

Correspondence analysis of transition matrices, with special attention to missing entries and asymmetry

PETER G. M. VAN DER HEIJDEN*†, HAN DE VRIES‡§ &

JAN A. R. A. M. VAN HOOFF‡§

**Department of Psychometrics and Research Methods, Leiden University, Postbus 9555, 2300 RB Leiden, The Netherlands*

‡*Laboratory of Comparative Physiology, University of Utrecht, Postbus 80.086, 3508TB Utrecht, The Netherlands*

Abstract. The use of correspondence analysis is proposed for the analysis of transition matrices. The practice of performing two factor analyses on a transition matrix, one for the rows and one for the columns, is better replaced by performing one correspondence analysis on the matrix. The chi-squared distance used by correspondence analysis is methodologically as well as ethologically a more valid association (or similarity) measure than the correlation coefficient used by factor analysis. Correspondence analysis is complementary to loglinear models. This is demonstrated, using a transition matrix, for the independence, the quasi-independence and the quasi-symmetry models. The method is illustrated with an analysis of Wiepkema's (1961 *Arch. Neerl. Zool.*, **14**, 103–199) data on the reproductive behaviour of bitterling, *Rhodeus amarus*.

The sequences of behavioural elements shown by individuals of a species have often been analysed to detect the organizational principles according to which the behaviour is structured. Another, less far reaching objective of the analysis of sequences is to describe the associations between every pair of behavioural elements. The behaviour of an animal is observed through time and a sequence of behavioural elements is scored. When durations of activities are not considered, the sequence can be analysed by summarizing the transitions in a two-way transition matrix where the row variable has the preceding activities as categories, and the column variable the following activities. This matrix can be used in discrete-state Markov-chain modelling, where the relevant parameters are the transition probabilities (see Slater 1973; Bishop et al. 1975; Cane 1978; Castellan 1979).

Another approach to investigating the transitions between activities, either within the same animal or between different animals, is based on information theory (e.g. Losey 1978). When the main focus of study is on the communication and information

transmission between animals, information theoretical measures like uncertainty, transmission and transmission efficiency can be used to describe the rather global associations between the sequences of behaviour of the two interactants. The measures can also be used to study the global sequential dependencies in individual sequences (van den Bercken & Cools 1980a, b).

More recently, much attention has been given to methods that also take into account the durations of activities. The work of Haccou and others (e.g. Metz et al. 1983; Haccou 1986, 1987) has shown that neglect of the durations can lead to false conclusions. This is the case when an element occurs with different durations and when long and short variants of this element are functionally different, i.e. when these are preceded and/or followed by different activities. In this case it is necessary to distinguish two or more new elements on the basis of their duration. This approach was followed by Thon & Chabaud (1986). Haccou & Metz (*loc. cit.*) have tried to analyse the sequential and temporal structure of behaviour by continuous-time Markov-chain models and generalizations thereof. These models seem to be very useful once the exploratory phase of research is over.

There are some cases for which it is justifiable (or unavoidable) to pay less attention to the time-aspect in the analysis of behavioural structure: for instance,

†Present address: Department of Development and Socialisation, Section Methods, University of Utrecht, Postbus 80140, 3508 TC Utrecht, The Netherlands.

§Present address: Department of Ethology and Socioecology, University of Utrecht, Postbus 80086, 3508 TB Utrecht, The Netherlands.

when one wants to conduct an exploratory investigation of the possible relations between a large number of activities shown by one or more individuals of a species. In this case one is often not interested in a detailed description of the ongoing behaviour, but in a global picture of the relations between the activities; one is searching for a classification (or categorization) of the (occasionally large number of) elements. For this purpose methods have been used that belong to the group of exploratory pattern detection methods, such as cluster analysis methods, principal components analysis (i.e. one form of factor analysis) and multidimensional scaling methods. The objective is to group together elements to the extent that these are related to one another in some respect, so as to form a more compact ethogram of lumped activities. Such empirically based categories may subsequently be used to describe complex social interactions. Thus in his study of polyadic strategies of chimpanzees, *Pan troglodytes*, de Waal (1978) applied the categories derived by van Hooff (1973).

Here we introduce correspondence analysis as a method especially suitable for the exploratory analysis of multidimensional structure in two-dimensional contingency tables, such as transition matrices. The similarity measure in correspondence analysis, the chi-squared distance, is a valid measure for the similarity between activities. This measure is preferable over the correlation measure as used in the conventionally applied method of factor analysis. Correspondence analysis is also closely connected to loglinear models, which have proven their usefulness in the analysis of transition matrices.

In ethology, correspondence analysis has been applied by Boy & Duncan (1979) and Duncan (1980, 1985) for the analysis of time-budgets. These are coded into matrices with individual animals in the rows and activities in the columns; in each cell of the matrix the proportion of time is found that an individual spent on a specific activity. Correspondence analysis of transition matrices has been explored by van der Heijden (1986, 1987) and Thon & Chabaud (1986). Traditionally correspondence analysis was suited only for ordinary two-way contingency tables, but it has been recently extended to contingency tables in which some cells might be uninteresting (such as the diagonal cells in some transition matrices) and to the study of asymmetry in square contingency tables (see van der Heijden et al. 1989). We apply these new methods in this

paper. But before doing so, we first discuss some related methods.

ANALYSIS OF TRANSITION MATRICES

A transition matrix is obtained in the course of investigating a sequence of activities. Such a sequence is then summarized in a two-way frequency matrix, with the preceding activities in the rows and the following activities in the columns. Let us assume that I different activities can be distinguished. Then a frequency f_{ij} denotes the number of times that category i is followed by category j . These frequencies f_{ij} can be collected into a matrix F having I rows and I columns. The matrix thus obtained has the special property that margins f_{i+} and f_{+i} of corresponding row and column category i , differ at most by 1 for one recorded sequence. This is because every following activity becomes a preceding activity in the next transition, except for the last element of the sequence.

We shall not give a thorough discussion of the 'correctness' of a reduction of a sequence to a two-way matrix F (compare Voeten 1985). Here we only emphasize that an implicit requirement for the correctness of such a reduction is that the transition process is stationary, i.e. that the transition probabilities $p(j|i)$ do not change over time. Here $p(j|i)$ denotes the conditional probability that, given that the animal is in activity i , activity j will follow.

We can also obtain transition matrices if we deal with a number of sequences. In this case we can work in the same way as discussed for one sequence. We can derive a transition matrix for each sequence, and sum these matrices. In this case f_{i+} may come to differ considerably from f_{+i} . When summing the matrices, we not only make the implicit requirement of stationarity, but also of homogeneity of sequences, i.e. we assume that the transition probabilities $p(j|i)$ are the same in each sequence (see also Castellán 1979). Preferably this assumption should be tested, for example with standard contingency table methods like loglinear analysis (Colgan & Smith 1978). However, a practical problem is that very often one of the conditions for these tests is not fulfilled, namely that the cell frequencies should be rather large (e.g. a rule of thumb is that they should be larger than 5). In fact, the occurrence of low cell frequencies is an important reason for summing matrices: in this way the cell entries become larger, and thus more stable.

Types of Association

With regard to the (first-order) sequential structure of behaviour three different types of association between each pair of activities can be discerned. For each type, a measure, expressing the strength of association between each pair, can be calculated from the transition matrix. The first type of association is the (directed) sequential association between each two elements, which indicates the degree to which one activity tends to follow the other activity. This measure is in general asymmetric. A second type of association is known as the precedent mutual replaceability, indicating the degree to which two activities have the same precursors (Douglas & Tweed 1979). Using a more interpretational term this type of association can also be called ‘internal causal association’. This association type is reflected by the similarity between each pair of columns of the transition matrix. A third type of association is the subsequent mutual replaceability, indicating the extent to which two activities have the same successors (ibid.), which can be interpreted as the ‘internal functional association’. It is reflected by the similarity between each pair of rows of the transition matrix. We now discuss some ways of analysing transition matrices, and show how the analyses deal with these types of association.

Loglinear Modelling

For ease of notation we now switch from frequencies f_{ij} to proportions $p_{ij} = f_{ij}/n$, where $n = f_{++}$.

A starting point in the analysis of transition matrices is often the comparison of observed proportions and proportions expected under the assumption that independence holds, i.e. when the occurrence of categories at time $t+1$ does not depend on the categories at time t . If the observed proportions are independent, the activity at time t does not predict the activity at time $t+1$.

We now define independence more formally. Let expected proportions be denoted as π_{ij} , and estimates of expected proportions by $\hat{\pi}_{ij}$. Under the independence model expected proportions π_{ij} can be estimated as

$$\hat{\pi}_{ij} = p_{i+}p_{+j} \tag{1}$$

The difference between observed proportions p_{ij} and estimates of expected proportions $\hat{\pi}_{ij}$ is tested with the chi-squared statistic

$$\chi^2 = n \sum_i \sum_j (p_{ij} - \hat{\pi}_{ij})^2 / \hat{\pi}_{ij} \tag{2}$$

with $df = (I - 1)^2$. If the chi-squared is not large enough given the number of degrees of freedom, we have no reason to reject the hypothesis of independence.

When the independence model fits badly, either a less restrictive model can be fitted to all cells (see below), or (standardized) residuals for individual cells can be studied for significant departure from independence. Standardized residuals are the quantities

$$n(p_{ij} - \hat{\pi}_{ij}) / \sqrt{\hat{\pi}_{ij}} \tag{3}$$

Quantities (3) are the square roots of the contributions to the chi-squared (2). They show whether behaviour i is followed by behaviour j more or less often than expected on the basis of independence, i.e. they show the strength of the (directed) sequential association between i and j .

In sequential analysis all standardized residuals (3) are often studied individually. However, this cannot be recommended, because the residuals are not independent statistically (Fagen & Mankovich 1980). This is easily seen from the fact that the raw residuals $p_{ij} - \hat{\pi}_{ij}$ add up to 0 for each row and column, indicating that in each row and column one entry is redundant. Another reason for not recommending the study of individual cells is that one may lose sight of the relationship that exists between significant cells or categories.

In later sections we discuss models that are less restricted than the independence model. For an introduction to the class of loglinear models, of which all the models discussed in this paper are special cases, see Colgan & Smith (1978).

Factor Analysis

A particular approach to revealing the interrelationships between the behavioural elements is to transform the transition matrix, and to factor analyse it subsequently. Examples and methodological discussions can be found in, among others, van Hooff (1970, 1973, 1982), Burton (1970), Balthazart (1972, 1973), Slater & Ollason (1972), Slater (1973), Hoekstra & Reddingius (1981), Vodegel (1978), Baerends (1984) and Loots (1985).

Several (slightly different) procedures have been followed. Wiepkema (1961) was the first to apply factor analysis to transition matrices in a study of the reproductive behaviour of the bitterling, *Rhodeus amarus*. He transformed the transition

matrix to a matrix with elements $p_{ij}/\hat{\pi}_{ij}$, i.e. observed proportions divided by expected proportions under the independence model (1). Subsequently Spearman rank correlations were computed between the rows, and between the columns, yielding two correlation matrices. On these matrices factor analyses were performed.

Others first constructed a symmetric transition matrix by adding all pairs of cells that lie symmetrically on each side of the diagonal. Sometimes standardized residuals (3) were taken instead of values $p_{ij}/\hat{\pi}_{ij}$. A third difference in the applications is that Kendall rank correlations or Pearson product-moment correlations are calculated instead of Spearman rank correlations.

In these cases factor analysis (in fact principal components analysis) was used for the detection of groups of related behaviour patterns by inspecting and interpreting the matrix of component loadings. Hoekstra & Reddingius (1981) have argued that the usefulness of principal components analysis in this context is due to the close connection that exists between it and principal coordinates analysis (Gower 1966; see also Digby & Kempton 1987, page 89).

Principal components analysis is used as a pattern detection method using the matrix of correlations between the pairs of elements as a starting point. The correlation coefficients are used as measures of similarity (association) between the elements. The different steps made (diagonal defined or undefined, calculation of the discrepancy measure $p_{ij}/\hat{\pi}_{ij}$ or $(p_{ij} - \hat{\pi}_{ij})/\hat{\pi}_{ij}^{\frac{1}{2}}$, the use of Pearson, Spearman or Kendall correlation coefficient) of course influence the resulting solution of the principal components analysis. The large number of possibilities marks the absence of clear rationales, and makes this approach rather arbitrary. From a methodological point of view, taking a Pearson correlation as a measure to describe the association between two variables presupposes that these variables follow a bivariate normal distribution. However, on the contrary, there is no reason to assume this in the case of transition matrices. So we can only hope that the different steps made (see above) will result in a situation that is not too far from bivariate normality.

We believe, therefore, that factor analysis should no longer be used, now that methods are available that use association measures in a methodologically and ethologically more appropriate manner. The main objection against factor analysis is that

the use of correlation coefficients is difficult to justify ethologically and incorrect from a statistical point of view. One is never sure whether factor analysis of an arbitrary type of correlations will lead to the correct conclusions. Below we show that correspondence analysis uses a dissimilarity measure that is better motivated, both methodologically and ethologically. Correspondence analysis circumvents the problems and arbitrary choices involved in factor analysis.

CORRESPONDENCE ANALYSIS

Correspondence analysis is formally identical to a number of other data analysis techniques, such as canonical analysis of contingency tables and reciprocal averaging. Consequently, since each of these approaches focuses on specially selected aspects of the technique, there are many ways to describe correspondence analysis. We give a description of correspondence analysis with an emphasis on the geometrical properties of the technique, because these are the most often mentioned as an explanation for its popularity. For details, see Benzécri et al. (1973), Nishisato (1980), Gifi (1981), or Greenacre (1984). For computational details see the Appendix.

Correspondence analysis is a technique for finding a multi-dimensional representation of the dependence between the rows and columns of a two-way contingency table. This representation is found by allocating scores to the row and column categories. These scores can be normalized in such a way that distances between row points and/or between column points in Euclidean space are equal to so-called chi-squared distances.

Chi-squared distances can be computed between rows as well as between columns. We proceed by considering chi-squared distances between rows. These distances are computed on the profiles of the rows of a matrix, where the profile of row i is the vector of conditional proportions $p(j|i)$ which sum to 1: $\sum_j p(j|i) = 1$. The chi-squared distance between rows i and i' is defined as

$$\delta^2(i, i') = \sum_j \frac{[p(j|i) - p(j|i')]^2}{p_{+j}} \quad (4)$$

Equation (4) shows that $\delta^2(i, i')$ is a measure for the differences between the profiles of row i and i' : for each column j , the difference between the conditional row proportions is squared. When i and i' have the same profile, so that $p(j|i) = p(j|i')$ for each

j , the result is that $\delta^2(i,i')=0$. The differences between the conditional row proportions of i and i' are corrected for marginal column proportions p_{+j} . The profile of marginal column proportions p_{+j} can also be interpreted as the mean row profile.

We can represent the I row profiles graphically as I points, where Euclidean distances between the points are equal to chi-squared distances. The mean row profile lies in the origin: $\sum_i (p_{i+})r_{ia} = 0$, where r_{ia} is the coordinate of row point i on dimension a . The weighted variance of the coordinates is defined to be equal to $\sum_i (p_{i+})r_{ia}^2 = \lambda_a^2$. The configuration of I row points is located in a Euclidean space of dimension $(I-1)$. Correspondence analysis results in a solution in which this configuration of I points is rotated in such a way that the values $\lambda_1^2, \dots, \lambda_{I-1}^2$ are subsequently maximized. Note that rows with a higher marginal proportion p_{i+} play a more important role in this rotation. Only in the $(I-1)$ -dimensional space are Euclidean distances equal to chi-squared distances. But because the weighted variances λ_a^2 of the distances to the origin are maximized for subsequent dimensions, it is useful to study projections onto lower-dimensional spaces.

Up to now we have discussed only the situation for the row categories. Correspondence analysis is symmetric in the sense that similar results hold for the columns. So it is also possible to construct a $(J-1)$ -dimensional space in which the mean column profile is placed in the origin.

An important property of correspondence analysis is that the solution for the row points is related to the solution for the column points. They are related in the following way. First, weighted variances λ_a^2 of the two solutions are equal for each dimension a . Furthermore, the row coordinates and column coordinates are related as follows

$$c_{ja} = (\lambda_a^{-1}) \sum_i p(i|j)r_{ia} \tag{5a}$$

$$r_{ia} = (\lambda_a^{-1}) \sum_j p(j|i)c_{ja} \tag{5b}$$

where r_{ia} (c_{ja}) is the coordinate for category point i (j) on dimension a . So, using the column profiles, column points can be computed as the weighted averages of the row points, and using the row profiles, row points can be computed as the weighted averages of the column points (multiplied by the constant λ_a^{-1}). It follows that both the row and the column space have the same dimensionality, equal to minimum $(I-1, J-1)$. Equations (5a, b) are known as the 'transition equations', because they indicate how to go from the solution for the row

points to the solution for the column points and vice versa. These transition equations define the rationale for an approach called 'reciprocal averaging', a technique well known in ecology (see Digby & Kempton 1987), and formally identical to correspondence analysis.

There is an important relation between the chi-squared for testing independence and the squared singular values λ_a^2

$$\sum_a \lambda_a^2 = \chi^2/n \tag{6}$$

The quantity in (6) is often called the 'total inertia'. It shows that correspondence analysis decomposes the chi-squared value of a matrix. This value measures whether the observed values differ from independent values. The differences between observed and independent values, the residuals, are decomposed by correspondence analysis, into a number of dimensions. As we have already indicated, the decomposition of residuals makes sense only when these residuals are not merely a result of random variation from independence. We can test this with the chi-squared test (2).

Correspondence analysis is usually used as a technique for exploratory data analysis. Inferential properties are seldom used. An overview of these properties is given in Gifi (1981) and Greenacre (1984). Goodman (1986) has introduced a maximum likelihood version of correspondence analysis in which various aspects can be tested. We do not discuss this any further, first because we use correspondence analysis as an exploratory technique model for residual analysis, and second because one of the assumptions for these inferential properties is always that the observations in the contingency table are independent, and this assumption is most often violated in the context of transition matrices.

Interpretation

Clouds of points based on chi-squared distances (4) can be interpreted as follows: when two row points (or two column points) are near each other, their profiles are similar. When profiles differ to a large extent, the distance between the points is large. The profiles of both the marginal row and the marginal column proportions are placed in the origin. When the distance of a category point to the origin is small, the profile of this category point does not differ much from the mean profile. Note that chi-squared distances are defined only among the set of rows and among the set of columns, and not between rows and columns.

The distance of row i to column j can be interpreted with the transition equations (5a, b): roughly one can say that i and j will be near each other when row profile element $p(j|i) \gg p_{+j}$ (this implies that at the same time also column profile element $p(i|j) \gg p_{i+}$), and that i and j are far apart when $p(i|j) \ll p_{+j}$ (this implies that at the same time also $p(j|i) \ll p_{i+}$).

The importance of dimension a can be evaluated by the ratio of the inertia of dimension a and the total inertia: $\tau_a = \lambda_a^2 / \sum_a \lambda_a^2$. This quantity τ_a can be interpreted as the proportion of 'explained' inertia for dimension a , or the proportion of chi-squared that is decomposed in dimension a (compare 6). The rationale for deciding how many dimensions are to be interpreted is the same as in conventional methods of principal components analysis. Three criteria can be used: one can take the absolute value of τ_a as a criterion. One can also use the 'elbow criterion': when we construct a two-dimensional plot where we set out the eigenvalues vertically, and their number horizontally, we study only the dimensions corresponding with eigenvalues that do still differ to some extent for subsequent dimensions (i.e. only dimensions before the elbow are studied). A third criterion is simply the interpretability of dimensions: dimensions are studied that are interpretable. The less the dimensions studied explain from the total inertia, the more careful one should be in interpreting the solution, since much of the inertia will then be hidden in higher dimensions.

An Alternative to Factor Analysis

We propose the use of correspondence analysis in preference to factor analysis for the analysis of transition matrices. The rationale for applying correspondence analysis to transition matrices runs as follows.

We have seen that the profiles of conditional probabilities p_{ij}/p_{i+} are very important in the context of transition matrices: when two row profiles are the same, the two activities occupy the same place in the sequence with regard to the behaviour patterns that follow, i.e. they are 'subsequent mutually replaceable'. Likewise the two activities are 'precedent mutually replaceable' if the two column profiles are the same. In the case of subsequent mutual replaceability we can say that the two activities serve the same purpose within the sequence, while activities that are precedent mutually replaceable have the same cause within the sequence.

In correspondence analysis also, profiles represent one of the central concepts. If the data are perfectly independent, i.e. the behaviour at time t is independent of the behaviour at time $t + 1$, then the profiles are all identical. If independence does not give an adequate description of the data, it is interesting to study whether and how row (or column) profiles differ. This can be done by correspondence analysis. The chi-squared distances can be interpreted as a measure for the difference between rows.

If one is interested in subsequent mutual replaceability, then the correspondence analysis solution of chi-squared distances between row points should be studied; if one is interested in precedent mutual replaceability, then the solution of chi-squared distances between column points should be studied. In addition to the aspect of mutual replaceability the cloud of points offers insight into the (directed) sequential association (see above). For this, one has to consider the solution for the rows and the columns together. We conclude that the dissimilarity measures used in correspondence analysis correspond very nicely with the concepts of subsequent and precedent mutual replaceability, and the relation between the solution for the rows and the solution for the columns with (directed) sequential association. Therefore we have not only a methodological rationale for preferring this method over conventional methods of principal components analysis, but also an ethological one.

The chi-squared distance used in correspondence analysis is preferable to the correlation measure used in factor analysis. The methodological rationale for using correlations on (scaled or transformed) frequencies is formally incorrect, whereas such a rationale can easily be given for the chi-squared distance. Nevertheless, performing factor analysis on correlations derived from frequencies seems to work out fine in many cases. This is because the method is similar to correspondence analysis in that both techniques work with differences between rows or between columns of the transition matrix, and both techniques work with residuals from independence (scaled in some way). However, whereas a valid representation of the data is guaranteed by correspondence analysis, it can only be hoped for when one uses factor analysis.

In conclusion: correspondence analysis of a transition matrix is a formally correct procedure resulting in a picture in which all the theoretically

Table I. Wiepkema's data on courtship behaviour of the bitterling

	jk	tu	hb	chs	fl	qu	le	hdp	sk	sn	chf	ffl	Total
JK	654	128	172	56	27	25	1	28	0	46	14	18	1169
TU	101	132	62	27	5	1	1	11	0	8	5	9	362
HB	171	62	197	130	0	25	0	50	14	18	14	12	693
CHS	60	22	152	135	0	8	0	43	16	15	12	4	467
FL	19	2	0	0	419	19	0	2	0	17	5	11	494
QU	36	1	18	5	12	789	119	295	26	70	1	14	1386
LE	4	0	0	0	0	57	167	73	0	8	0	0	309
HDP	22	9	40	37	5	245	7	171	287	53	8	13	897
SK	3	2	7	38	0	120	8	134	19	28	4	0	363
SN	42	2	17	16	20	70	11	67	9	225	12	12	503
CHF	18	3	10	13	6	5	0	8	0	24	97	9	193
FFL	27	3	6	5	10	13	0	18	0	10	8	29	129
Total	1157	366	681	462	504	1377	314	900	371	522	180	131	6965

Rows are preceding behaviours, columns following behaviours. Abbreviations are described in the text.

relevant aspects of behavioural structure that are present in a transition matrix are represented in one integrative view.

Example

Examples of correspondence analysis of transition matrices can be found in van der Heijden (1986, 1987) and Thon & Chaboud (1986). Here we analyse Wiepkema's (1961) data as an illustration. The example deals with the reproductive behaviour of 13 male bitterlings. The activities are jerking (JK), turning beats (TU), head butting (HB), chasing (CHS), fleeing (FL), quivering (QU), leading (LE), head-down posture (HDP), skimming (SK), snapping (SN), chafing (CHF) and fin-flickering (FFL). The matrix is shown in Table I. The data from different individuals are pooled, since Wiepkema could not provide us with the unpooled data. Anyway, the many small frequencies would not allow us to test the assumption underlying pooling the matrices for the 13 bitterlings properly. Note also that almost all diagonal elements are prominent, compared with the off-diagonal elements, indicating that behavioural states are often followed by themselves.

Correspondence analysis describes the departure from independence, so we first have to check the possibility that independence gives an adequate description of the data. If this is the case, we do not need correspondence analysis to describe the data. However, the deviation from independence is highly significant: $\chi^2 = 16\,985$, $df = 121$.

Correspondence analysis might be useful here for the interpretation of deviations from independence. The chi-squared will be decomposed into a number of dimensions and thus the correspondence analysis approach stays close to the approach to test for independence.

The first five singular values λ_a (with proportion of chi-squared τ_a) are 0.698 (0.286), 0.630 (0.259), 0.292 (0.120), 0.268 (0.110) and 0.167 (0.068). A representation of the first two dimensions is shown in Fig. 1. The first dimension is dominated by the fact that fleeing is followed almost exclusively by itself. This frequency causes the (row and column) fleeing profiles to be very different from all other profiles, hence the distances to all other categories must be large. The second dimension can be interpreted as an agonistic-sexual component: turning beats, jerking, head butting and chasing are agonistic movements, while leading, quivering, head-down posture and skimming are considered sexual movements. Note that because of high diagonal frequencies all corresponding row and column activities fall approximately in the same position. The first two dimensions account for 55% of chi-squared ($\tau_1 + \tau_2 = 0.545$). The third dimension, not shown here, distinguishes mainly chafing, fin-flickering and snapping on the positive side from leading on the negative side. Furthermore, we find the row point for head-down posture near the column point for skimming, indicating that in the sexual activities specific orders can be found. We come back to this aspect below.

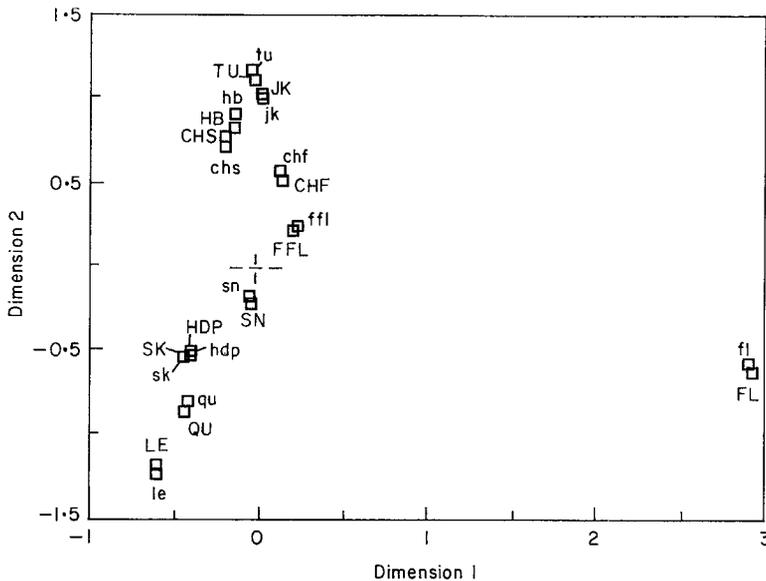


Figure 1. Correspondence analysis of data in Table I: dimensions 1 and 2. Capital letters represent preceding behaviours, small letters following behaviours. Distances between row points and between column points are small when their profiles are similar (see equation 4). The distance between row i and column j is small when $p(i|j) \gg p_{i+}$ (see equation 5). See text for further details.

Wiepkema (1961), on performing a factor analysis procedure with quadrimax rotation, concluded that three main factors underlie the activities of the bitterling. He interpreted these as follows: an agonistic factor, split into an aggressive tendency (head butting and chasing) and a flight tendency (turning beats, jerking and fleeing). Remarkably, the type element for the fleeing tendency is not found on the agonistic axis, but shares an independent axis with fin-flickering, chafing and snapping; Wiepkema interpreted this second factor as the non-reproductive factor. Finally there is a third factor, the sexual factor (skimming, head-down posture, quivering and leading). These three factors, found for the columns, accounted for 90% of the variance. As Wiepkema already noted himself, the principal components analysis solution for the rows gives surprisingly different results from that for the columns. In the solution for the rows, chasing has a high positive loading and fleeing and fin-flickering a high negative loading on the sexual dimension, in contrast to the solution for the columns in which these activities have low loadings. The element turning beats has a low loading on the sexual factor for the row solution and a high negative loading for the column solution. The anomalous position of chasing in the row solution

posed a particular problem for Wiepkema. The high positive correlation between chasing and skimming (as preceding elements) influences the positions of the other activities. This points to interactions between the causes underlying the aggressive and sexual tendencies, according to Wiepkema.

The second dimension of our analysis shows Wiepkema's factors ordered from aggressive, to non-reproductive, to sexual, where fleeing takes a special place. Compared with Wiepkema, our solution displays less of the total inertia χ^2/n , namely 77.5% in four dimensions. This is due in large part to the fact that the Spearman rank order correlations used by Wiepkema smooth out large irregularities in the data, because values $p_{ij}/\hat{\pi}_{ij}$ are replaced by rank order numbers. This leads to a neglect of the fact that, for example, fleeing is followed so often by itself. A more detailed comparison of Wiepkema's results with ours is given below, where the influence of the diagonal elements of the transition matrix is suppressed.

BEHAVIOURAL CHANGES

In the last section correspondence analysis was used to decompose the residuals from independence. Another independence model with which

transition matrices are often analysed is the quasi-independence model. In this model, for some cells in the matrix the expected proportions π_{ij} are set equal to observed proportions p_{ij} , and for other cells independence should hold. In the context of transition matrices this model is useful when substantive interest is restricted to the off-diagonal cells, particularly when the diagonal cell frequencies are either very high or extremely low. Extreme diagonal frequencies can be the result of the sampling strategy that is employed: for example, for sequences based on time sampling (e.g. the observed behaviour is recorded every 3 s), they can be very high. When only the transitions to other states are registered (sequences of unrepeatd events), the off-diagonal frequencies will be about the same, whereas the diagonal frequencies are zero (see also Chatfield & Lemon 1970; Chatfield 1973).

If for the diagonal cells expected proportions $\hat{\pi}_{ii}$ are taken as $\hat{\pi}_{ii} = p_{ii}$, the quasi-independence model can be written as

$$\hat{\pi}_{ij} = a_i b_j \text{ if } i \neq j \tag{7}$$

where a_i and b_j have to be estimated iteratively (see Bishop et al. 1975). Compared with the ordinary independence model (1) the expected proportions are determined by a product of a row and a column term, but now these terms are not equal to the observed marginal proportions. Model (7) can be tested with the chi-squared statistic (2), having $df = (I-1)(I-2) - 1$. The quasi-independence model can of course also be used when there are off-diagonal structural zeros, e.g. in Slater & Ollason (1972) where certain transitions are excluded by design.

Incomplete Tables

De Leeuw & van der Heijden (1988; see also van der Heijden et al. 1989) have shown that it is possible to decompose the residuals from quasi-independence with a program for ordinary correspondence analysis by analysing an adapted transition matrix. The transition matrix has to be adapted in the following way. Consider the matrix with observed proportions p_{ij} . We fit iteratively independent values to the cells we want to exclude from the analysis (here the diagonal cells), like

$$p_{ii}^{m+1} = p_{i+}^m p_{+i}^m \tag{8}$$

where p_{ij}^m is the estimate of p_{ij} at step m . The resulting matrix has independent values in the adapted cells and the original observed values in the non-

adapted cells. This resulting matrix is now analysed with correspondence analysis.

It can be shown that the relation between the chi-squared (for testing quasi-independence) and the sum of the squared singular values again holds as in (6) for independence. Because $\hat{\pi}_{ii} = p_{ii}$, the residuals are zero for these cells. Therefore, diagonal cells do not contribute to the total inertia $\sum_a \lambda_a^2$. Thus by using the quasi-independence model the influence of the diagonal cells is eliminated from the solution.

The interpretation of correspondence analysis decomposing residuals from quasi-independence is straightforward: it is just the ordinary correspondence analysis of the matrix with adapted values on the diagonal. Hence we can interpret correspondence analysis distances using the ordinary formula for chi-squared distances (4) and the transition equations (5a, b). We have only to keep in mind that the diagonal cells have independent values.

We have adapted correspondence analysis in such a way that it fits nicely into the ethological work of Lemon & Chatfield (1971) who used the quasi-independence model for transition matrices. Thus correspondence analysis still gives a good representation of the sequential association. The interpretation of chi-squared distances as a measure for mutual replaceability is slightly hampered though, because if we want to ignore the diagonal frequencies, the mutual replaceability for two activities i and j is best expressed by a similarity measure calculated across $I-2$ pairs of frequencies, i.e. not only excluding the pairs (i,i) , and (j,j) but also (i,j) and (j,i) .

Example

Examples of correspondence analysis of incomplete tables are shown in van der Heijden (1986, 1987), de Leeuw & van der Heijden (1988), and van der Heijden et al. (1989). In van der Heijden (1986) the data of Slater & Ollason (1972), which also have off-diagonal structural zeros, have been analysed. Here we discuss the Wiepkema data again.

We have seen that the diagonal elements dominate the solution of ordinary correspondence analysis. We now focus on off-diagonal cells. First we tested the quasi-independence model, with $p_{ii} = \pi_{ii}$. This model has a chi-squared value of 4156, with $df = 109$. We can see that the chi-squared is much lower than 16 985, the value obtained for the independence model. However, the current value of

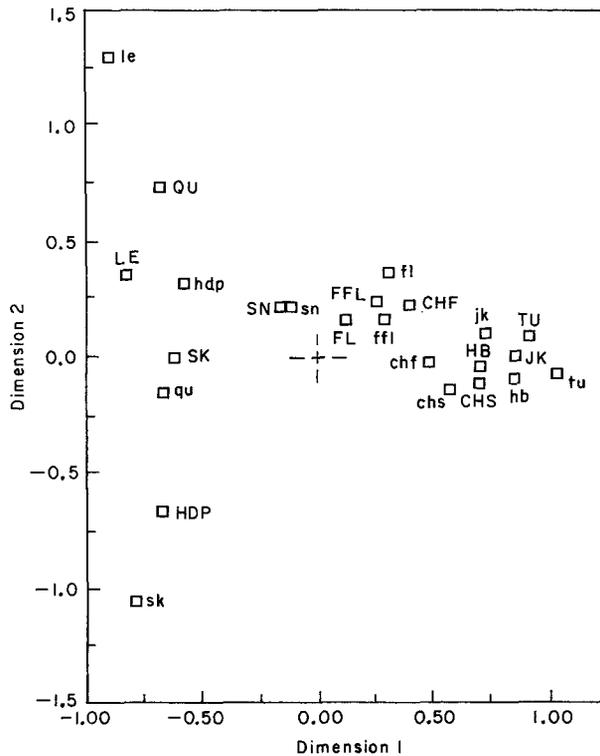


Figure 2. Correspondence analysis of data in Table I: decomposition of residuals from quasi-independence; dimensions 1 and 2. Capital letters represent preceding behaviours, small letters following behaviours. The solution can be interpreted as a correspondence analysis on the matrix with adjusted diagonal elements (see equation 8 and Fig. 1). See text for further details.

chi-squared is still significant, so a complementary correspondence analysis to study the structure in the departure from quasi-independence remains interesting.

The first five singular values of the correspondence analysis solution (with proportions of explained inertia) are 0.685 (0.507), 0.435 (0.204), 0.298 (0.096), 0.249 (0.067) and 0.218 (0.051). The first two dimensions explain 71% of chi-squared, which is much better than the result obtained in the previous analysis, where only 54.5% was decomposed. A plot is shown in Fig. 2. On the first dimension we find the agonistic activities jerking, turning beats, head butting and chasing on the right; in the upper part near the centre are the so-called 'non-reproductive activities' snapping, fleeing, fin-flickering and chafing; and the sexual activities are on the left. The second dimension is dominated by the sexual behaviours, head-down posture, skimming, quivering and leading. This dimension also reveals a strong sequential asymmetry by head-

down posture being followed by skimming (HDP row point close to SK column point in Fig. 2), and, similarly, quivering followed by leading. According to the elbow criterion, we can neglect the third dimension.

We now compare Wiepkema's results with ours, both the correspondence analysis performed on the complete transition matrix (see above; call this 'CA-Ind', from decomposition of residuals from independence) and the correspondence analysis performed on the transition matrix with imputed values on the diagonal (call this 'CA-QInd', from decomposition of residuals from quasi-independence).

Wiepkema decided upon three more or less independent factors: agonistic, non-reproductive and sexual. While the three sets of movements that make up these factors can be clearly recognized in our figures, our solutions suggest that there is one main factor consisting of two oppositely directed tendencies: the sexual tendency and the agonistic

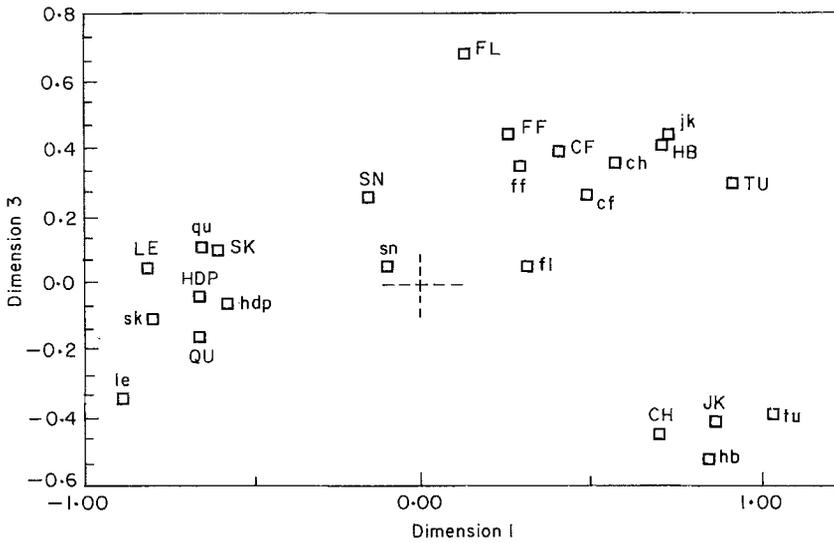


Figure 3. Correspondence analysis of data in Table I: decomposition of residuals from quasi-independence; dimensions 1 and 3. Capital letters represent preceding behaviours, small letters following behaviours. The solution can be interpreted as a correspondence analysis on the matrix with adjusted diagonal elements (see equation 8 and Fig. 1). See text for further details.

tendency, with the agonistic elements at extreme positions and the so-called non-reproductive activities at intermediary positions (see factor 2 in CA-Ind, Fig. 1 and factor 1 in CA-QInd, Fig. 2). This should be interpreted as follows: behavioural changes occur mainly within the three behaviour sets, while changes from one set to another are more unlikely between the sexual and the agonistic set than from these sets to the non-reproductive set.

Within the set of agonistic movements, Wiepkema made a distinction between two activities expressing a pure aggressive tendency (head butting and chasing) and three activities expressing more or less a flight tendency (turning beats, jerking and fleeing). Closely related with this distinction is Wiepkema's claim that there is an array of activities FL-JK-TU-HB-CHS in his principal components analysis solution which corresponds with an increase of the aggression/flight ratio. We are not convinced that this array can be recognized in his data. In addition neither this distinction nor the array is conspicuously visible in our solutions CA-Ind and CA-QInd.

From his principal components analysis Wiepkema inferred the existence of an inhibitive relationship between the causal factors corresponding with the aggressive set and corresponding with the two sexual elements: leading and quivering. The

same inhibitive relationship was inferred between the sexual tendency and the flight tendency (i.e. turning beats, jerking and fleeing). As shown above, our solutions show a general inhibitive relationship between the sexual tendency and the agonistic tendency.

Although factor 3 of the CA-QInd solution explains only 10% of the total inertia and although the elbow criterion provides an argument for not considering it, we have nevertheless taken it into consideration, because Wiepkema has extracted three dimensions as well. Examination of factor 3 suggests a meaningful interpretation, because the agonistic elements become widely spaced (Fig. 3). At one extreme we find two agonistic elements, chasing and jerking, which both lead rather predictably to the attack elements head butting and turning beats. Once the animal has shown these elements (on many occasions with the result that its adversary has fled) it may enter various other behavioural states, as demonstrated by the intermediary position of head butting and turning beats (preceding elements) on this axis. At the other extreme of the axis we find flight as a preceding element. This means that once an animal has fled, there is a small chance that it will show either aggressive or sexual behaviour again. Flight as a following element takes a rather central position,

indicating that animals can be brought to fleeing when they are in various behavioural states. We believe that this representation is more understandably in accordance with the 'non-factorial evidence' Wiepkema presents than his principal components analysis solution.

The solution of the CA-QInd shows very clearly the existence of asymmetries in the sequential associations among the sexual movements: head-down posture is followed by skimming and quivering by leading. This aspect of the directed sequential associations is not represented in the principal components analysis.

The positions of the elements chasing as well as some other activities (fleeing, fin-flickering and turning beats) in the principal components analysis solution for the rows is different from those in the solution for the columns, which caused interpretative difficulties. The performance of a correspondence analysis on a transition matrix results in a solution in which both the rows and the columns as well as the associations between rows and columns are presented. In the CA-QInd solution we do find different positions for the row and column elements in the set of sexual elements (factor 2) as well as in the set of agonistic elements and we could interpret these differences as being due to asymmetries in the sequential associations (see above).

One way to check the discrepancies between Wiepkema's solution and correspondence analysis is by going back to the data. We studied the standardized residuals from quasi-independence and this showed that correspondence analysis pointed to the correct conclusions. Of course, this was not surprising, since these measures were decomposed directly by correspondence analysis.

ASYMMETRIC RELATIONS

In the foregoing section we have studied the departure from quasi-independence. On the second correspondence analysis dimension some asymmetries in the data matrix were visible. In addition to these asymmetries, we might still be interested in whether there are other important asymmetries. Therefore, what we need is a solution in which the asymmetries are displayed clearly. In this section we discuss such a solution. Other ways of analysing asymmetry can be found in Digby & Kempton (1987).

When we are interested in asymmetry, we can calculate the matrix with values $\frac{1}{2}(p_{ij} + p_{ji})$. These

values are the averages of the values above and below the diagonal of the square matrix, and they are the maximum likelihood estimates under the so-called loglinear symmetry model. This model has $df = I(I-1)/2$. If the matrix with observed values p_{ij} equals the matrix with values $\frac{1}{2}(p_{ij} + p_{ji})$, then activity i is followed as often by activity j as the other way around, and clearly there is no asymmetry. By studying the differences between the observed values p_{ij} and the expected values $\frac{1}{2}(p_{ij} + p_{ji})$, we see whether activity i follows more or less often activity j than the other way around. If activity i is followed more often by activity j than the other way around, this can indicate that activity j triggers or causes the occurrence of activity i . In this way a study of residuals from symmetry can give some indication concerning causality in behaviour.

A drawback of this approach is that these values also reflect the difference between the marginal row frequency and column frequency of activity i . For a transition matrix like the Wiepkema data this is not often of interest because it merely reflects the design of the study (see above).

So we are interested in asymmetry due to aspects other than unequal margins. With a different symmetry model, namely the quasi-symmetry model, it is possible to study whether this asymmetry is sufficiently important to command attention. This model can be formulated as

$$\pi_{ij} = a_i b_j s_{ij} \quad (9)$$

where the s_{ij} are restricted as $s_{ij} = s_{ji}$. When we compare the quasi-symmetry model with the quasi-independence model (7), we see that the parameters a_i and b_j correspond again to the margins of the table, and the new parameters s_{ij} correspond to the symmetry in the table, where the interaction term above the diagonal must be equal to the interaction below the diagonal. Estimates of expected proportions have to be calculated iteratively (Bishop et al. 1975). Again we perform two consecutive steps. First, the chi-squared test (2) can be performed to test whether the quasi-symmetry model holds, with $df = (I-1)(I-2)/2$. If the model does not hold, i.e. when there is asymmetry in the data that cannot be neglected, it makes sense to take the second step and to study the residuals from this model.

The matrix with residuals from quasi-symmetry has an interesting property: it is skew-symmetric, i.e. a residual above the diagonal is equal to minus the residual below the diagonal. We employ a

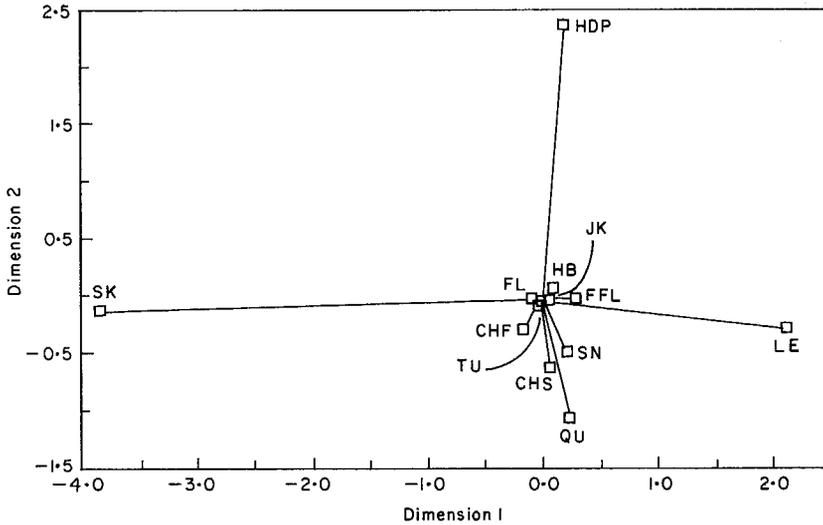


Figure 4. Correspondence analysis of data in Table I: decomposition of residuals from quasi-symmetry; first pair of dimensions. Going counter-clockwise, the residuals are positive. Two points (categories) that form a large triangle with the origin have a large residual (asymmetry). Points are close to the origin when the residuals with other points are small.

special correspondence analysis that uses this property. We discuss only the interpretational aspects of this correspondence analysis; for details see van der Heijden et al. (1989).

As in ordinary correspondence analysis, here the residuals are scaled to reduce the influence of the marginal frequencies. For ordinary correspondence analysis the residuals were scaled by the marginal frequencies p_{i+} and p_{+j} . For the special form of correspondence analysis we employ here, the residuals are scaled by the average of the row and column proportions, i.e. by $d_i = (p_{i+} + p_{+i})/2$.

Due to the skew-symmetry of the residual matrix, the resulting correspondence analysis solution has the singular values ordered in pairs: $\lambda_1, \lambda_1, \lambda_2, \lambda_2, \dots$. Therefore plots are made of these paired dimensions. To simplify the discussion, let us assume that the first paired dimension shows all there is, i.e. $\lambda_2 = 0$. In the first paired dimensions each activity is represented by one point (instead of two points, one for the row and one for the column). In this plot we interpret not distances between points, but areas: in fact the area ijO of the triangle of activity i , activity j , and the origin O is equal to the residual for activity i and j , i.e. $\frac{1}{2} d_i d_j n_{ij}$ where n_{ij} is the residual for row activity i and column activity j . This scaled residual for activity i and j is positive if the rotation from i to j is counter-clockwise, and it is negative when the rotation is clockwise (since $\frac{1}{2} d_i d_j n_{ij} = -\frac{1}{2} d_i d_j n_{ji}$).

If the first paired dimension does not suffice to show the complete asymmetry, then the areas of triangles in the first pair of dimensions represent only approximations of elements of $\frac{1}{2} d_i d_j n_{ij}$.

Two general principles for interpretation that follow from the area interpretation are, first, that activities lying on a line through the origin have no asymmetric relation because area ijO is zero, and, second, that points near the origin have unimportant asymmetries.

Example

We now perform a correspondence analysis decomposition of residuals from quasi-symmetry for the Wiepkema data. Thus we obtain a clear view of the asymmetry in the data. For the Wiepkema data the quasi-symmetry model fits badly: $\chi^2 = 293$ for $df = 55$. We conclude that there is still reason to inspect the residuals. A plot of the first paired dimension, with singular value 0.172 (proportion of explained inertia 0.935) is shown in Fig. 4. The correspondence analysis solution shows that, as far as asymmetries are present in the data, they are found predominantly in the cluster of sexual activities: head-down posture is followed relatively more often by skimming than the other way around; similarly for leading and head-down posture, quivering and leading, and skimming and quivering. Only two asymmetries were already revealed in

Fig. 2, namely head-down posture leading more to skimming than the other way around, and quivering leading more to leading behaviour than the other way around. The solution also shows that there are no important asymmetries between sexual activities on the one hand and non-reproductive activities and aggressive activities on the other hand, since these latter activities are all near the origin. We conclude that it made sense to study the residuals from quasi-symmetry in a separate solution: this clear picture of the asymmetries was revealed neither by our earlier analyses, nor by Wiepkema's.

CONCLUSIONS

We have shown that correspondence analysis is a useful tool for the analysis of transition matrices. The first reason is that in transition matrices the profile concept is important, and that in correspondence analysis distances between profiles are projected onto low-dimensional space. Ethologically, the comparison of profiles is used in the concepts precedent and subsequent mutual replaceability, so correspondence analysis can show which activities are mutually replaceable. Second, correspondence analysis can help to find relations between cells for which the residuals differ significantly from the selected model. Ethologically, this implies that correspondence analysis can give us a summary of the (directed) sequential association between the elements. Third, correspondence analysis is a flexible tool in the sense that we can restrict attention to particular aspects in the data such as the off-diagonal dependence (i.e. attention is restricted to the behavioural changes only) and asymmetries in behavioural changes. These arguments become more important to the extent that the number of categories is large. We believe that correspondence analysis should replace one of the current practices of performing two factor analyses on a transition matrix. Factor analysis does not have the above-mentioned advantages. Furthermore, factor analysis uses correlation coefficients as measures of similarity. A correlation is a good measure for similarity when the data stem from a bivariate normal distribution. Since this assumption cannot be defended for (scaled) rows and columns of contingency tables, wrong conclusions can result from using correlation coefficients (see above).

Correspondence analysis has many similarities in objectives with other data analysis tools. We have

presented correspondence analysis as a multi-dimensional scaling technique (compare Spence 1978) in which chi-squared distances are approximated. It is also closely related to model-fitting approaches such as loglinear analysis, as we have shown above (see also van der Heijden et al. 1989). Correspondence analysis has in common with factor analysis and principal components analysis that a data matrix is approximated by a matrix of lower rank. Compared with cluster analysis (Morgan et al. 1976; de Gheff 1978), correspondence analysis can also be helpful for the classification of the rows or columns of contingency tables.

In the literature correspondence analysis is often used together with some form of cluster analysis in the same way as the latter is applied together with other techniques like factor analysis or multidimensional scaling (compare de Gheff 1978). In general, results from cluster analysis can simplify the interpretation of the correspondence analysis solution. Results of correspondence analysis and cluster analysis can become quite similar if in cluster analysis the chi-squared distance is used as the similarity measure.

APPENDIX

Computation

We give a more technical description of correspondence analysis here. For details see Greenacre (1984) and van der Heijden et al. (1989). We need some matrix notation to explain the computation of correspondence analysis. The correspondence analysis solution can be found as follows: let \mathbf{D}_r and \mathbf{D}_c be diagonal matrices with, respectively, marginal row proportions p_{i+} and column proportions p_{+j} ; $\hat{\mathbf{I}}$ is the matrix with estimates under independence (1). We then compute the singular value decomposition of the matrix with standardized residuals (3), divided by n . In matrix terms this gives

$$\mathbf{D}_c^{-\frac{1}{2}} (\mathbf{P} - \hat{\mathbf{I}}) \mathbf{D}_r^{-\frac{1}{2}} = \mathbf{U} \mathbf{A} \mathbf{V}' \quad (10)$$

where $\mathbf{U}'\mathbf{U} = \mathbf{I} = \mathbf{V}'\mathbf{V}$, and \mathbf{A} is a diagonal matrix with singular values λ_a in descending order. The row and column scores are normalized as follows

$$\mathbf{R} = \mathbf{D}_r^{-\frac{1}{2}} \mathbf{U} \mathbf{A} \quad (11a)$$

$$\mathbf{C} = \mathbf{D}_c^{-\frac{1}{2}} \mathbf{V} \mathbf{A} \quad (11b)$$

By substituting (11) in (10), one finds

$$\mathbf{P} = \hat{\mathbf{I}} + \mathbf{D}_r \mathbf{R} \mathbf{A}^{-1} \mathbf{C}' \mathbf{D}_c \quad (12)$$

which is known under the name 'reconstitution formula'. It shows that correspondence analysis decomposes the difference $(\mathbf{P} - \hat{\mathbf{P}})$, being the residuals from independence.

In generalized correspondence analysis the matrix \mathbf{P} is decomposed as

$$\mathbf{P} = \hat{\mathbf{P}} + \mathbf{S}_r \mathbf{R} \mathbf{A}^{-1} \mathbf{C}' \mathbf{S}_c \quad (13)$$

where \mathbf{S}_r and \mathbf{S}_c are diagonal, and $\hat{\mathbf{P}}$ is not necessarily independent (as was the case for ordinary correspondence analysis). For correspondence analysis of incomplete tables $\hat{\mathbf{P}}$ is a matrix with estimates under the quasi-independence model, \mathbf{S}_r and \mathbf{S}_c are diagonal with values a_i and b_j , respectively. Decomposition (13) can be found as in (10) and (11) by replacing \mathbf{D}_r by \mathbf{S}_r and \mathbf{D}_c by \mathbf{S}_c . An easier way to find the generalized correspondence analysis solution for quasi-independence is given by de Leeuw & van der Heijden (1988) and described in the present paper.

For the quasi-symmetry model the matrix $\hat{\mathbf{P}}$ is a matrix with estimates under quasi-symmetry, \mathbf{S}_r and \mathbf{S}_c are diagonal with values $(p_{i+} + p_{+j})/2$. Van der Heijden et al. (1989) show that the matrix $(\mathbf{P} - \hat{\mathbf{P}})$ is skew-symmetric, i.e. $(p_{ij} - \pi_{ij}) = -(p_{ji} - \pi_{ji})$, and discuss the implications of this for correspondence analysis (see Gower 1977 for a discussion of singular value decomposition of skew-symmetric matrices).

We conclude that the decompositions of residuals from independence and quasi-independence can both be performed using ordinary correspondence analysis programs. In the near future it will be possible to do ordinary correspondence analysis using program packages like BMDP, SAS or SPSS, using the moment-correspondence analysis routines being developed. The correspondence analysis decomposing residuals from quasi-symmetry has to be performed using a special purpose program. It does not cost too much time to write your own correspondence analysis program if you have a tool for singular value decomposition or eigenvalue decomposition available (such as in GENSTAT or SAS MATRIX). The crucial operations are equations (10) and (11). For all procedures APL-programs are available from us free of charge.

ACKNOWLEDGMENTS

We are grateful to Dr P. W. Colgan and Dr B. Thon for helpful comments on an earlier version of this

paper. We thank Mrs S. M. McNab for her advice on style and spelling.

REFERENCES

- Baerends, G. P. 1984. The organization of the pre-spawning behaviour in the cichlid fish *Aequidens portaleguensis* (Hensel). *Neth. J. Zool.*, **34**, 233-366.
- Balthazart, J. 1972. Validité de l'application de l'analyse factorielle à l'étude causale d'une séquence comportementale avec un essai d'analyse du comportement agonistique chez *Tilapia macrochir* (Boulenger 1912). *Annls Soc. r. zool.*, **102**, 3-34.
- Balthazart, J. 1973. Analyse factorielle du comportement agonistique chez *Tilapia macrochir* (Boulenger 1912). *Behaviour*, **46**, 37-72.
- Benzécri, J.-P. (& collaborateurs) 1973. *Analyse des Données. Vol. 2*. Paris: Dunod.
- van den Bercken, J. H. L. & Cools, A. R. 1980a. Information-statistical analysis of social interaction and communication: an analysis-of-variance approach. *Anim. Behav.*, **28**, 172-188.
- van den Bercken, J. H. L. & Cools, A. R. 1980b. Information-statistical analysis of factors determining ongoing behaviour and social interaction in Java monkeys (*Macaca fascicularis*). *Anim. Behav.*, **28**, 189-200.
- Bishop, Y. M. M., Fienberg, S. E. & Holland, P. W. 1975. *Discrete Multivariate Analysis*. Cambridge, Massachusetts: MIT Press.
- Boy, V. & Duncan, P. 1979. Time-budgets of Camargue horses. I. Developmental changes in the time-budgets of foals. *Behaviour*, **71**, 187-202.
- Burton, J. 1970. Etude critique de l'analyse factorielle de la rivalité territoriale chez *Tilapia mossambica* Peters (Poisson Cichlide). *Annals Soc. r. zool.*, **100**, 5-47.
- Cane, V. R. 1978. On fitting low-order Markov chains to behaviour sequences. *Anim. Behav.*, **26**, 332-338.
- Castellan, N. J., Jr. 1979. The analysis of behavior sequences. In: *The Analysis of Social Interaction: Methods, Analysis and Illustrations* (Ed. by R. B. Cairns), pp. 81-116. Hillsdale, New Jersey: Lawrence Erlbaum.
- Chatfield, C. 1973. Statistical inference regarding Markov chain models. *Appl. Statist.*, **22**, 7-20.
- Chatfield, C. & Lemon, R. E. 1970. Analysing sequences of behavioural events. *J. theor. Biol.*, **29**, 427-445.
- Colgan, P. W. & Smith, J. T. 1978. Multidimensional contingency table analysis. In: *Quantitative Ethology* (Ed. by P. W. Colgan), pp. 145-174. New York: John Wiley.
- Digby, P. G. N. & Kempton, R. A. 1987. *Multivariate Analysis of Ecological Communities*. London: Chapman & Hall.
- Douglas, J. M. & Tweed, R. L. 1979. Analysing the patterning of a sequence of discrete behavioural events. *Anim. Behav.*, **27**, 1236-1252.
- Duncan, P. 1980. Time-budgets of Camargue horses. II. Time-budgets of adult horses and weaned sub-adults. *Behaviour*, **72**, 26-49.
- Duncan, P. 1985. Time-budgets of Camargue horses. III. Environmental influences. *Behaviour*, **91**, 188-208.

- Fagen, R. M. & Mankovich, N. J. 1980. Two-act transitions, partitioned contingency tables, and the 'significant cells' problem. *Anim. Behav.*, **28**, 1017-1023.
- de Gheff, V. J. 1978. Hierarchical cluster analysis. In: *Quantitative Ethology* (Ed. by P. W. Colgan), pp. 115-144. New York: John Wiley.
- Gifi, A. 1981. *Non-linear Multivariate Analysis*. Leiden: Leiden University, Department of Data Theory.
- Goodman, L. A. 1986. Some useful extensions of the usual correspondence analysis approach and the usual loglinear models approach in the analysis of contingency tables. *Int. Statist. Rev.*, **54**, 243-309.
- Gower, J. C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, **53**, 325-338.
- Gower, J. C. 1977. The analysis of asymmetry and orthogonality. In: *Recent Developments in Statistics* (Ed. by J. R. Barra, F. Brodeau, G. Romier & B. van Cutsem), pp. 109-123. Amsterdam: North Holland.
- Greenacre, M. J. 1984. *Theory and Applications of Correspondence Analysis*. New York: Academic Press.
- Haccou, P. 1986. Analysis of behaviour by means of continuous time Markov chain models and their generalizations. In: *Quantitative Models in Ethology* (Ed. by P. W. Colgan & R. Zayan), pp. 81-96. Toulouse: Privat, I.E.C.
- Haccou, P. 1987. Statistical methods for ethological data. Ph.D. thesis, Centrum voor Wiskunde en Informatica, Amsterdam.
- van der Heijden, P. G. M. 1986. Transition matrices, model fitting, and correspondence analysis. In: *Data Analysis and Informatics IV* (Ed. by E. Diday), pp. 221-226. Amsterdam: North Holland.
- van der Heijden, P. G. M. 1987. *Correspondence Analysis of Longitudinal Categorical Data*. Leiden: D.S.W.O.-Press.
- van der Heijden, P. G. M., de Falguerolles, A. & de Leeuw, J. 1989. A combined approach to contingency table analysis using correspondence analysis and log-linear analysis. *Appl. Statist.*, **38**, 249-292.
- Hoekstra, J. A. & Reddingius, J. 1981. Over de weergave van associaties tussen gedragselementen. Wageningen: Instituut Toegepast Natuurwetenschappelijk Onderzoek voor Wiskunde, Informatica, Verwerking en Statistiek.
- van Hooff, J. A. R. A. M. 1970. A component analysis of the structure of the social behaviour of a semi-captive chimpanzee group. *Experientia*, **26**, 549-550.
- van Hooff, J. A. R. A. M. 1973. A structural analysis of the social behaviour of a semi-captive group of chimpanzees. In: *Expressive Movement and Nonverbal Communication* (Ed. by M. von Cranach & I. Vine), pp. 75-162. London: Academic Press.
- van Hooff, J. A. R. A. M. 1982. Categories and sequences of behaviour: methods of description and analysis. In: *Handbook of Methods in Nonverbal Behaviour Research* (Ed. by K. R. Scherer & P. Ekman), pp. 362-439. Cambridge: Cambridge University Press.
- de Leeuw, J. & van der Heijden, P. G. M. 1988. Correspondence analysis of incomplete tables. *Psychometrika*, **53**, 223-233.
- Lemon, R. E. & Chatfield, C. 1971. Organization of song in cardinals. *Anim. Behav.*, **19**, 1-17.
- Loots, G. M. P. 1985. Social relationships in groups of children: an observational study. Ph.D. thesis, Free University of Amsterdam.
- Losey, G. S. 1978. Information theory and communication. In: *Quantitative Ethology* (Ed. by P. W. Colgan), pp. 43-78. New York: John Wiley.
- Metz, J. A. J., Dieneske, H., de Jonge, G. & Putters, F. 1983. Continuous time Markov chains as models for animal behaviour. *Bull. math. Biol.*, **45**, 643-658.
- Morgan, B. J. T., Simpson, M. J. A., Hanby, J. P. & Hall-Craggs, J. 1976. Visualizing interaction and sequential data in animal behaviour: theory and application of cluster-analysis methods. *Behaviour*, **56**, 1-43.
- Nishisato, S. 1980. *Analysis of Categorical Data: Dual Scaling and its Applications*. Toronto: University of Toronto Press.
- Slater, P. J. B. 1973. Describing sequences of behavior. In: *Perspectives in Ethology* (Ed. by P. P. G. Bateson & P. H. Klopfer), pp. 131-153. New York: Plenum Press.
- Slater, P. J. B. & Ollason, J. C. 1972. The temporal pattern of behaviour in isolated male zebra finches: transition analysis. *Behaviour*, **42**, 248-269.
- Spence, I. 1978. Multidimensional scaling. In: *Quantitative Ethology* (Ed. by P. W. Colgan), pp. 175-217. New York: John Wiley.
- Thon, B. & Chabaud, C. 1986. Act duration as a potential informational parameter in the sequential analysis of behaviour: visualization by correspondence analysis method. In: *Quantitative Models in Ethology* (Ed. by P. W. Colgan & R. Zayan), pp. 129-142. Toulouse: Privat, I.E.C.
- Vodegel, N. 1978. A causal analysis of the behaviour of *Pseudotropheus zebra*. Ph.D. thesis, University of Groningen.
- Voeten, M. J. M. 1985. Sequential analysis of teacher-student interaction. Ph.D. thesis, Catholic University of Nijmegen.
- de Waal, F. B. M. 1978. Exploitative and familiarity-dependent support strategies in a colony of semi-free living chimpanzees. *Behaviour*, **66**, 268-312.
- Wiepkema, P. R. 1961. An ethological analysis of the reproductive behaviour of the bitterling (*Rhodeus amarus* Bloch). *Archs Néerl. Zool.*, **14**, 103-199.

(Received 16 September 1988; initial acceptance
11 January 1989; final acceptance 5 May 1989;
MS. number: 3288)