j (a_{ij} - a_{i.}) y_j \quad (i = 1, \ldots, m); \tag{5}$$

$$y_j = \sum_i (a_{ij} - a_{i.}) x_i \quad (j = 1, \ldots, n); \tag{6}$$

where $a_{i.}$ is the $i^{\text{th}}$ row-mean, $x_i$ is the loading of the $i^{\text{th}}$ species and $y_j$ is the score of the $j^{\text{th}}$ stand. In matrix notation this reduces to

$$\mathbf{x}' = \mathbf{S}\mathbf{x} \tag{7}$$

where $\mathbf{S}$ is the matrix of deviation sums of squares and products of the rows of $\mathbf{A}$, and $\mathbf{x}'$ is the new value of $\mathbf{x}$. Now eqns (5) and (6) can be rewritten

$$x_i = \sum_j (a_{ij} - a_{i.})\, y_j \quad (i = 1, \ldots, m); \tag{8}$$

$$y_j = \sum_i a_{ij} x_i \quad (j = 1, \ldots, n); \tag{9}$$

which has the effect merely of shifting the origin of co-ordinates for the stand scores $y_j$. Eqn (9) shows an important difference between principal components and reciprocal averaging. This is that the stand scores are with principal components the sum of a set of species scores (loadings) for those species which occur in the stand, while with reciprocal averaging an average of a set of species scores is used. This means that principal components is likely to emphasize species-rich stands and to give species-poor ones a rather intermediate position. Reciprocal averaging corrects for species-richness. Another important feature is shown by eqn (8). This is that a much greater weighting is given to species of intermediate commonness than to species which are either very rare or very common. Thus with very rare or very common species, $(a_{ij} - a_{i.})$ is very seldom appreciably different from zero, while with intermediate species it is always different. Hence, principal components using unstandardized data pays very little attention to rare species; not only do they occur seldom, but they will probably have small loadings.

Principal components using standardized data, on the other hand, puts the species on a more equal footing. Let $S_i$ be the sum of squared deviations from the mean of the $i^{\text{th}}$ row. Then the reciprocal scoring system can be written

$$x_i = \sum_j (a_{ij} - a_{i.}) y_j / \sqrt{S_i} \quad (i = 1, \ldots, m); \tag{10}$$

$$y_j = \sum_i (a_{ij} - a_{i.}) x_i / \sqrt{S_i} \quad (j = 1, \ldots, n). \tag{11}$$

The stand scores $y_j$ are centred to have zero mean, and the species scores $x_i$ are the usual species loadings. Now the reciprocal scoring system,

$$x_i = \sum_j (a_{ij} - a_{i.}) y_j / S_i \quad (i = 1, \ldots, m); \tag{12}$$

$$y_j = \sum_i a_{ij} x_i \quad (j = 1, \ldots, n); \tag{13}$$

determines essentially the same problem, but generates uncentred stand scores $y_j$ and

differently normalized species scores $x_i$. In accordance with the notation used above, let $r_i$ be the row-total of the $i^{\text{th}}$ row. Then with presence–absence data

$$S_i = r_i(m - r_i)/m;$$

so that for relatively uncommon species—i.e. $r_i$ less than about $\frac{1}{4}m$—the sums of squared deviations $S_i$ are approximately proportional to the row-totals $r_i$. This largely explains why reciprocal averaging gives ordinations which resemble those derived from principal components using standardized data. Provided that most of the species are not excessively common, and provided that the stands are not too variable in species-richness, eqns (12) and (13) are roughly proportional to eqns (1) and (2).

It remains to describe the algorithm which has been used to compute reciprocal averaging ordinations at Bangor. It is a direct iteration algorithm similar to that of Jennings (1967) and Clint & Jennings (1970) but simpler. Eqn (4) shows that the computation amounts to extracting the eigenvectors of a symmetric matrix. Eqns (1) ad (2) show that provided that a direct iteration algorithm is used, there is no need ever to evaluate the matrix explicitly; nothing need be stored in the computer except the raw data and a few vectors. Sparse presence–absence data can be stored efficiently, permitting a great saving of space. All the algorithm does is to take a trial vector $\mathbf{x}_1$ and to iterate three times, deriving

$$\mathbf{x}_2 = \mathbf{L}\mathbf{x}_1, \mathbf{x}_3 = \mathbf{L}\mathbf{x}_2, \mathbf{x}_4 = \mathbf{L}\mathbf{x}_3,$$

where $\mathbf{L}$ is our symmetric matrix. The Schmidt orthogonalization procedure is then applied to $\mathbf{x}_4$, $\mathbf{x}_3$ and $\mathbf{x}_2$ so as to derive an orthonormal triad. The co-ordinates of $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ and $\mathbf{x}_4$ with respect to this triad are used to obtain a representation of $\mathbf{L}$ restricted to the space spanned by the triad. The representation is a $3 \times 3$ symmetric matrix and the problem is solved explicitly in the restricted space to obtain a new trial vector $\mathbf{x}_1$. Using the tolerance criterion

$$\| \mathbf{L}\mathbf{x} - \hat{\lambda}\mathbf{x} \| < \frac{1}{1000} \| \mathbf{L}\mathbf{x} \|,$$

where

$$\hat{\lambda} = \| \mathbf{L}\mathbf{x} \| / \| \mathbf{x} \|$$

is an estimate of the eigenvalue, it usually takes about thirty-five iterations of $\mathbf{L}$ (i.e. seventy passes of the data) to extract three eigenvectors.

## APPENDIX 2

A worked example of the calculations of reciprocal averaging is given here.

| | (i) | (ii) | (iii) | (iv) | (v) | (vi) | (vii) | (viii) | (R) | (1) | (2) | (2a) | (3) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (i) | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 4 | 100 | 52·5 | 55 | 44·3 |
| (ii) | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 4 | 0 | 37·5 | 0 | 36·2 |
| (iii) | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 4 | 100 | 65·0 | 100 | 63·4 |
| (iv) | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 6 | 0 | 43·3 | 21 | 39·3 |
| (v) | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 4 | 100 | 56·7 | 70 | 47·2 |
| (vi) | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 46·7 | 33 | 46·0 |
| (C) | 5 | 4 | 2 | 3 | 3 | 2 | 1 | 4 | 24 | | | | |
| (1) | 60·0 | 50·0 | 0·0 | 66·7 | 33·3 | 50·0 | 100·0 | 50·0 | | | | | |
| (2) | 55·8 | 47·8 | 10·5 | 48·7 | 36·3 | 50·0 | 100·0 | 36·5 | | | | | |
| ..... | | | | | | | | | | | | | |
| (11) | 31·8 | 56·5 | 48·4 | 19·7 | 10·0 | 86·0 | 100·0 | 32·7 | | | | | |
| (11a) | 24 | 52 | 42 | 11 | 0 | 84 | 100 | 25 | | | | | |

The calculations are represented schematically in the foregoing table. The data-matrix is given in the top left-hand corner, and (R) and (C) are the row (species) and column (stand) totals respectively. Column (1) is an arbitrarily chosen set of starting scores. In practice these should be chosen to reflect what is suspected of being the main gradient. A good choice will much reduce the amount of calculation required.

Row (1) is derived from column (1) by averaging. Thus the entry in row (1) column (v) is 33·3, being the average of 100, 0 and 0, which are the scores in column (1) corresponding to the non-zero entries of column (v). Column (2) is defined similarly. Thus the entry in column (2) row (i) is the average of 60·0, 66·7, 33·3 and 50·0—these being the scores in row (1) corresponding to the non-zero entries of row (i). Column (2a) is derived from column (2) by rescaling, and is given by the formula:

$$\text{column (2a)} = 100 \times (\text{column (2)} - 37{\cdot}5)/27{\cdot}5.$$

This ensures that the range of column (2a) is 0 to 100, since 27·5 is the range of column (2) and 37·5 is its minimum value. By continuing in this manner, the following sequence of species (row) scores is obtained.

| (1) | (2a) | (3a) | (4a) | (5a) | (6a) | (7a) | (8a) | (9a) | (10a) | (11a) | (12a) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 55 | 30 | 8 | 0 | 0 | 0 | 2 | 3 | 4 | 5 | 5 | 23·5 |
| 0 | 0 | 0 | 6 | 23 | 40 | 52 | 60 | 66 | 70 | 72 | 73 | 55·9 |
| 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 68·6 |
| 0 | 21 | 11 | 0 | 3 | 10 | 14 | 18 | 21 | 23 | 25 | 26 | 33·2 |
| 100 | 70 | 40 | 18 | 12 | 16 | 19 | 24 | 26 | 28 | 29 | 30 | 35·1 |
| 0 | 33 | 36 | 26 | 16 | 10 | 5 | 0 | 0 | 0 | 0 | 0 | 20·9 |

It takes eleven iterations to reach stability of the scores, but this is the result of making a bad initial choice. Three or four iterations should normally suffice if a good initial choice is made. The final stand (column) scores are derived by rescaling row (11) to form row (11a) as indicated in the original table. The eigenvalue (latent root) corresponding to the first axis is a measure of how much the range of the scores contracts in one iteration. The range of column (12) (shown after column (12a)) is 47·7, and it is derived from column (11a) which has a range of 100. Hence the estimate of the eigenvalue is 0·477. These calculations should be done with the data on one piece of quadrille paper and the scores on another, matching the two side by side.

When the first axis has been obtained, the second is considered. A good starting point for the scores of the second axis is obtained by using a set of scores which were fairly near to the final ones for the first axis. In this case column (8a) is used. Before iteration, these scores have to be adjusted by subtracting a multiple of the final first axis. This multiple is estimated as follows.

| $z$ | R | R$z$ | R$\bar{z}$ | $x$ | $y$ | (13) | (13a) | (14a) | (15a) |
|---|---|---|---|---|---|---|---|---|---|
| 5 | 4 | 20 | 165 | −145 | 2 | −3·0 | 71 | 62 | 59 |
| 73 | 4 | 292 | 165 | 127 | 60 | −12·4 | 0 | 0 | 0 |
| 100 | 4 | 400 | 165 | 235 | 100 | 0·8 | 100 | 94 | 89 |
| 26 | 6 | 156 | 247 | −91 | 18 | −7·8 | 35 | 34 | 33 |
| 30 | 4 | 120 | 165 | −45 | 24 | −5·8 | 50 | 45 | 41 |
| 0 | 2 | 0 | 82 | −82 | 0 | 0 | 94 | 100 | 100 |
| | 24 | 988 | | −1 | | | | | |

The column $z$ is the first axis; R is the row totals and $y$ is the set of scores to be adjusted (in this case equal to column (8a)). Multiply R by $z$ to form R$z$. Form $\bar{z}$ a weighted mean value of $z$ by taking

$$\bar{z} = \sum Rz / \sum R.$$

In this case,

$$\bar{z} = 988/24 = 41{\cdot}17.$$

Form a column $R\bar{z}$ by multiplying R by $\bar{z}$; then subtract $R\bar{z}$ from $Rz$ to derive $x = Rz - R\bar{z}$. (A check at this point is that, apart from round-off error, $x$ should sum to zero.) The multiple of $z$ to be subtracted from $y$ is given by

$$\sum xy / \sum xz,$$

which in this case is 0·992. Column (13) is therefore $y - 0{\cdot}992z$, and after rescaling to derive column (13a) the iterations are continued in the usual way. The first axis will slowly re-establish itself if the appropriate multiple of $z$ (i.e. $\Sigma xy'/\Sigma xz$) is not at intervals subtracted from subsequent scores $y'$; but this need not be done very often. The column (15a) derived after two iterations from (13a) has not been further corrected for the first axis, but it may nonetheless be taken as a reasonable estimate of the second. The estimate of the second eigenvalue, derived from column (15) (not shown), is 0·305.

These calculations are rather laborious. They would be worth the trouble if a good ordination were required in the absence of a computer.