

## CORRESPONDENCE ANALYSIS USED COMPLEMENTARY TO LOGLINEAR ANALYSIS

PETER G. M. VAN DER HEIJDEN

DEPARTMENT OF METHODOLOGY  
UNIVERSITY OF LEIDEN

JAN DE LEEUW

DEPARTMENT OF DATA THEORY  
UNIVERSITY OF LEIDEN

Loglinear analysis and correspondence analysis provide us with two different methods for the decomposition of contingency tables. In this paper we will show that there are cases in which these two techniques can be used complementary to each other. More specifically, we will show that often correspondence analysis can be viewed as providing a decomposition of the difference between two matrices, each following a specific loglinear model. Therefore, in these cases the correspondence analysis solution can be interpreted in terms of the difference between these loglinear models. A generalization of correspondence analysis, recently proposed by Escofier, will also be discussed. With this decomposition, which includes classical correspondence analysis as a special case, it is possible to use correspondence analysis complementary to loglinear analysis in more instances than those described for classical correspondence analysis. In this context correspondence analysis is used for the decomposition of the residuals of specific restricted loglinear models.

Key words: correspondence analysis, data analysis, loglinear analysis, multi-dimensional scaling.

### 1. Introduction

For the analysis of contingency tables, loglinear analysis is already a very popular technique in the English speaking countries. Standard references are Andersen (1980), Bishop, Fienberg and Holland (1975), and Fienberg (1980). In the last few years there is a growing interest in correspondence analysis, which has been the most important data analytic technique in France for many years. When one has full command of the French language, the basic works are those of the group around Benzécri (1973, 1980). In the English speaking world the growing interest is apparent from works written by de Leeuw (1984), Gifi (1981), Greenacre (1984), and Nishisato (1980). Apart from these books, the number of articles and contributions at conferences is growing rapidly.

Strangely enough, correspondence analysis was already known in the English literature for a long time but under several other names. Nishisato (1980) gives a full survey of all these names and references which have appeared in the history of correspondence analysis. Greenacre (1984) accentuates that these other approaches have a different rationale and interpretation. He discusses this for the approaches "reciprocal averaging," "dual (or optimal) scaling," "canonical correlation analysis," and "simultaneous linear regressions." The recent flourishing of correspondence analysis as a data analytic technique is probably due to the heavy emphasis on the geometrical aspects of the method. Canonical correlation analysis of categorical data (Kendall & Stuart, 1973, pp. 588–598), which is

Requests for reprints should be sent to P. G. M. van der Heijden, Department of Methodology, University of Leiden, Hooigracht 15, 2312 KM Leiden, THE NETHERLANDS.

proved by de Leeuw (1971) to be formally identical to correspondence analysis, emphasizes the quantification aspect.

With the growing interest in correspondence analysis, there is a natural interest in relations between correspondence analysis and loglinear analysis. Until now, the literature on this topic can be divided into two approaches. The first approach deals with the RC-model (RC for "row, column") of Goodman (1979, 1981a, 1981b) and generalizations of this model (Agresti, 1983; de Leeuw, 1983). The RC-model is a model for a two-way contingency table in which the loglinear interaction parameter has a multiplicative form. Goodman (1981b) shows that, when the frequencies in a two-way table are generated from an underlying discretized bivariate normal distribution (or a bivariate distribution which is bivariate normal after a proper transformation), the estimates of the multiplicative row and column parameters are approximately the same as the row and column scores found for the first dimension in correspondence analysis. See also Fienberg and Meyer (1983) and Israëls and Sikkel (1982) for other relations between the RC-model and correspondence analysis.

Contributions in the second approach discuss the complementary use of correspondence analysis and loglinear analysis. Examples are Daudin and Trécourt (1980) and Israëls and Sikkel (1982). They conclude that loglinear analysis is the method most apt to trace higher-order interactions, although it is recognized that the interpretation of these interactions is often difficult. When the number of variables is small, loglinear analysis can be complemented by correspondence analysis. In these two papers complementary use of the techniques is advocated only in a general way. It is not made explicit how exactly these methods can be used in a complementary way. It is our purpose to propose more specific ways of combination in this paper.

First we give a short introduction to correspondence analysis. Following the French tradition, heavy emphasis will be placed upon the geometrical aspects of the technique. Correspondence analysis as such is a technique for the analysis of two-way tables. We will discuss the most usual way to analyse higher-way tables with correspondence analysis. We will not deal with correspondence analysis of higher-way contingency tables by means of the so-called Burt matrix (see Gifi, 1981, pp. 134–162; Greenacre, 1984, chap. 5). In section 3 we will discuss loglinear analysis. Special attention is paid to the case of the three-way contingency table. In section 4 we will present our main results on the complementary relation of correspondence analysis and loglinear analysis. It will be shown that the correspondence analysis solution is based on the decomposition of the difference between two loglinear models, which will be specified there. The results will be presented for two-way, three-way and higher-way contingency tables respectively. Our results will be illustrated by an example. A three-way table on suicide behavior is analyzed. In section 5 a generalization of correspondence analysis is discussed. The generalization is used to make it possible to use correspondence analysis complementary to loglinear analysis in more cases than these discussed in section 4. We will end with some conclusions.

## 2. Correspondence Analysis of Contingency Tables

We will treat correspondence analysis briefly here, with an emphasis on the geometrical aspects. For details and proofs we refer to Gifi (1981) or Greenacre (1984), and the references mentioned there.

Correspondence analysis is a technique with which it is possible to find a multi-dimensional representation of the dependence between the row and column variable of a two-way contingency table. This representation can be constructed using the scores found for row and column categories as coordinates for the category points. These scores can be normalized in such a way that distances between row points or between column points in euclidian space are equal to chi-squared distances. By emphasizing the chi-squared dis-

tances, Heiser and Meulman (1983) treat correspondence analysis within the multi-dimensional scaling frame work.

Consider a two-way contingency table  $F$  with elements  $f_{ij}$ , having  $I$  rows ( $i = 1, \dots, i, \dots, i', \dots, I$ ) and  $J$  columns ( $j = 1, \dots, j, \dots, j', \dots, J$ ). An index is replaced by “+” when summed over the corresponding variable, for example  $\sum_j f_{ij} = f_{i+}$ . The chi-squared distances are computed on the profiles of the corresponding rows or columns, where for example the profile of row  $i$  is the row of values  $f_{ij}/f_{i+}$ . So  $\sum_j f_{ij}/f_{i+} = 1$ . The chi-squared distance between rows  $i$  and  $i'$  is defined as

$$\delta^2(i, i') = \sum_j \frac{\left(\frac{f_{ij}}{f_{i+}} - \frac{f_{i'j}}{f_{i'+}}\right)^2}{\frac{f_{+j}}{n}}, \tag{1}$$

where  $n = f_{++}$ . Equation (1) shows that  $\delta^2(i, i')$  is a measure of the difference between the profiles of row  $i$  and  $i'$ ; when  $i$  and  $i'$  have the same profile,  $\delta^2(i, i') = 0$ .

The correspondence analysis solution can be found as follows: let  $X$  be the matrix to be analysed;  $D_r$  and  $D_c$  diagonal matrices with marginal row frequencies  $x_{i+}$  and column frequencies  $x_{+j}$  respectively;  $E = D_r \mathbf{t} \mathbf{t}' D_c / n$ , where  $n = f_{++}$  and  $\mathbf{t}$  is a vector with ones, the length of which depends on the context. Elements of  $E$  have the following form:

$$e_{ij} = \frac{x_{i+} x_{+j}}{n}. \tag{2}$$

Subsequently the singular value decomposition of the matrix  $D_r^{-1/2}(X - E)D_c^{-1/2}$  is computed. Elements of this matrix have value  $(1/n^{1/2})(x_{ij} - e_{ij}/e_{ij}^{1/2})$ , which are standardized residuals, scaled by  $(1/n)^{1/2}$ . These residuals are decomposed with (3):

$$D_r^{-1/2}(X - E)D_c^{-1/2} = U \Lambda V', \tag{3}$$

where  $U'U = I$ ,  $V'V = I$ , and  $\Lambda$  is a diagonal matrix with singular values  $\lambda_\alpha$  in descending order;  $\alpha$  is the index for dimension. The dimensionality of the solution is equal to  $\min(I - 1, J - 1)$ . For the remaining dimensions  $\lambda_\alpha = 0$ .

$U$  and  $V$  contain scores corresponding with the row and column categories. These scores are normalized as follows:

$$\begin{aligned} R &= D_r^{-1/2} U n^{1/2}, & \text{and} \\ C &= D_c^{-1/2} V n^{1/2}. \end{aligned} \tag{4}$$

So  $R'D_r R = nI$  and  $C'D_c C = nI$ . Furthermore  $\mathbf{t}' D_r R = 0$  and  $\mathbf{t}' D_c C = 0$ : For each dimension row scores and column scores have a weighted variance of 1 and a weighed average of 0.

One can make a simultaneous representation of row and column points in three ways (Gifi, 1981, pp. 134–151): (a) The first can be made by using scores  $R$  and  $\tilde{C} = CA$  as coordinates, so that the euclidean distances between column points are equal to chi-squared distances. The weighted variance of the coordinates of the column points equals  $\lambda_\alpha^2$  for each dimension. (b) The second is made by using  $\tilde{R} = RA$  and  $C$  as coordinates, so that the analogous result holds for the row points. (c) The third simultaneous representation is made by using  $RA^{1/2}$  and  $CA^{1/2}$ , so that a symmetric representation of row and column points is chosen. Row scores can be derived from column scores (and column scores from row scores) with the so called “transition formulas”:

$$\begin{aligned} R &= D_r^{-1} X C A^{-1}, & \text{and} \\ C &= D_c^{-1} X' R A^{-1}. \end{aligned} \tag{5}$$

Bringing  $\Lambda$  from the right to the left side of (5), it is easily seen that in the above mentioned simultaneous representation, (b), the row scores  $\tilde{R}$  are in the weighted average of the column scores  $C$ , and in (a) the column scores  $\tilde{C}$  in the weighted average of the row scores  $R$ . This property is called the "barycentric principle." Here the weighting is done by the column and row profiles. The transition formulas define the rationale for the "reciprocal averaging" approach.

Chi-squared distances between row  $i$  and  $i'$ , and their approximations, can be computed with (6):

$$(I_i - I_{i'})D_r^{-1}XD_c^{-1}X'D_r^{-1}(I_i - I_{i'})n = (\tilde{r}_i - \tilde{r}_{i'})/(\tilde{r}_i - \tilde{r}_{i'}), \quad (6)$$

where  $I_i$  is the  $i$ -th row of the identity matrix,  $\tilde{r}_i$  the  $i$ -th row of matrix  $\tilde{R}$ . With the right term of (6) one can compute approximations of the chi-squared distance by dropping the last column(s) of  $R$ .

The so-called reconstitution formula (Benzécri, 1973, 1980; Greenacre, 1984, p. 93) can be found by substituting (4) in (3):

$$D_r^{-1}(X - E)D_c^{-1}n = R\Lambda C',$$

so that

$$X = E + D_r R\Lambda C' D_c n^{-1} = n^{-1}D_r(\mathbf{t}\mathbf{t}' + R\Lambda C')D_c. \quad (7)$$

Elements of  $R\Lambda C'$  are equal to  $(x_{ij} - e_{ij})/e_{ij}$ . Equation (7) shows that correspondence analysis decomposes the departure from independence in a matrix. This decomposition has the following relation with the well-known Pearson goodness-of-fit  $X^2$  statistic, which will be defined in (14):

$$\text{trace } \Lambda^2 = X^2/n, \quad (8)$$

where  $\text{trace } \Lambda^2$  is called the total "inertia." So, correspondence analysis decomposes the  $X^2$ -value of a matrix (Kendall & Stuart, 1971, pp. 588-594). The importance of dimension  $\alpha$  can be evaluated by the ratio of the inertia of dimension  $\alpha$  and the total inertia:  $\lambda_\alpha^2 / \sum_\alpha \lambda_\alpha^2$ . This quantity can be interpreted as the proportion "explained" inertia for dimension  $\alpha$ , or the proportion of  $X^2$  that is decomposed in dimension  $\alpha$ .

Clouds of points can be interpreted using chi-squared distances: when two row points (or two column points) are near each other, their profiles are similar. When profiles differ considerably, the distance between the points is large. The profiles of the marginal frequencies of  $X$  are projected into 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. The distance of row  $i$  to column  $j$  can be interpreted with the transition formulas; roughly one can say that  $i$  and  $j$  will be near each other when  $x_{ij} \gg e_{ij}$ , and that  $i$  and  $j$  are far apart from  $x_{ij} \ll e_{ij}$ .

An important aid in the interpretation of a solution is the property that the sum of the weighted squared distances of the row points (or column points) to the origin, is equal to  $\lambda_\alpha^2$  for dimension  $\alpha$ :

$$\lambda_\alpha^2 = \sum_i \left( \frac{f_{i+}}{n} \right) \tilde{r}_{i\alpha}^2 = \sum_j \left( \frac{f_{+j}}{n} \right) \tilde{c}_{j\alpha}^2. \quad (9)$$

With (9) one can evaluate the relative contribution of row  $i$  to dimension  $\alpha$  with the ratio  $((f_{i+}/n)\tilde{r}_{i\alpha}^2)/\lambda_\alpha^2$ , which can also be interpreted as the proportion of variance in dimension  $\alpha$  accounted for by row  $i$ . The same holds for column point  $j$ , when one uses the last term of (9). Using the theorem of Pythagoras, it is also possible to compute for row  $i$  on dimension  $\alpha$  the ratio of the squared projected distance and the squared total distance to the origin. With this ratio it is possible to evaluate how good the total chi-squared distance of

row  $i$  to the origin is represented on dimension  $\alpha$ . We will use these ratios when we interpret the examples.

In the introduction it was stated that correspondence analysis is formally identical to canonical correlation analysis of contingency tables. From this follows the special relation between  $\lambda_1$  and the Pearson product-moment correlation coefficient: The correlation between the row and column variable is, under all possible rescalings of the row and column categories, maximal and equal to  $\lambda_1$ , when as quantification of the categories of both variables the scores for the first dimension are taken.  $\lambda_2$  is equal to the maximal correlation of the quantified variables, where the quantification is restricted to be orthogonal to the quantification for the first dimension, and so forth (Kendall & Stuart, 1973, pp. 588–594). Correspondence analysis thus finds the maximal canonical correlations between the row and column variables.

Correspondence analysis as explained above is a method for the analysis of two-way contingency tables. A higher-way contingency table can be analysed in several ways. We will differentiate between three approaches. The first approach takes “slices” of a data block as its starting point: one can analyze separate tables for each category of other variables, for example when the three way table  $F$  is of order  $I \times J \times K$ , one can analyze two-way tables of order  $I \times J$  for every category of the third variable. It is also possible to concatenate these slices, thus creating “interactive variables.” Interactive variables are constructed by merging two or more original variables. For a three-way table  $F$  one can find three two-way matrices having order  $I \times (J \times K)$ ,  $J \times (I \times K)$  and  $K \times (I \times J)$  respectively. As far as we know, Gifi (1981, p. 151) is the only English standard reference in which this alternative is actually applied. On the other hand, in the French literature this is the method used most often to study interaction in higher-way tables. The two-way tables that are formed in this way are called “tableaux multiples” (e.g. Benzécri, 1973, 1980). In the sequel we will confine ourselves to this alternative and refer to these tables as “multiple tables.” Multiple tables will be notated by placing the variables, which constitute the interactive variables, between brackets, for example in a three-way table the three possible multiple tables are  $F^{1 \times (2 \times 3)}$ ,  $F^{2 \times (1 \times 3)}$  and  $F^{3 \times (1 \times 2)}$ .

A second approach for the analysis of higher way tables takes adding up over variables as its starting point. One can analyze marginal tables of two variables, but it is also possible to analyze concatenations of these marginal tables, for example in the three-way table one possibility is a table of order  $I \times (J + K)$  formed by concatenating the marginal tables of order  $I \times K$  and  $I \times J$ . Leclerc (1975) discusses properties of correspondence analysis solutions. A third possibility in this approach is to analyze the so-called Burt matrix, which consists of concatenations of uni- and bivariate marginal matrices, and is of order  $(I + J + K) \times (I + J + K)$  (Greenacre, 1984, p. 140–143). Correspondence analysis of a Burt matrix comes to the same as homogeneity analysis (Gifi, 1981) or multiple correspondence analysis (Greenacre).

In the third approach a higher way analogue of singular value decomposition is performed on standardized residuals of the higher-way matrix (Kroonenberg, 1983). For four and higher way tables the three approaches can of course be mixed, leaving us with an abundance of possibilities. Further research is needed to get a better understanding of all these possibilities, especially as to which possibility gives the best answer to which question.

Practice seems to show that in the two variable case correspondence analysis is an especially good method to gain insight into the relation between the variables of a contingency table when the number of categories is large.

### 3. Loglinear Analysis

Loglinear analysis is a well-known method to study the structural relations between variables in a contingency table. We treat here only some main points of loglinear analy-

sis. For details and proofs we refer to Andersen (1980), Bishop et al. (1975), and Fienberg (1980). Loglinear analysis decomposes the logarithm of the frequencies using the linear model. For example, in the case of a three-way matrix, the unrestricted loglinear model (called "saturated model") has the form

$$\log f_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)} + u_{123(ijk)}. \quad (10)$$

In the saturated case one has for a frequency  $f_{ijk}$  one parameter for the mean; three sets of parameters for the margins; three sets of parameters for the so-called first order interaction (or two-factor interaction); and one set of parameters for second order interaction. The  $u$ -parameters follow the usual ANOVA constraints: they sum to zero for each subscript.

The purpose of loglinear model fitting is to investigate whether, for a given dataset, certain  $u$ -parameters can be restricted to have value zero, for all  $i, j$  and  $k$ . The procedure is to specify a restricted model, to compute expected frequencies under this model, and to evaluate whether the difference between the expected frequencies and the observed frequencies is large. This evaluation usually takes place with the aid of Pearson's goodness-of-fit statistic  $X^2$ , or the likelihood-ratio chi-squared test  $G^2$ . Consult the above mentioned reference for details. Very often attention is restricted to the class of hierarchical models; When a  $u$ -term is zero, all higher order  $u$ -terms having the same indices are also zero; for example when  $u_{12(ij)} = 0$ , then  $u_{123(ijk)} = 0$ . The primary reason for this restriction is interpretive (Fienberg, 1980, p. 43).

Maximum likelihood-estimation is the most often used estimation method for loglinear models. The estimates are unique, and the estimation formulas do not differ when the frequencies form a sample from a (product-) multinomial or Poisson distribution. ML-estimation amounts to the same thing as fitting marginal tables to the full table. For example in the restricted model

$$\log f_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{23(jk)}, \quad (11)$$

and expected frequencies  $m_{ijk}$  have the following structure:

$$m_{ijk} = \frac{m_{i+} + m_{+jk}}{n}. \quad (12)$$

Complete minimal sufficient statistics for  $m_{i+}$  and  $m_{+jk}$  are  $f_{i+}$  and  $f_{+jk}$  respectively, so for model (11)

$$\hat{m}_{ijk} = \frac{f_{i+} + f_{+jk}}{n}. \quad (13)$$

It is easy to see that for the fitted marginals  $f_{i+} = \hat{m}_{i+}$  and  $f_{+jk} = \hat{m}_{+jk}$ , that is for the fitted margins the observed marginal frequencies are equal to margins of estimates of the expected frequencies. This is a special property of ML-estimation. For (11), it is possible to estimate the expected frequencies directly. This is not always the case: Sometimes expected frequencies have to be estimated iteratively when they can not be written out in closed form. The procedure mostly used is called "iterative-proportional fitting." For the three variable case this is only needed in case of hypothesis  $u_{123} = 0$ . See Reynolds (1977, chap. 6) for details on this procedure, and for general specifications when direct estimates cannot be found.

Evaluation of the fit of models is done with the Pearson goodness-of-fit chi-squared statistic  $X^2$ ,

$$X^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}, \quad (14)$$

where the sum is taken over all frequencies; or the likelihood-ratio chi-squared statistic  $G^2$ :

$$G^2 = 2 \sum \text{Observed} \log (\text{Observed}/\text{Expected}), \tag{15}$$

where again the sum is taken over all frequencies. The values of  $X^2$  and  $G^2$  are asymptotically distributed as a chi-squared variate when the specified model is true. The number of degrees of freedom is equal to the number of cells with non-zero fitted values minus the number of fitted independent parameters. A good review of goodness-of-fit measures is Bonett and Bentler (1983).

We code LL-models by placing the variables that constitute the highest fitted marginals between brackets: For the example above we write the model as [1][23], for the saturated model we write [123].

It is not usual to interpret individual  $u$ -parameters. This has the following reasons: Unlike in the ANOVA situation, there is no underlying empirical scale for a dependent variable. The “dependent variable” here is the logarithm of a frequency. Furthermore, the number of  $u$ -parameters that must be interpreted often becomes very large, especially when there are higher order interactions present, or when the number of categories is large. Moreover, it is not necessary to estimate these  $u$ -parameters if we only want to investigate whether variables are related or not.

Loglinear analysis seems especially a useful method to trace relations between variables when the number of variables and the number of categories is not too large.

#### 4. Correspondence Analysis Decomposing the Difference Between Two Specific Loglinear Models

Equation (6) showed that correspondence analysis decomposes the difference  $(X - E)$ , where the computation of values of  $E$  is based on the margins of  $X$ , see (2). We will show that values  $x_{ij}$  and  $e_{ij}$  are equal to ML-estimates of expected frequencies, computed under a specified loglinear model. This is the case for two-way contingency tables, and multiple tables formed from three-or-higher way tables. This result makes it possible to interpret the correspondence analysis solution in terms of loglinear models, and to circumvent the often arising interpretation problems with loglinear  $u$ -parameters. We will consider these applications more thoroughly at the end of this section. First we will discuss the situation for a two-way table, then for a three-way table, and finally for higher-way tables.

For a two-way contingency table with elements  $f_{ij}$  it is easy to show that values  $x_{ij}$  and  $e_{ij}$  are equal to ML-estimates of expected frequencies following models [12] and [1][2] respectively. For  $x_{ij}$  this is evident, since  $X = F$ : we take as  $X$ —the matrix to be analysed—the matrix of observed frequencies  $F$ , and in case of the saturated model [12] estimates of the expected frequencies are equal to the observed frequencies,  $\hat{m}_{ij} = f_{ij}$ . Values  $e_{ij}$  are equal to estimates of the expected frequencies following the independence model [1][2]: For this model  $\hat{m}_{ij} = f_{i*}f_{*j}/n$ , and since  $e_{ij} = x_{i*}x_{*j}/n$ , (see (2), and  $X = F$ , it follows that  $x_{i*} = f_{i*}$  and  $x_{*j} = f_{*j}$ . So, in case of a two-way contingency table the correspondence analysis solution can be interpreted in terms of the difference between the loglinear models [12] and [1][2]. Of course, this result is not very amazing. Without explicit reference to loglinear models, it is stated in almost all correspondence analysis literature that correspondence analysis decomposes the departure from independence of the row and column variable.

Results become more interesting in case of analysis of three-or-higher-way tables by means of multiple tables. First consider the three-way table  $F$  with frequencies  $f_{ijk}$ , from which (without loss of generality) the multiple table  $F^{1 \times (2 \times 3)}$  is constructed. We will

index the elements of  $X$  and  $E$  as  $x_{im}$  and  $e_{im}$ , where each distinct combination of variable 2 and 3 forms a different category of the interactive variable, which is indexed by  $m$ . Of course, since  $X = F^{1 \times (2 \times 3)}$ , elements  $x_{im}$  are equal to ML-estimates of expected frequencies for the saturated model [123]. Values  $e_{im}$  are equal to ML-estimates of expected frequencies for model [1][23]. These expected frequencies  $\hat{m}_{ijk}$  are estimated with (13),  $\hat{m}_{ijk} = f_{i**}f_{*jk}/n$ . We know that  $e_{im} = x_{i*}x_{*m}/n$ . Since  $X = F^{1 \times (2 \times 3)}$ , it follows that  $x_{i*} = f_{i**}$  and  $x_{*m} = f_{*jk}$  and consequently  $e_{im} = \hat{m}_{ijk}$ . So when the three-way matrix  $F$  is flattened to the multiple matrix  $F^{1 \times (2 \times 3)}$ , the correspondence analysis solution can be interpreted in terms of the difference between the loglinear models [123] and [1][23]. When multiple tables  $F^{2 \times (1 \times 3)}$  or  $F^{3 \times (1 \times 2)}$  would have been constructed, values  $e_{ij}$  would follow model [2][13] or [3][12] respectively.

When we have higher-way contingency table  $F$ , we can construct multiple table  $F^{A \times B}$ , where  $A$  and  $B$  are two groups of variables which together constitute the higher-way table, where  $A \cap B = 0$ . So, from the variables of group  $A$  and  $B$  two interactive variables are constructed. (When a group has only one variable, this variable is the new interactive variable.) In this case the correspondence analysis solution can be interpreted in terms of the difference between loglinear models  $[A \cup B]$  and  $[A][B]$ : elements  $x_{ab}$  of  $X$  are equal to elements  $f_{ab}$  of  $F^{A \times B}$ ; elements  $e_{ab}$  of  $E$  are equal to  $e_{ab} = x_{a*}x_{*b}/n$ . For model  $[A][B]$  expected frequencies  $m_{ab}$  are estimated with the formula  $\hat{m}_{ab} = f_{a*}f_{*b}/n$ , and since  $X = F^{A \times B}$ , we find that  $x_{a*} = f_{a*}$ ,  $x_{*b} = f_{*b}$ , and  $e_{ab} = \hat{m}_{ab}$ .

For three- and higher-way tables we can conclude that, when one analyses multiple tables with correspondence analysis, the correspondence analysis solution can be interpreted in terms of the difference between two loglinear models. We have also shown that, when one inspects the correspondence analysis solution, one does not see all the dependence in the data matrix, but only a part of it. For example, for the multiple table  $F^{1 \times (2 \times 3)}$  constructed from the three-way table  $F$ , the first-order interaction between Variables 2 and 3 does not influence the solution. One can use this fact in two ways when one has to decide which multiple table can best be chosen. Let us consider again a three-way table. Firstly, when some first-order interaction does *not* seem interesting, one can code the corresponding two variables interactively. Secondly, when one is especially interested in the relation of some variable with the other two, this variable should not be coded interactively with another variable. The correspondence analysis solution shows the two first-order interactions of this variable with the other two, and the second-order interaction.

Another implication of our results is that it is in some cases possible to use correspondence analysis and loglinear analysis complementary to each other. One can use the two techniques to find answers to different questions: loglinear analysis can answer the question of which variables are related: correspondence analysis can answer the question of how these variables are related, that is, which categories occur more often together than expected, and which categories less often. Or, to put it differently, loglinear analysis answers questions about interaction between variables on the "variable level," and correspondence analysis on the "category level." Not only for two-way tables but also for three- or higher-way tables, some of the computable Pearson goodness-of-fit measures can be decomposed with correspondence analysis. One circumvents problems with the interpretation of  $u$ -parameters in this way. For this reason we would advise one to do a loglinear analysis, and evaluate the acceptability of models with the  $X^2$ -statistic first. Secondly, we decide which multiple matrix or matrices one should choose, and thirdly we perform the correspondence analysis on these matrices. Correspondence analysis decomposes the values of the  $X^2$ -statistics, computed in the loglinear analysis. We will illustrate this procedure with an example.

*Example: Suicide Behavior*

A contingency table with data on suicide behavior is analysed. The number of successful suicides is arranged by the variables "age" (18 categories)  $\times$  "sex"  $\times$  "cause of death" (10 categories). In a cell one finds the number of times persons of a certain age and sex, committed suicide, using one of ten methods. The data are collected by the German Office for Statistics in Western Germany for the years 1974 to 1977, and can be found in a book on suicide prevention by Heuer (1979, Table 1). We have omitted one age-group (5–10 years old;  $n = 1$ ) and one cause of death ("died later because of suicide attempt";  $n = 5$ ) from the analyses because of their small marginal frequency.

In literature on suicide, official statistics are often said to be not completely reliable (Diekstra, 1981, pp. 63–84, Douglas, 1973, pp. 163–231; Heuler, 1979, pp. 63–72). The most important reason is that suicides are sometimes not reported as such, but as accidents, due for example to different definitions on suicide and attempts to conceal them. This will occur more often with certain methods and certain ages (especially with children). Apart from case histories, theory on suicide has for various reasons relied heavily on official statistics (Douglas, pp. 164–167). One reason is that a good theory on suicide should be capable to explain these statistics. A precise understanding of the structure of these statistics is therefore important.

Loglinear analysis is performed with the program BMDP4F (Dixon, 1981), with which all possible hierarchical models are examined. The results can be found in Table 2. All goodness-of-fit statistics are significantly different from zero. This is not surprising, since chi-squares are proportional to  $n$  ( $n = 53,210$ ). The three first order interactions seem to be important (which can be seen from the difference in  $G^2$  and  $X^2$  between all models with two first-order interactions, and the model with three first order interactions), and even the second order interaction cannot be neglected. We can conclude that there is good reason to investigate in which way the three variables are related.

Which multiple tables should be analysed with correspondence analysis? Here are some possible considerations. When one first-order interaction would have been less important, we could have used this as an argument to construct from the corresponding two variables the interactive variable; but this is not the case. There is no first order interaction in which we have a priori interest, so that consideration can not be used here. We are most interested in the relation of variable "cause-of-death" with the two background variables age and sex: in what way is the background related with the chosen cause of death? This points to the multiple table  $F^{M \times (A \times S)}$ .

Correspondence analysis of  $F^{S \times (M \times A)}$  shows us a one-dimensional solution, since "sex" has only two categories; we cannot benefit from the nice geometrical aspects of correspondence analysis in the one-dimensional case, and therefore we skipped this analysis. The analyses of  $F^{M \times (A \times S)}$  and  $F^{A \times (M \times S)}$  remain, of which the first one seems most interesting. We will represent results using symmetric normalization. The solutions are computed using APL.

Correspondence analysis of matrix  $F^{M \times (S \times A)}$  decomposes the value of  $X^2$ , which is 9995 for model  $[M][SA]$ , in 8 dimensions. The first two dimensions are shown in Figure 1. The singular values are .312 and .268, and they "explain" 52 percent and 38 percent of the  $X^2$  value respectively. These singular values can be interpreted as maximal canonical correlations between the background variables and the cause of death. The first order interaction between age and sex does not influence the solution. Roughly, the first dimension stresses the differences in behavior of men and women, the second dimension stresses the different use of methods by people of different ages. In Table 3 the contributions to the eigenvalues for the individual categories can be found, as well as the contributions of

Table 1: Suicide behavior: age by sex by cause of death

Labels for cause of death-categories:

1. Suicide by solid or liquid matter (MATT)
2. Suicide by toxification of gas at home (GASH)
3. Suicide by toxification with other gas (GASO)
4. Suicide by hanging, strangling, suffocating (HANG)
5. Suicide by drowning (DROW)
6. Suicide with guns and explosives (GUNS)
7. Suicide with knives etc. (STAB)
8. Suicide by jumping (JUMP)
9. Suicide with other methods (OTHE)

	MATT	GASH	GASO	HANG	DROW	GUNS	STAB	JUMP	OTHE	TOTAL
<u>Men</u>										
10-15	4	0	0	247	1	17	1	6	9	285
15-20	348	7	67	578	22	179	11	74	175	1461
20-25	808	32	229	699	44	316	35	109	289	2561
25-30	789	26	243	648	52	268	38	109	226	2399
30-35	916	17	257	825	74	291	52	123	281	2836
35-40	1118	27	313	1278	87	293	49	134	268	3567
40-45	926	13	250	1273	89	299	53	78	198	3179
45-50	855	9	203	1381	71	347	68	103	190	3227
50-55	684	14	136	1282	87	229	62	63	146	2703
55-60	502	6	77	972	49	151	46	66	77	1946
60-65	516	5	74	1249	83	162	52	92	122	2355
65-70	513	8	31	1360	75	164	56	115	95	2417
70-75	425	5	21	1268	90	121	44	119	82	2175
75-80	266	4	9	866	63	78	30	79	34	1429
80-85	159	2	2	479	39	18	18	46	19	782
85-90	70	1	0	259	16	10	9	18	10	393
90+	18	0	1	76	4	2	4	6	2	113
TOT	8917	176	1913	14740	946	2945	628	1340	2223	33828
<u>Women</u>										
10-15	28	0	3	20	0	1	0	10	6	68
15-20	353	2	11	81	6	15	2	43	47	560
20-25	540	4	20	111	24	9	9	78	67	862
25-30	454	6	27	125	33	26	7	86	75	839
30-35	530	2	29	178	42	14	20	92	78	985
35-40	688	5	44	272	64	24	14	98	110	1319
40-45	566	4	24	343	76	18	22	103	86	1242
45-50	716	6	24	447	94	13	21	95	88	1504
50-55	942	7	26	691	184	21	37	129	131	2168
55-60	723	3	14	527	163	14	30	92	92	1658
60-65	820	8	8	702	245	11	35	140	114	2083
65-70	740	8	4	785	271	4	38	156	90	2096
70-75	624	6	4	610	244	1	27	129	46	1691
75-80	495	8	1	420	161	2	29	129	35	1279
80-85	292	3	2	223	78	0	10	84	23	715
85-90	113	4	0	83	14	0	6	34	2	256
90+	24	1	0	19	4	0	2	7	0	57
TOT	8648	77	241	5637	1703	172	309	1505	1090	19382

Table 2: Loglinear analysis of data in table 1

Only hierarchical models with all main effects are reported. .1 is added to each cell for all analyses. Labels M, S, A denote variables method, sex and age.

<u>Model</u>	<u>df</u>	<u>G<sup>2</sup></u>	<u>X<sup>2</sup></u>
[M][S][A]	280	12337	12304
[M][SA]	264	10313	9995
[S][MA]	152	7780	7198
[A][MS]	272	6858	6522
[MS][MA]	144	2300	2253
[MA][SA]	136	5756	5369
[MS][SA]	256	4834	4519
[MS][MA][SA]	128	429	436

the first two dimensions to the squared distances of the categories to the origin. Note that for the methods with high marginal frequencies almost the complete distance is projected on the first two dimensions. Combining the results from Figure 1 and Table 3, interpretation becomes easy: JUMP is done more by women, MATT more by older women, and DROW by younger women. Men use more “violent” methods: HANG is done more by older men and boys—this last group almost exclusively commits suicide by HANG, and GASO and GUNS by younger men. (See Table 1 for an explanation of the cause-of-death labels.)

The correspondence analysis solution for matrix  $F^{A \times (M \times S)}$  can be found in Figure 2 and Table 4. The point for children from 10 to 15 years old is not plotted (coordinates are  $-.469, -2.878$ ) to make the structure of the other points more clear. This solution is not influenced by the first order interaction between method and sex. The first two singular values are .315 and .100, with respective “explained” percentages of  $X^2$  81 percent and 8 percent—the first dimension is very dominant. On this dimension a gradual shift in use of methods can be found from young persons to persons older than 65. Most women points lie on the left, most men points on the right: women commit suicide relatively more often when they are older. At younger age men predominantly commit suicide using methods as MATT, GASO, OTHE and GUNS. Men use HANG most often when they are children, middle-aged or old. Women use HANG and DROW especially when they are old. The second dimension is dominated by children up to 15 years (contribution is .55), for which the boys commit almost exclusively suicide by HANG (contribution is .60). We saw this result already in the previous analysis.

In this section we have tried to show how correspondence analysis can complement loglinear analysis by decomposing the  $X^2$ -statistics found with loglinear analysis. The correspondence analysis plots prove to be a very useful aid in the interpretation of loglinear interaction, especially when used together with tables of contributions of individual category points. Because of the large number of categories, and the presence of first- and

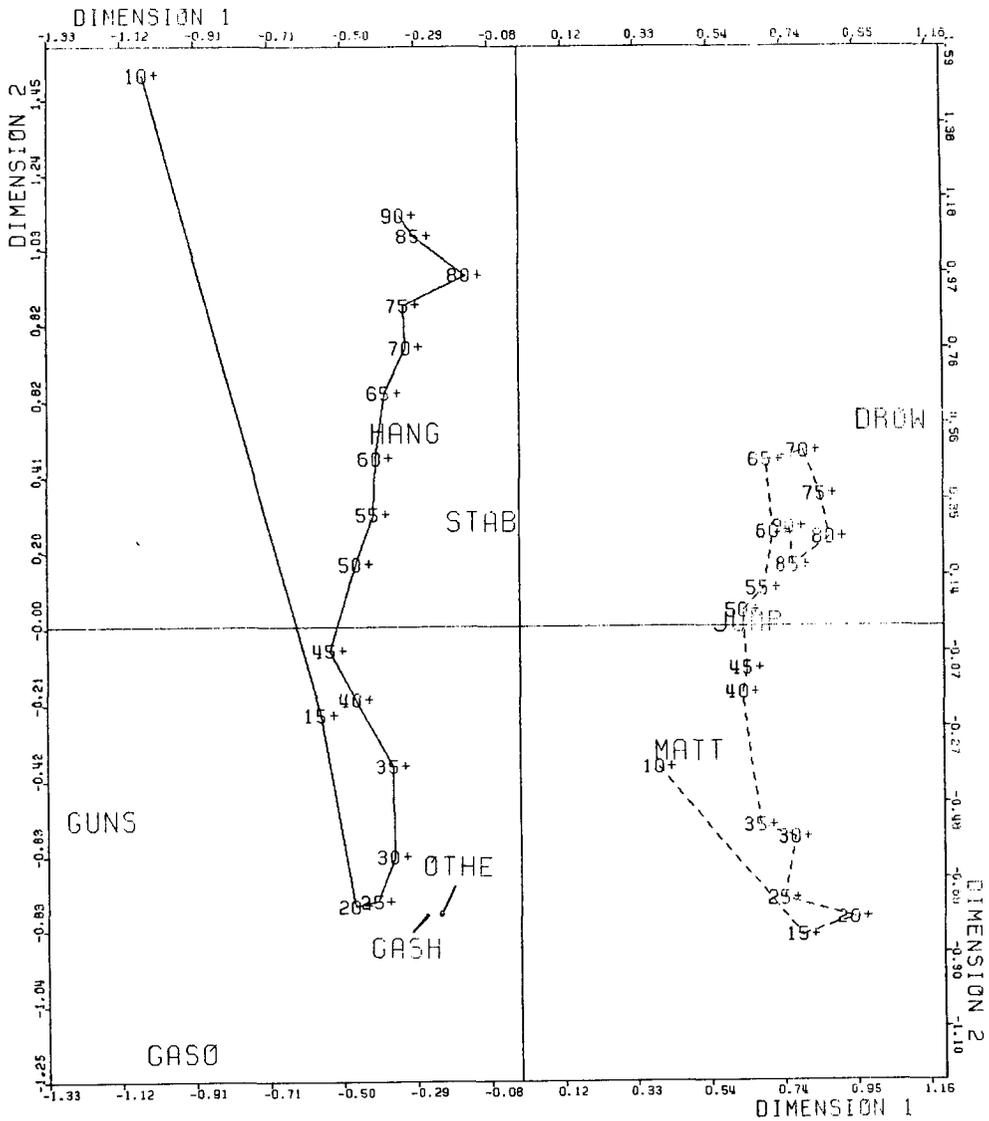


Figure 1: Analysis of  $F^{M \times (S \times A)}$ , first two dimensions. The line for men is solid. Singular values with relative contributions:  
 $\lambda_1 = .312 (.519)$ ,  $\lambda_2 = .268 (.381)$ .

second-order interaction effects, the interpretation of  $u$ -parameters would have been a cumbersome job.

### 5. Using a Generalization of Correspondence Analysis for the Decomposition of the Difference Between Other Loglinear Models

In the previous section it was shown that, when analyzing multiple tables, the correspondence analysis solution can be interpreted in terms of the difference between the saturated model (the observed frequencies) and a specific loglinear model. When the dif-

Table 3: Contributions of categories to eigenvalues, analysis  $F^{M \times (G \times A)}$

Contributions of methods to dimensions:		Proportion squared projected distance to the origin:	
MATT	.240 .142	.666	.290
GASH	.001 .009	.036	.384
GASO	.119 .207	.401	.514
HANG	.127 .404	.297	.695
DROW	.176 .058	.608	.148
GUNS	.261 .061	.824	.140
STAB	.001 .005	.052	.320
JUMP	.071 .000	.720	.000
OTHE	.005 .113	.048	.766
	1.000 1.000		

	Contributions of sex-age categories to dimensions:		Proportion squared projected distances to the origin:			
	Men	Women	Men	Women	Men	Women
10-15	.019 .046	.001 .001	.333 .586	.149 .122		
15-20	.028 .006	.021 .028	.545 .087	.390 .378		
20-25	.033 .106	.046 .039	.282 .653	.532 .328		
25-30	.024 .096	.028 .033	.238 .710	.485 .419		
30-35	.022 .079	.036 .023	.261 .692	.604 .285		
35-40	.028 .036	.037 .027	.449 .430	.599 .330		
40-45	.041 .009	.030 .003	.752 .120	.871 .061		
45-50	.056 .001	.037 .001	.920 .011	.871 .024		
50-55	.034 .006	.052 .000	.827 .100	.949 .004		
55-60	.020 .013	.047 .001	.602 .288	.908 .017		
60-65	.023 .035	.065 .010	.463 .515	.823 .090		
65-70	.021 .069	.061 .031	.272 .662	.647 .236		
70-75	.013 .089	.066 .028	.161 .798	.653 .199		
75-80	.009 .078	.056 .012	.132 .836	.774 .119		
80-85	.001 .051	.033 .003	.025 .920	.836 .055		
85-90	.002 .032	.009 .001	.076 .864	.554 .022		
90-+	.001 .010	.002 .000	.084 .812	.533 .057		
	1.000 1.000	1.000 1.000				

ference between such a model and the observed frequencies is significant, it makes sense to complement loglinear analysis with correspondence analysis. Unfortunately, it is possible to complement loglinear analysis only in a limited number of cases, for example in a three-way matrix when one wants to analyze the difference between the observed frequencies and one of the models [1][23], [2][13] or [3][12]. In this section we will describe a generalization of correspondence analysis proposed by Escofier (1983), which we will use to analyze the difference between other restricted loglinear models and the saturated model. At the end of this section we will illustrate this with two examples.

Escofier (1983) generalizes correspondence analysis by computing the singular value decomposition of the matrix  $S_r^{-1/2}(G_1 - G_2)S_c^{-1/2}$  instead of the matrix  $D_r^{-1/2}(X - E)D_c^{-1/2}$ , (3), to find row and column scores  $R$  and  $C$ , and singular values  $\Lambda$ . Here  $G_1$  and  $G_2$  are matrices of the same size, and  $S_r$  and  $S_c$  are diagonal matrices with weights for

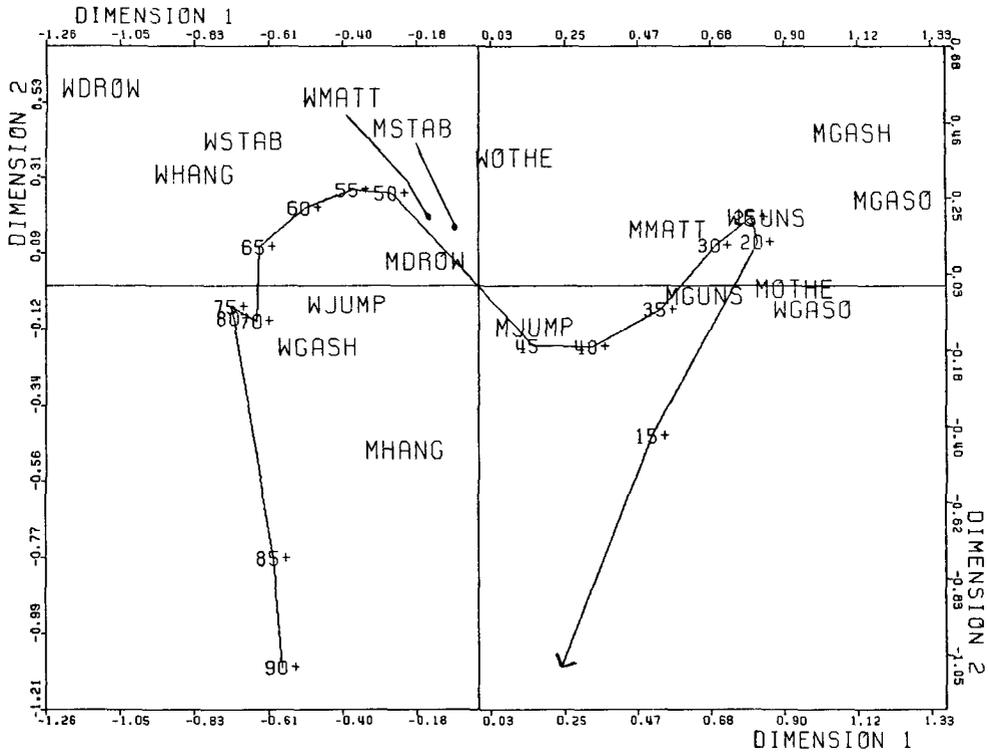


Figure 2: Analysis of  $F^{Ax(MxS)}$ , first two dimensions. Singular values with relative contributions:  $\lambda_1 = .315 (.808)$ ,  $\lambda_2 = .100 (.082)$

row and column categories. In contrast to classical correspondence analysis,  $S_r$ ,  $S_c$ ,  $G_1$  and  $G_2$  are not necessarily related in the way that  $D_r$ ,  $D_c$  and  $E$  are to the matrix  $X$ . So,

$$S_r^{-1/2}(G_1 - G_2)S_c^{-1/2} = U\Lambda V', \tag{16}$$

where  $U$ ,  $V$  and  $\Lambda$  follow the same restrictions as in (3). Row and column scores are found as—compare with (4):

$$R = S_r^{-1/2}U n_r^{1/2} \quad \text{and} \tag{17}$$

$$C = S_c^{-1/2}V n_c^{1/2},$$

where  $n_r = \text{trace } S_r$  and  $n_c = \text{trace } S_c$ . The transition formulas are

$$R = S_r^{-1}(G_1 - G_2)C A^{-1} \quad \text{and} \tag{18}$$

$$C = S_c^{-1}(G_1 - G_2)R A^{-1}.$$

The transition equations, (5), are more simple, since  $D_r^{-1}E C A^{-1} = 0$  and  $D_c^{-1}E' R A^{-1} = 0$ . The reconstitution formula is

$$G_1 - G_2 = S_r R A C' S_c n_r^{-1/2} n_c^{-1/2} \tag{19}$$

Comparable with (6), the chi-squared distance between row  $i$  and  $i'$ , and their approximations, can be computed as

$$(I_i - I_{i'}) S_r^{-1} (G_1 - G_2) S_c^{-1} (G_1 - G_2)' S_r^{-1} (I_i - I_{i'})' n = (\tilde{r}_i - \tilde{r}_{i'})' (\tilde{r}_i - \tilde{r}_{i'}). \tag{20}$$

Table 4: Contributions of categories to eigenvalues, analysis  $F^{Ax(M \times G)}$

Contributions of age categories to dimensions:				Proportion squared projected distance to the origin:	
10-15	.005	.547		.071	.860
15-20	.031	.069		.509	.115
20-25	.138	.010		.925	.007
25-30	.124	.023		.945	.017
30-35	.110	.009		.951	.008
35-40	.082	.005		.899	.005
40-45	.029	.025		.674	.060
45-50	.007	.026		.383	.143
50-55	.019	.063		.514	.174
55-60	.029	.050		.708	.123
60-65	.069	.040		.911	.055
65-70	.111	.010		.951	.009
70-75	.097	.008		.913	.007
75-80	.085	.002		.926	.002
80-85	.046	.002		.832	.005
85-90	.014	.073		.493	.263
90-+	<u>.003</u>	<u>.038</u>		<u>.361</u>	<u>.412</u>
	1.000	1.000		1.000	1.000

	Contributions of methods-sex categories to dimensions:				Proportion squared projected distances to the origin:			
	Men		Women		Men		Women	
MATT	.165	.043	.014	.070	.934	.025	.392	.207
GASH	.013	.006	.001	.000	.702	.035	.227	.010
GASO	.170	.021	.014	.000	.942	.012	.784	.001
HANG	.042	.603	.232	.107	.401	.583	.941	.044
DROW	.001	.001	.125	.101	.178	.011	.877	.072
GUNS	.077	.001	.007	.001	.871	.001	.513	.009
STAB	.000	.005	.009	.010	.033	.135	.734	.084
JUMP	.002	.004	.014	.001	.166	.020	.507	.003
OTHE	<u>.114</u>	<u>.000</u>	<u>.001</u>	<u>.027</u>	.886	.000	.061	.240
	1.000	1.000	1.000	1.000				

We will now discuss some principles important for the interpretation. A row point (or column point) is placed far from the origin when the corresponding row (or column) in the left term of (16) has high values. Two row points (or two column points) deviate in the same direction from the origin when the corresponding rows (or columns) have roughly the same values. A row point is placed near a column point when the corresponding value in the left term of (16) is high. When for  $G_1$  and  $G_2$  the marginal row or column frequencies are different, Escofier (1983) advises to take as weights for  $S_r$  and  $S_c$  the marginal frequencies of  $G_1$  or of  $G_2$  or of  $(G_1 + G_2)/2$ . When the marginal frequencies of the rows (or columns) of  $G_1$  and  $G_2$  are very different, the weighted average of the column points (or row points) is far away from the origin. Escofier indicates correctly that formulas (16)

to (20) are difficult to interpret in their most general form, that is when  $S_r$ ,  $S_c$ ,  $G_1$  and  $G_2$  are not related. Interpretation becomes easier when one of the following conditions holds:

1.  $G_1t = G_2t$ ,  $G_1$  and  $G_2$  have equal marginal row frequencies;
2.  $G_1't = G_2't$ ,  $G_1$  and  $G_2$  have equal marginal column frequencies;
3. as in Condition 1, but also  $G_1t = S_r t$ , as weights for the rows the marginal row frequencies are taken;
4. as in Condition 2, but also  $G_1't = S_c t$ , as weights for the columns the marginal column frequencies are taken;
5.  $G_2 = G_1 t t' G_1' / n$ , that is,  $G_1 = X$  and  $G_2 = E$ .

When Condition 1 holds,  $t'S_c C = 0$ : the weighted average for the column scores is zero for each dimension when the marginal row frequencies of  $G_1$  and  $G_2$  are equal. For Condition 2 comparable results hold for the row scores. When Condition 3 holds, (17) can be interpreted in terms of the barycentric principle: row points are in the weighted average of the column points (when for the simultaneous representation  $\bar{K}$  and  $C$  are taken as coordinates). A row point represents the difference between the profiles of the row in  $G_1$  and  $G_2$ . When Condition 4 holds, we find comparable results for the column points. When Condition 5 holds, Condition 1 and Condition 2 also hold. When in addition Condition 3 and Condition 4 also hold, we deal with classical correspondence analysis. Escofier (1983) shows that when only Condition 3 and Condition 4 both hold, the generalized correspondence analysis solution can be found with a classical correspondence analysis program analysing the matrix  $(G_1 - G_2 + E)$ . Comparing (3) and (16), this will be clear.

This generalization can be used for the analysis of residuals of various sorts of models. Our experience with the generalization of correspondence analysis is promising in cases that  $G_1$  and  $G_2$  have equal marginal frequencies—which often occurs, since it is not unusual in model fitting to condition on the marginals. Examples are Goodman's RC-model (1979) and quasi-symmetry models for mobility tables, confusion matrices, import-export tables, et cetera. For pairs of loglinear models, multiple tables with expected frequencies also have often the same marginals. When one or both marginals are unequal, this fact may dominate the solution, especially when the difference is large. In the sense of (19), in the first few dimensions first the difference between these marginals is reconstituted. One is usually not interested in the reconstitution of this part of the difference between  $G_1$  and  $G_2$ . This is also the case in our context, and therefore we will discuss here only examples for which Conditions 3 and 4 hold. Another important property is that a point represents the difference between the *profiles* of the category in  $G_1$  and  $G_2$ .

### Examples

For the suicide data, we will show two examples. The first example shows the difference between the saturated model (observed frequencies) and the model with two first order interactions  $[MA][SA]$ , using multiple table  $F^{M \times (S \times A)}$ . In the resulting solution the relation between sex and method will be shown; that is, how this relation is the same over ages (first order interaction) and how it differs over ages (second order interaction). The second example shows the difference between the models  $[MSA]$  and  $[MS][MA][SA]$ , using the same multiple table. Here the solution should be interpreted in terms of the second order interaction only.

The first two dimensions of the first example are plotted in Figure 3. Singular values are .305 and .095, and they explain 87 percent and 10 percent of the total inertia respectively. For the first dimension this solution is roughly the same as the solution in Figure 1: this dimension is dominated by the first order interaction between sex and method. The

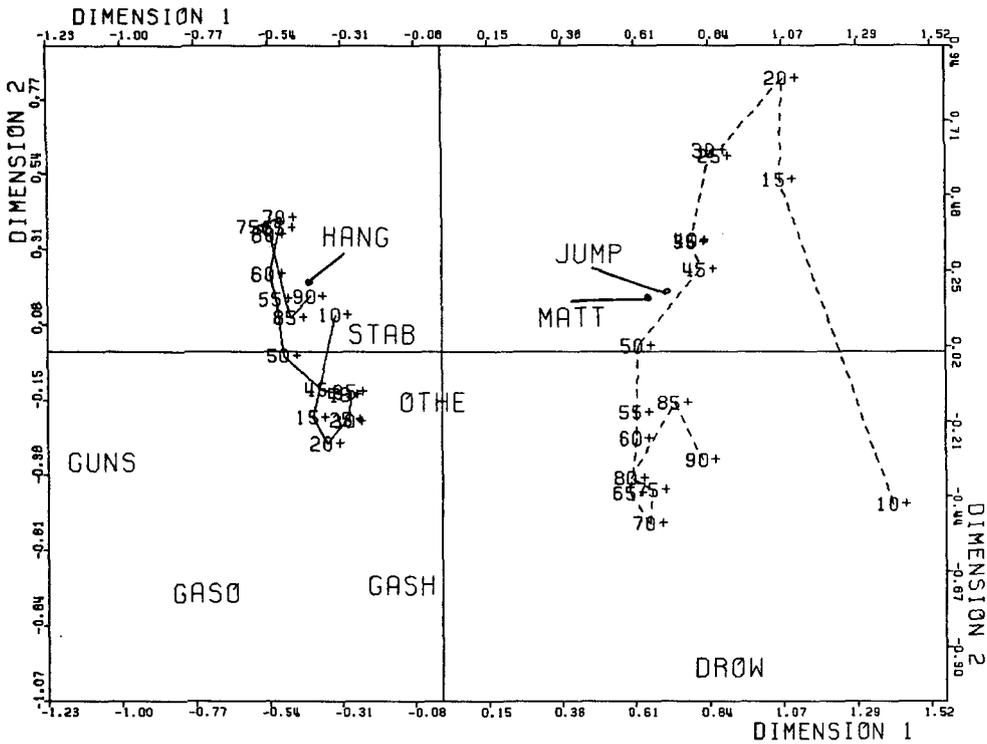


Figure 3: Analysis of  $F^{M \times (S \times A)}$ . Difference between  $[MSA]$  vs.  $[MA][SA]$ . Singular values  $\lambda_1 = .305 (.865)$ ,  $\lambda_2 = .095 (.085)$ .

second dimension is dominated by the second order interaction. This is clear from the fact that the age lines for men and women are reflected—roughly—in the origin: for example, apart from the first order interaction that women drown themselves more often than men, we can see that the second order interaction is that older women and younger men drown themselves more than older men and younger women. Contributions of individual categories can be computed in the same way as for classical correspondence analysis. For reasons of space we do not give them here.

Figure 4 shows the first two dimensions of the second example. Singular values are .056 and .040, which explain 49 percent and 26 percent of the total inertia. It makes sense to study the third dimension, but for reasons of space we skip the discussion of this part of the analysis. This solution should be interpreted in terms of the second order interaction only. Again, the men and women age line is roughly reflected in the origin. The point for women from 10 to 15 years old is not plotted (coordinates are 2.808 and .308) to make the structure of the other points more clear. A study of contributions of method categories shows that the first dimension is dominated by HANG and MATT, and the second dimension by DROW and JUMP. For example, part of the second order interaction is that solid and liquid matter is used more by younger and older women, and men of moderate age.

### 6. Conclusions

It is shown that correspondence analysis of multiple tables can be seen as the decomposition of the difference between two matrices, each following a specific loglinear

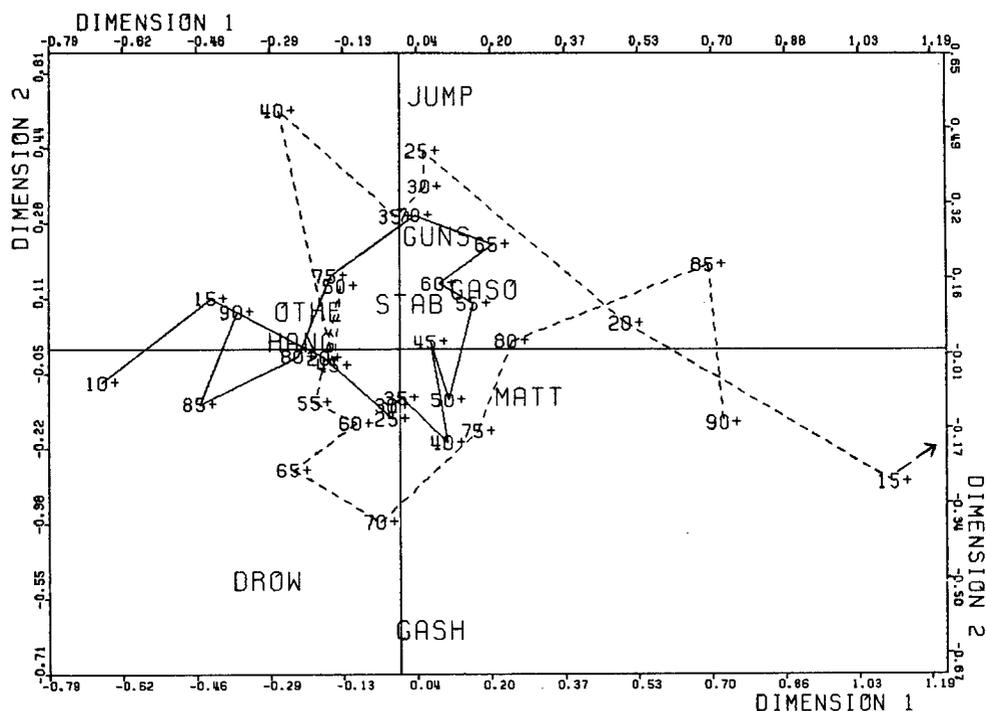


Figure 4: Analysis of  $F^{M \times (S \times A)}$ . Difference between [MSA] vs. [MS][MA][SA]. Singular values  $\lambda_1 = .056 (.491)$ ,  $\lambda_2 = .040 (.255)$ .

model. Apart from the fact that this is of theoretical interest, it can lead to a better understanding of what correspondence analysis shows us, and it can lead to a complementary use of loglinear analysis and correspondence analysis: instead of the cumbersome interpretation of  $u$ -parameters, one has the disposal of a nice geometrical representation of the data. This complementary relation is made more complete by the introduction of Escofier's generalization of correspondence analysis (1983).

We have shown that the analysis of a three-way table using multiple tables works out fine. Our experience is that this is usually the case, when marginal row and column frequencies of the multiple table are not too low. Low marginal row or column frequencies can cause instability of (parts of) the solution. When the number of variables is larger than 3, and the number of categories of the row and column variable becomes large, the interpretation of the correspondence analysis solution can become difficult. Computing weighted averaged points over the interactive variable (where the weighting is done over the marginal frequencies, and the averages are taken over the categories of the original variables) is often helpful in these cases. For example in the first example, 2 weighted sex-categories and 17 weighted age-categories can be computed.

#### References

- Agresti, A. (1983). A survey of strategies for modeling cross-classifications having ordinal variables. *Journal of the American Statistical Association*, 78, pp. 184-198.
- Andersen, E. B. (1980). *Discrete statistical models with social science applications*. Amsterdam: North Holland.
- Benzécri, J. P. (1973). *Analyse des données [The analysis of Data]* (2 vols.). Paris: Dunod.
- Benzécri, J. P. (1980). *Pratique de l'analyse des données [The Practice of Data Analysis]* (3 vols.). Paris: Dunod.

- Bishop, Y. M., Y., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge: MIT press.
- Bonett, D. G., & Bentler, P. M. (1983). Goodness-of-fit procedures for the evaluation and selection of loglinear models. *Psychological Bulletin*, 93(1), pp. 149–166.
- Daudin, J. J., & Trécourt, P. (1980). Analyse factorielle des correspondances et modèle log-linéaire: Comparaison des deux méthodes sur un exemple [Correspondence analysis and the loglinear model: Comparison of two methods on an example]. *Revue de Statistique Appliquée*, 28(1), 5–24.
- de Leeuw, J. (1971). *Canonical analysis of contingency tables* (RB-002-71). Leiden: R.U.L./F.S.W., Department of Datatheory.
- de Leeuw, J. (1983). Models and methods for the analysis of correlation coefficients. *Journal of Econometrics*, 22, 113–128.
- de Leeuw, J. (1984). *Canonical analysis of categorical data*. (Doctoral dissertation, University of Leiden, 1973). Leiden: DSWO-press.
- Dijkstra, R. F. W. (1981). *Over suicide* [about suicide]. Alphen a/d Rijn: Samson.
- Dixon, W. J. (Ed.). (1981). *BMDP statistical software 1981*. Berkeley: University of California Press.
- Douglas, J. D. (1973). *The social meanings of suicide*. Princeton: Princeton University Press.
- Escofier, B. (1983). Analyse de la différence entre deux mesures sur le produit de deux mêmes ensembles [Analysis of the difference between two measures on the product of two identical sets]. *Cahiers de l'Analyse des Données*, 3, 325–329.
- Fienberg, S. E. (1980). *The analysis of cross-classified categorical data* (2nd ed.). Cambridge: MIT Press.
- Fienberg, S. E., & Meyer, M. M. (1983). Loglinear models and categorical data analysis with psychometric and econometric applications. *Journal of Econometrics*, 22, 191–214.
- Gifi, A. (1981). *Non-linear multivariate analysis*. Leiden: R.U.L./F.S.W., Department of Datatheory. (2nd. ed. in press, Leiden: DSWO press).
- Goodman, L. A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association*, 74, 537–552.
- Goodman, L. A. (1981a). Association models and the bivariate normal for contingency tables with ordered categories. *Biometrika*, 68, 347–355.
- Goodman, L. A. (1981b). Association models and canonical correlation in the analysis of cross-classifications having ordered categories. *Journal of the American Statistical Association*, 76, 320–334.
- Greenacre, M. J. (1984). *Theory and applications of correspondence analysis*. London: Academic Press.
- Heiser, W. J., & Meulman, J. (1983). Analyzing rectangular tables by joint and constrained multi-dimensional scaling. *Journal of Econometrics*, 22, 139–167.
- Heuer, J. (1979). *Selbstmord bei Kinder und Jugendlichen* [Suicide of children and youth]. Stuttgart, Ernst Klett Verlag.
- Israëls, A. Z., & Sikkels, D. (1982). *Correspondence analysis and comparisons with other techniques*. Voorburg: Centraal Bureau voor Statistiek.
- Jacobs, L. (1971) *Adolescent suicide*. New York: Wiley.
- Kendall, M. G., & Stuart, A. (1973). *The advanced theory of statistics*, Vol. 2, 3rd. ed. London: Griffin.
- Kroonenberg, P. M. (1983). *Three-mode principal component analysis*. Leiden: DSWO-press.
- Leclerc, A. (1975). L'analyse des correspondances sur juxtaposition de tableaux de contingence [Correspondence analysis of concatenated contingency tables]. *Revue de Statistique Appliquée*, 23(3), 5–16.
- Nishisato, S. (1980). *Analysis of categorical data: Dual scaling and its applications*. Toronto: University of Toronto Press.
- Reynolds, H. T. (1977). *The analysis of cross-classifications*. Glencoe, IL: The Free Press.

Manuscript received 12/7/84

Final version received 5/15/85