ELSEVIER Computational Statistics & Data Analysis 24 (1997) 337–356

COMPUTATIONAL STATISTICS & DATA ANALYSIS

Three-factor association models for three-way contingency tables

Roberta Siciliano[a,*], Ab Mooijaart[b]

[a] *Department of Mathematics and Statistics, University of Naples, Monte S. Angelo, via Cintia, 80126 Napoli, Italy*

[b] *Department of Psychology, Leiden University, P.O. Box 9555, 2300 RB, Leiden, The Netherlands*

Received May 1995; revised June 1996

Abstract

A general formulation of association models is introduced for the analysis of three-way contingency tables. The two-factor and three-factor interaction matrices are decomposed into matrices of lower rank. In particular, the three-factor interaction is decomposed by the PARAFAC model. Various restricted models can be used to validate special assumptions for the data such as departures from conditional independence in the context of sets of contingency tables. The problem of identification is discussed. Two sets of data are analyzed to illustrate the versatility in the interpretation and the advantages of the models and methods developed here.

Keywords: Log-linear analysis; Log-trilinear decomposition; PARAFAC; Maximum likelihood estimation

1. Introduction

In the framework of multi-way cross-classifications, hierarchical log-linear modeling has proved to be a feasible and widely used method for validating specific assumptions concerning the relations among categorical variables (see for example Bishop et al. 1975; Haberman, 1978, 1979; Agresti, 1990).

Nevertheless, this approach is no longer useful in two-way tables when the log-linear model with no interaction does not fit to the data and only the saturated model fits to the data. A fruitful attempt to define the somewhat "in between"

* Corresponding author. E-mail: r.sic@dmsna.dms.unina.it.

PII S0167-9473(96)00067-7

independence model and the saturated model has been given by the association models for two-way tables (Goodman, 1979, 1985, 1986; Andersen, 1980). The main idea is to constrain the interaction parameters of a log-linear model to be equal to one or more bilinear terms and thus specifying the so-called RC(P)-association model. This model provides a meaningful description of the association in the data when the categories of the variables are ordered and the order is not necessarily known in advance. In fact, the order among the categories can be investigated by considering scores of the categories which are optimally found under the maximum likelihood criterion which is equivalent to the maximum information criterion (Gilula et al., 1988; Gilula and Haberman, 1988).

Similarly, for analyzing three-way tables hierarchical log-linear modeling cannot be feasible in two cases. In the first case, the model with no three-factor interaction does not fit to the data and the saturated model is the only model that obviously fits to the data. In the second case, the model of conditional independence does not fit to the data and the model with no three-factor interaction which fits to the data is the most parsimonious model. In order to define suitable models which are "in between" such restricted models, Becker (1989, 1992) introduced a general family of association model for three-way contingency tables enabling to deal with several hypotheses on the three-factor interaction among the variables. When appropriate constraints are specified upon the parameters several restricted versions can be specified and some of them have been already considered in literature; see, for instance, Agresti and Kezouch (1983), Goodman (1986), Gilula and Haberman (1988), Becker and Clogg (1989), Lauro and Siciliano (1989), Siciliano et al. (1990), Mooijaart (1992), Mooijaart and van der Heijden (1992). The main idea of Becker's approach consists of specifying conditional association between two variables given the level of the third variable which can be separated into unconditional two-way and three-way factor interactions.

In this paper we deal with this general family of association models and we present some restricted versions in an alternative formulation. The idea is to specify bilinear decompositions of two-factor interactions and trilinear decompositions of three-factor interactions. In particular, we use the PARAFAC/CANDECOMP model (Harshman, 1970; Carroll and Chang, 1970) for specifying the three-factor interaction. The PARAFAC model has been considered also by Kroonenberg (1983) and D'Ambra and Kiers (1990) to analyze the residuals of a log-trilinear model by a least squares loss-function.

We describe the estimation procedure by maximum likelihood method, the identification of the model parameters, the calculation of the degrees of freedom, the testing of various restricted models.

An important restricted model is considered to analyze the conditional association between two variables given a stratifying variable in a set of two-way tables. This model will allow us to answer questions such as does the association between the two variables differ across the groups? Furthermore, if there are differences, what is the pattern of their association? Is it possible to rank the groups from the one in which the variables have the highest association to the one in which they have the lowest association?

We also discuss the relationships and the differences with other kinds of models for conditional association proposed in literature; see in particular Becker and Clogg (1989), Xie (1992), Wong (1995).

In order to illustrate the versatility in the interpretation and the advantages of the models and methods developed here we present the results in two tables. Table 1 is a three-way table that concerns a sample of 16 236 children in the Netherlands. This sample is part of the data that were published in Meester and de Leeuw (1983). The variables are the scores on a Test for Intellectual Capacity (TIC),

Table 1
Three cross-classifications of education, intelligence test score and sex

	Sex													
	Boys							Girls						
TIC Education	1	2	3	4	5	6	7	1	2	3	4	5	6	7
DO	75	77	105	125	89	38	17	51	60	115	123	78	56	9
LBO	216	305	495	522	389	168	34	144	223	382	370	290	107	26
MAVO	67	144	267	368	339	194	54	60	134	288	424	442	266	72
MBO	51	84	239	345	301	208	65	75	167	320	458	428	258	72
HAVO	26	65	200	332	383	258	98	23	68	211	373	450	402	169
VWO	12	27	104	216	325	321	178	5	9	77	183	307	326	209

Table 2
Four cross-classifications of housing, influence, satisfaction and contact

Contact		Low			High		
Satisfaction		Low	Medium	High	Low	Medium	High
Housing	Influence						
Tower blocks	Low	21	21	28	14	19	37
	Medim	34	22	36	17	23	40
	High	10	11	36	3	5	23
Apartments	Low	61	23	17	78	46	43
	Medium	43	35	40	48	45	86
	High	26	18	54	15	25	62
Atrium houses	Low	13	9	10	20	23	20
	Medium	8	8	12	10	22	24
	High	6	7	9	7	10	21
Terraced houses	Low	18	6	7	57	23	13
	Medim	15	13	13	31	21	13
	High	7	5	11	5	6	13

Sex and the Level of Education attained after four years of secondary education. The TIC scores were recorded in seven classes where each class corresponds to an interval number of correct answers. The levels of education are: (1) dropped out (DO), (2) junior level of education for professions (LBO), (3) medium level of general education (MAVO), (4) senior level of education for professions (MBO), (5) high level of general education (HAVO) and (6) general education preparing for university (VWO). This table has been already analyzed by van der Heijden et al., (1992) with constrained latent budget analysis, by Mooijaart and van der Heijden (1992) with log-trilinear models and by Siciliano and van der Heijden (1994) with simultaneous latent budget analysis. Table 2 presents a four-way table concerning a sample of 1681 residents of 12 areas in Copenaghen. The table describes the interrelations among type of housing (tower blocks, apartments, atrium houses, terraced houses), degree of contact with other residents (low, high), feeling of influence on apartment management (low, medium, high) and satisfaction with housing conditions (low, medium, high). This table has been already considered by Agresti (1984) for logit-linear analysis.

2. Three-factor association models

2.1. Model definition

Let n_{ijk} denote the observed frequency in the (i, j, k)-th cell of an $I \times J \times K$ contingency table, with $i = 1, \ldots, I$; J; $k = 1, \ldots, K$. The corresponding expected frequency will be denoted by m_{ijk}. The (one-dimensional) marginals of the first, second and third variables of the table with proportions will be denoted as p_{i++}, p_{+j+}, and p_{++k}, respectively. The model we assume for the logarithm of these expected frequencies is

$$\log m_{ijk} = \mu + \mu_{1(i)} + \mu_{2(j)} + \mu_{3(k)} + \mu_{12(ij)} + \mu_{13(ik)} + \mu_{23(jk)} + \mu_{123(ijk)}, \quad (1)$$

where the μ parameters are the main effects or interaction parameters which satisfy the usual identifiability conditions as in log-linear analysis. Several restricted models can be specified, i.e., by setting one or more sets of parameters equal to zero.

We are interested in special cases of model (1) where the two-factor and the three-factor interactions are decomposed into matrices of lower rank. In the following, we first discuss how to decompose the three-factor interactions and then how to decompose the two-factor interactions.

2.2. Modeling the three-factor interactions

There are several ways for decomposing the three-factor interactions. Becker (1989, 1992) discussed a general approach for modeling the three-factor interaction by extending some results of Goodman (1986). This can be written as

$$\mu_{123(ijk)} = \sum_{s=1}^{S} \lambda_{s(k)}^{12(3)} \, x_{is(k)}^{1(3)} \, y_{js(k)}^{2(3)}, \quad (2)$$

where the parameters $x_{is(k)}^{1(3)}$ and $y_{js(k)}^{2(3)}$ are the row and column scores of the conditional table $I \times J$ given a particular layer k of the third variable, and the parameters $\lambda_{s(k)}^{12(3)}$ are the intrinsic associations between variables 1 and 2 given the level of variable 3.

We choose a special case of (2) that can be obtained by restricting the parameters $x_{is(k)}^{1(3)}$ and $y_{js(k)}^{2(3)}$ to be homogeneous over the K tables, i.e., $x_{is(k)}^{1(3)} = x_{is}$ and $y_{js(k)}^{2(3)} = y_{js}$, and rewriting $\lambda_{s(k)}^{12(3)}$ into a multiplicative factor such as $\lambda_{s(k)}^{12(3)} = \lambda_s z_{ks}$. So we can write

$$\mu_{123(ijk)} = \sum_{s=1}^{S} \lambda_{123(s)} x_{is} y_{js} z_{ks}. \tag{3}$$

This decomposition is known in psychometrics as PARAFAC (parallel factor analysis) and it has been developed so far for continuous data, in particular for three-way matrices that include, observations of J numerical variables on I individuals in K different occasions (Harshman, 1970; Kroonenberg, 1983; Harshman and Lundy, 1984a, b). However, the trilinear decomposition (3) has never been discussed in details and neither examples have been given in the literature when dealing with categorical data. The main problem that arises in case of categorical data is that ranks conditions are more crucial than in case of continuous data. Indeed, in continuous data we choose a rank that is (much) lower than the dimension of the three-way table. This problem will be discussed in Section 3.

We impose the following centering and weighting restrictions upon the parameters in (3):

$$\sum_{i=1}^{I} p_{i++} x_{is} = \sum_{j=1}^{J} p_{+j+} y_{js} = \sum_{k=1}^{K} p_{++k} z_{ks} = 0,$$

and

$$\sum_{i=1}^{I} p_{i++} x_{is}^2 = \sum_{j=1}^{J} p_{+j+} y_{js}^2 = \sum_{k=1}^{K} p_{++k} z_{ks}^2 = 1.$$

Notice that we only fix the scales for the parameters x_{is}, y_{js} and z_{ks}, we do not restrict parameters of x_{is}, y_{js} and/or z_{ks} to be orthogonal for different s. A problem of the PARAFAC model may be the existence of degenerate solutions, see Kruskal et al. (1989). A typical degeneracy is that subsequent solutions for $s = 1,2, \ldots ,$ are very similar. To overcome such degeneracies one can assume orthogonality restrictions for one or two sets of the parameters of x_{is}, y_{js} and/or z_{ks}.

Another special case of (2) is known in psychometrics as TUCKER3 model (Tucker, 1964, 1966; Kroonenberg, 1983)

$$\mu_{123(ijk)} = \sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{u=1}^{U} \lambda_{123(stu)} x_{is} y_{jt} z_{ku}. \tag{4}$$

A problem with this decomposition is that the identification of the parameters is not simple. Another problem is that many factors should be considered for the interpretation of the three-factor interactions. Note that the decompositions (3) and (4) are strictly related in the sense that for example we can derive (3) from (4) by imposing that all λ's are zero except for $s = t = u$.

Other special cases of (2) can be considered to model the conditional association between two variables given the level of a stratifying variable; see for instance Becker and Clogg (1989), Xie (1992) and Wong (1994). A restricted version of model (2) for conditional association will be considered in Section 2.4.

2.3. *Modeling the two-factor interactions*

In three-dimensional tables there are three types of two-factor interactions. In the most general form we can write these interactions as

$$\mu_{12(ij)} = \sum_{p=1}^{P} \lambda_{12(p)}\, a_{ip}\, b_{jp}, \qquad \mu_{13(ik)} = \sum_{q=1}^{Q} \lambda_{13(q)}\, c_{iq}\, d_{jq},$$

$$\mu_{23(jk)} = \sum_{r=1}^{R} \lambda_{23(r)}\, e_{jr}\, f_{kr}. \tag{5a–c}$$

In order to identify the model we impose the following centering restrictions:

$$\sum_{i=1}^{I} p_{i++}\, a_{ip} = \sum_{i=1}^{I} p_{i++}\, c_{iq} = \sum_{j=1}^{J} p_{+j+}\, b_{jp} = \sum_{j=1}^{J} p_{+j+}\, e_{jr}$$

$$= \sum_{k=1}^{K} p_{++k}\, d_{kq} = \sum_{k=1}^{K} p_{++k}\, f_{kr} = 0,$$

and the following weighting restrictions:

$$\sum_{i=1}^{I} p_{i++}\, a_{ip}\, a_{ip'} = \sum_{j=1}^{J} p_{+j+}\, b_{jp}\, b_{jp'} = \delta^{pp'},$$

$$\sum_{i=1}^{I} p_{i++}\, c_{iq}\, c_{iq'} = \sum_{k=1}^{K} p_{++k}\, d_{kq}\, d_{kq'} = \delta^{qq'},$$

$$\sum_{i=1}^{J} p_{+j+}\, e_{jr}\, e_{jr'} = \sum_{k=1}^{K} p_{++k}\, f_{kr}\, f_{kr'} = \delta^{rr'},$$

where $\delta^{ss'} = 1$ if $s = s^1$, and $\delta^{ss'} = 0$ otherwise. We have considered a weighting system given by the one dimensional marginals of the table with proportions. Other weights can also be assumed without changing the estimates of the expected frequencies (see, e.g., Becker and Clogg, 1989).

The final model that we will consider in the following can be defined by (1), (3) and (5). This model is more general than the log-trilinear model discussed by Choulakian (1988) since we do not restrict the scores of the categories of a variable to be the same in the two-factor decomposition and in the three-factor decomposition. In fact, the parameters a_{ip}, b_{jp}, c_{iq}, d_{kq}, e_{jr}, f_{kr}, x_{is}, y_{js}, z_{ks}, can be conceived as scores of the categories of the three variables. Notice that in this way the category i may have three different scores: P times by the scores a_{ip}, Q times by the scores c_{ip} and S times by the scores x_{ip}. For a more restricted model we can impose, that the scores of a variable are equal in all two-factor and three-factor interactions: $a_{ip} = c_{ip} = x_{ip}$, $b_{jp} = e_{jp} = y_{jp}$ and $d_{kp} = f_{kp} = z_{kp}$. In this more restricted model the order for each of these matrices is equal, i.e., $P = Q = R = S$. A proposal for such a model was given by Choulakian (1988). In our approach we deal with the less restrictive model that the two-factor and three-factor interactions can have different decompositions specified by (5a–c) and (3), respectively, and thus they can have different orders. However, by imposing the restrictions that the categories have the same scores in all decompositions we can deal with more restricted models as well.

2.4. Three-factor interaction model for a set of two-way tables

A restricted version of the model (1), (3), (5) can also be fruitfully used to analyze the conditional association between two variables given a stratifying variable of a set of two-way tables. The model of conditional independence between the variables 2 and 3 given the variable 1 can be specified when $\mu_{123(ijk)} = \mu_{23(ij)} = 0$. We analyze the departure from conditional independence with the following model

$$\log m_{ijk} = \mu + \mu_{1(i)} + \mu_{2(j)} + \mu_{3(k)} + \mu_{12(ij)} + \mu_{13(ik)} + \sum_{s=1}^{S} \lambda_{23(s)} e_{js} f_{ks} + \sum_{s=1}^{S} \lambda_{23(s)} x_{is} e_{js} f_{ks}. \tag{6}$$

In this model we specify that the scores of the categories of the second and third variables are equal in the two-factor and three-factor interaction terms. In this way, the parameters x_{is} in the three-factor term refer to the differences among the categories of the stratifying variable. The two-factor interactions $\mu_{12(ij)}$ and $\mu_{13(ik)}$ can also be decomposed according to the Eqs. (5a) and (5b).

Two important special cases can be considered. When $\lambda_{23(s)} = 0$ for all s, we obtain the model of conditional independence. When x_{is} is equal to a constant for all s and i, the association is homogeneous over the I groups.

The model (6) can also be written as

$$\log m_{ijk} = \mu + \mu_{1(i)} + \mu_{2(j)} + \mu_{3(k)} + \mu_{12(ij)} + \mu_{13(ik)} + \sum_{s=1}^{S} \lambda_{23(s)} (x_{is} + 1) e_{js} f_{ks}. \tag{7}$$

Note that multiplying $\lambda_{23(s)}$ with a constant and dividing $(x_{is} + 1)$ by the same constant, results in the same expected frequency m_{ijk}. Therefore, the scores x_{is} of one category may be fixed to some constant values and this allows to simplify the interpretation of the intergroup differences.

Our approach deals with trilinear terms to specify the conditional association. A more general formulation is provided by Becker and Clogg (1989) through the definition of a multiple-group model for simultaneous analysis of a set of two-way contingency tables. As special case Xie (1992) has proposed a log-multiplicative model to compare mobility tables.

In our formulation we provide the scores of the categories of the stratifying variable. This turns out to be particularly convenient for testing intergroup differences. As an example, we can fix the restrictions $x_{1s} = x_{2s}$ for all s in order to test that the three-factor interaction due to the stratifying factor does not change in the first group with respect to the second group.

Following the general formulation of Becker and Clogg (1989), Wong (1995) defined a special class of conditional association models with linear and quadratic constraints on the intrinsic association parameters for the particular case that the stratifying variable indexes temporal order.

3. Identification of the parameters

In this section we discuss the identification of the parameters. In addition we compute the maximum number of independent parameters. This number is important in computing the number of degrees of freedom which we need in testing the models.

In Table 3 we show the number of independent parameters to be estimated in the unrestricted and restricted models as well as the maximum rank of the matrices of two-factor and three-factor interactions.

First of all we collect all parameters in matrices. For instance, the parameters a_{ip} will be collected in the matrix $\boldsymbol{A}$ of order $I \times P$, the parameters b_{jp} will be collected in the matrix $\boldsymbol{B}$ of order $J \times P$. Analogously, we define the matrices $\boldsymbol{C}$, $\boldsymbol{D}$, $\boldsymbol{E}$, $\boldsymbol{F}$, $\boldsymbol{X}$, and $\boldsymbol{Z}$.

3.1. Two factor interactions

As an example we investigate the identification of the parameters a_{ip} and b_{jp}. Since $I > P$ and $J > P$, matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are of rank P. Taking into account the centering restrictions, the number of free elements in $\boldsymbol{A}$ and $\boldsymbol{B}$ are $P(I-1) + P(J-1) = P(I+J-2)$. However, the product $\boldsymbol{A}\boldsymbol{B}'$ is equal to $\boldsymbol{A}\boldsymbol{S}\boldsymbol{S}^{-1}\boldsymbol{B}'$ for any nonsingular matrix $\boldsymbol{S}$ of order $P \times P$. So matrix $\boldsymbol{A}$, and so the elements a_{ip}, can be identified up to a nonsingular linear tranformation $\boldsymbol{S}$. It follows that matrix $\boldsymbol{B}$, and so the elements b_{jp}, can be identified up to a nonsingular linear tranformation $\boldsymbol{S}^{-1}$.

The total number of independent parameters is $(I+J-2)P - P^2$. Note that if $P = \min(I-1, J-1)$, the number of parameters in the reduced case is equal to the number of parameters in the unrestricted case, namely $(I-1, J-1)$. Obviously, analogous results hold for the parameters c_{iq}, d_{kq}, e_{jr}, f_{kr}.

Table 3
Number of independent parameters to be estimated in three-factor association models ($I \leq J \leq K$)

Parameter	Number of independent parameters		
	Unrestricted model	Restricted model	Maximum rank
μ	1	1	1
$\mu_{1(i)}$	$(I-1)$	$(I-1)$	$(I-1)$
$\mu_{2(j)}$	$(J-1)$	$(J-1)$	$(J-1)$
$\mu_{3(k)}$	$(K-1)$	$(K-1)$	$(K-1)$
$\mu_{12(ij)}$	$(I-1)(J-1)$	$P(I+J-2)-P^2$	$(I-1)$
$\mu_{13(ik)}$	$(I-1)(K-1)$	$Q(I+K-2)-Q^2$	$(I-1)$
$\mu_{23(jk)}$	$(J-1)(K-1)$	$R(J+K-2)-R^2$	$(I-1)$
$\mu_{123(ijk)}$	$(I-1)(J-1)(K-1)$	For $I=2$: $(J+K-2)S-S^2$	$(I-1)$
		For $I>2$ and $S<J$: $(I+J+K-5)S$	Unknown

3.2. Three factor interactions

In this paragraph we assume that the variables are ordered in such a way that $I \leq J \leq K$ holds. Besides the definition of the matrices $\boldsymbol{Y}$ and $\boldsymbol{Z}$, with orders $(J \times S)$ and $(K \times S)$, respectively, we define diagonal matrices where the diagonal of these matrices consist of row i of matrix $\boldsymbol{X}$. So there are I different matrices $\boldsymbol{D}_i$. Now we can formulate I different three-factor interaction matrices as $\boldsymbol{U}_i = \boldsymbol{Y}\boldsymbol{D}_i\boldsymbol{Z}'$, of order $(J \times K)$, which has elements u_{ijk}, for $j = 1, \ldots, J$ and $k = 1, \ldots, K$. By the centering restrictions for the elements x_{is}, one matrix $\boldsymbol{D}_i$ is redundant because it is just a simple function of the other $\boldsymbol{D}$ matrix. Therefore, we only investigate $\boldsymbol{U}_i$, for $i = 1, \ldots, I-1$. Furthermore, we define $S_Y = \min(S, J-1)$, which is the maximum rank of matrix $\boldsymbol{Y}$. Two cases have to be distinguished, Case 1: $I = 2$, and Case 2: $I > 2$.

Case 1: $I = 2$.

The only matrix equation we have to investigate is $\boldsymbol{U}_1 = \boldsymbol{Y}\boldsymbol{D}_1\boldsymbol{Z}'$ (here that $\boldsymbol{U}_2 = \boldsymbol{Y}\boldsymbol{D}_2\boldsymbol{Z}'$ holds is redundant). Because of the fact that $\boldsymbol{U}_1$ consists of a product of three matrices, matrix $\boldsymbol{D}_1$ is arbitrary, so without any loss of generality we may fix matrix $\boldsymbol{D}_1$ to the identity matrix. So we have $\boldsymbol{U}_1 = \boldsymbol{Y}\boldsymbol{Z}$. Now let $\boldsymbol{W}$ be a nonsingular matrix of order $S_Y \times S_Y$, then we also can write $\boldsymbol{U}_1 = \boldsymbol{Y}\boldsymbol{W}\boldsymbol{W}^{-1}\boldsymbol{Z}$. So matrix $\boldsymbol{Y}$, and so the elements y, can be identified up to a nonsingular linear tranformation $\boldsymbol{W}$. It follows that matrix $\boldsymbol{Z}$, and so the elements z, can be identified up to a nonsingular linear tranformation $\boldsymbol{W}^{-1}$.

The total number of independent parameters is $(J-1)S_Y + (K-1)S_Y - S_Y^2 = (J+K-2)S_Y - S_Y^2$. The maximum rank is $(J-1)$ which results in a total number of independent parameters $(J-1)(K-1)$. This number is equal to the total number of independent parameters in the unrestricted case.

Case 2: $I > 2$.

Here we have to investigate the matrices $\boldsymbol{U}_i = \boldsymbol{Y}\boldsymbol{D}_i\boldsymbol{Z}'$, for $i = 1, \ldots, I-1$. Obviously, one matrix $\boldsymbol{D}_i$ is arbitrary, and so may be fixed as, e.g., the identity matrix. Here we let $\boldsymbol{D}_1$ be the identity matrix. So we have the following equations: $\boldsymbol{U}_1 = \boldsymbol{Y}\boldsymbol{Z}'$, $\boldsymbol{U}_2 = \boldsymbol{Y}\boldsymbol{D}_2\boldsymbol{Z}'$, ... , $\boldsymbol{U}_{I-1} = \boldsymbol{Y}\boldsymbol{D}_{I-1}\boldsymbol{Z}'$. If we define new matrices $\boldsymbol{Y}$ and $\boldsymbol{Z}$ as $\boldsymbol{Y}^+ = \boldsymbol{Y}\boldsymbol{W}$ and $\boldsymbol{Z}^+ = \boldsymbol{Z}(\boldsymbol{W}')^{-1}$, then we have $\boldsymbol{U}_1 = \boldsymbol{Y}^+\boldsymbol{W}^{-1}\boldsymbol{W}\boldsymbol{Z}^{+\prime}$. Furthermore, we have $\boldsymbol{U}_i = \boldsymbol{Y}^+\boldsymbol{W}^{-1}\boldsymbol{D}_i\boldsymbol{W}\boldsymbol{Z}^{+\prime}$, so a new matrix $\boldsymbol{D}_1$ becomes $\boldsymbol{D}_i^+ = \boldsymbol{W}^{-1}\boldsymbol{D}_i\boldsymbol{W}$. Because matrix $\boldsymbol{D}_i^+$ must be diagonal, it follows that matrix $\boldsymbol{W}$ is diagonal. (Note the difference with case 1, where such a restriction is not necessary.) So matrix $\boldsymbol{Y}$, and so the elements y, can be identified up to a scaling of the columns. It follows that matrix $\boldsymbol{Z}$, and so the elements z, can be identified up to the corresponding inverse scaling.

The total number of independent parameters is $(J-1)S_Y + (K-1)S_Y + (I-2)S_Y - S_Y = (I+J+K-5)S_Y$.

In determining the maximum rank, which will yield a saturated model for the three factor interactions, we encounter a well known problem: the maximum rank of three dimensional tables is unknown. This means that in the case of $I \geq 3$ we cannot determine a value of S which will give a saturated model for the three factor interactions. An analogous problem exists in latent class analysis where which number of latent classes will give a perfect model is also unknown i.e., a model which describes the data perfectly. In general, we can state that the concept of rank of a three or higher dimensional table is much more complicated than in two dimensional tables and the mathematical properties of these tables are still unresolved completely. This problem was discussed by, e.g., Kruskal (1977), ten Berge et al. (1988), Kruskal (1989) and ten Berge (1991).

4. The maximum likelihood estimation and testing

4.1. The estimation procedure

The parameters of the model (1), (3), (5) will be estimated by the maximum likelihood method. Under Poisson sampling the kernel of the log likelihood function can be written as

$$\log L = \sum_{i=1}^{I}\sum_{j=2}^{J}\sum_{k=1}^{K} n_{ijk}\log m_{ijk} - \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} m_{ijk}. \tag{8}$$

This likelihood function should be optimized with respect to the parameters. The two- and three-factor interactions in the equation above are functions of the unknown parameters, see Eqs. (3) and (5). Estimation of these parameters will be done by some alternating estimation procedure.

But first, two remarks have to be made.

First, the centering restrictions are not important in the algorithm. For instance, let the scores x_{is} be shifted with a constant. Then (3) can be written as

$$\mu_{123(ijk)} = \sum_{s=1}^{S} \lambda_{123(s)} (x_{is} + \delta_s) y_{js} z_{ks} - \sum_{s=1}^{S} \lambda_{123(s)} \delta_s y_{js} z_{ks}. \tag{9}$$

In the first term on the right-hand side of (9) the scores x_{is} are shifted with a constant, whereas the second term on the right-hand side is in fact a two-factor interaction between the second and third variable. Depending on restrictions which hold for the two-factor interactions $\mu_{23(jk)}$ a simple correction may be applied without changing the expected frequencies m_{ijk} as given in (1). For instance, when the two-factor interaction parameters $\mu_{23(jk)}$ are completely free, the two-factor interaction $\mu_{23(jk)}$ has to be corrected with the second term of (9). Similarly, the same correction has to be applied when the scores of variables 2 and 3 are equal in both the two and three-factor interactions, i.e., $S = R$, and furthermore $e_{js} = y_{js}$, and $f_{ks} = z_{ks}$. Obviously, analogous results hold for a shift of the scores y_{js} and z_{ks} for each s.

The second remark we make is that the estimation of the λ parameters is not necessary in the first part of the algorithm, because the λ parameters are in fact just the result of some specific scaling of the scores. For instance, according to (2a) we can write $\mu_{12(ij)} = \sum_p \lambda_{12(p)} a_{ip} b_{jp} = \sum_p a^*_{ip} b^*_{jp}$ where $a^*_{ip} = a_{ip}/\lambda^{1/2}_{12(p)}$. So it is sufficient to estimate a^*_{ip} and b^*_{jp}, and afterwards decomposing $\sum_p a^*_{ip} b^*_{jp}$ to find a_{ip} and b_{jp} which meet the centering and weighting restrictions mentioned before. Such a decomposition of $\sum_p a^*_{ip} b^*_{jp}$ can be carried out by a generalized singular value decomposition (see Greenacre, 1984).

The basic algorithm now runs as follows: Fix all the parameters for the scores, except one set of parameters, for example x_{i1}. Find a better estimate of the elements of these parameters, and repeat the whole procedure but now by improving the elements of some other set of parameters. Repeat the whole procedure till no change of the elements occur.

As an example we show how to estimate the elements x_{is}. The method we use here is the unidimensional Newton–Raphson method, as was proposed by Goodman (1986); see also Becker (1990).

The first and the second derivatives of the log likelihood function (see (8)) with respect to the parameters x_{is}, are

$$l'(x_{is}) \equiv \frac{\partial \log L}{\partial x_{is}} = \sum_{j=1}^{J} \sum_{k=1}^{K} (n_{ijk} - m_{ijk}) y_{js} z_{ks}, \tag{10a}$$

$$l''(x_{is}) \equiv \frac{\partial^2 \log L}{\partial x_{is}} = - \sum_{j=1}^{J} \sum_{k=1}^{K} m_{ijk} y^2_{js} z^2_{ks}. \tag{10b}$$

According to the unidimensional Newton–Raphson method a new update of the scores x_{is} can be obtained by

$$x^+_{is} = x_{is} - \frac{l'(x_{is})}{l''(x_{is})}, \tag{11}$$

where the superscript " + " denotes the updated score x_{is}. So the updated score x_{is} can be written as

$$x_{is}^{+} = x_{is} + \frac{\sum_{j=1}^{J} \sum_{k=1}^{K} (n_{ijk} - m_{ijk}) y_{js} z_{ks}}{\sum_{j=1}^{J} \sum_{k=1}^{K} m_{ijk} y_{js}^2 z_{ks}^2}. \tag{12}$$

Analogous formulae can be derived for other parameters. Standard errors of parameter estimates can be calculated easily by resampling methods such as jackknife and bootstrap.

The estimation method above can easily be generalized to models in which scores of categories of a variable are equal in different interactions. For instance, if $a_{is} = c_{is} = x_{is}$, some terms have to be added to the formulae for the first and second derivatives as given in (10a) and (10b). (These additional terms are very easy to derive.) Then applying (11) and (12) give update formulae which are analogous to the formulae given above. In order to deal with linear restrictions we refer to Siciliano et al. (1993).

4.2. Testing the model

Models can be tested by the likelihood ratio test, i.e., after the algorithm has been converged we compute

$$G^2 = -2\{\sum_i \sum_j \sum_k n_{ijk} \log m_{ijk} - \sum_i \sum_j \sum_k n_{ijk} \log n_{ijk}\}.$$

Under the assumption that we have the "correct" model the G^2 value is chi-square distributed with the number of degrees of freedom depending on the given model.

It is common practice to use differences of test statistics corresponding to different nested models in order to compare models. Such a method can be used to select a suitable model. However, generally it holds that such differences of (conditional) test statistics do not have a chi-square distribution.

5. Examples

5.1. Analysis of the data in Table 1

We analyze the data in Table 1. Table 4 gives a summary of some models with their corresponding test statistics. In Table 4 "u" means that the corresponding interactions are unrestricted. The numbers 0, 1 and 2 mean that the corresponding interaction marix has rank 0, 1 and 2, respectively. Note that the maximum rank of the interaction matrices (1,2) and (1,3) is 1. So for these interactions "unrestricted" is equal to rank 1. The sign " * " means that some special restrictions are imposed.

The table shows that M_1, the standard log-linear model with no three-factor interactions, does not fit. Thus, all models with restricted two-factor interactions and no three-factor interactions will not fit as well. We need to model the three-factor interaction and we start considering the most parsimonious model such as

Table 4
Models and test statistics for the data in Table 1

Model	Interaction				G^2	d.f.	Significance
	$\mu_{12(ij)}$	$\mu_{13(ik)}$	$\mu_{23(jk)}$	$\mu_{123(ijk)}$			
M_1	u	u	u	0	61.62	30	s
M_2	u	u	u	1	19.91	20	ns
M_3	u	u	1	1	45.71	40	ns
M_4	u	u	1*	1*	73.46	50	s
M_5	u	u	2*	2*	42.02	40	ns

the model M_2, with unrestricted two-factor interactions and only one component for the three-factor interactions, i.e., $S = 1$; this model does fit. Therefore, we try to restrict further this model. Model M_2 can be restricted further by imposing restrictions on the interaction matrix of variables 2 and 3. The table shows that model M_3, the model in which the rank of the interaction matrix of variables 2 and 3 is just 1 instead of the maximum rank 5, does fit the data adequately.

An interesting question is the following: we find an interaction between Test for Intellectual Capacity (TIC) and Final Educational Level; however, are the patterns of this interaction equal for boys and girls, more specifically, does this pattern of interaction for boys and girls only differ in strength?

For answering this question, we consider the model of conditional association (6), where it is specified that the scores of the categories of the second and third variable are equal in the two and three-factor interaction terms. The parameters x_{is} in the three-factor interaction term refer to the differences between the sexes. From model M_1 we know that these parameters cannot be 0, because a model in which the three-factor interactions vanish does not fit. In Table 3 model M_4 and M_5 refer to the model above with $S = 1$ and 2, respectively. From the test statistics we see that model M_5 does fit the data and the two dimensions have a nice interpretation. The solutions of the parameters e_{js}, f_{ks} and x_{is} are given in Table 5. The λ parameters are 0.403 and 0.028 for dimensions 1 and 2, respectively.

Figs. 1 and 2 give the plots of the solutions for the category scores of the two variables. The first dimension gives the ordering of the categories of the variable Test for Intellectual Capacity illustrating the Intelligence, from low intelligence to high intelligence. This ordering is just in line with the ordering of the Educational levels of the different school types. The second dimension shows the extremes against the intermediate levels of intelligence as well as a difference between the Vocational Categories (in particular MBO) and 'drop out', the general education lying in between. The second dimension shows the well-known "intensity dimension" as often is the case for the analysis of categorical data. The only exception seems to be category 6 of variable Intelligence that has a low value with respect to the second dimension.

Table 5
Scores of the categories according to model M_4 (data in Table 1)

TIC	Dimension		Education	Dimension		Sex	Dimension	
1	– 2.231	0.455	DO	– 1.502	1.322	Boys	0.000	0.000
2	– 1.591	1.663	LBO	– 1.180	– 0.242	Girls	0.239	– 1.444
3	– 0.682	– 1.422	MAVO	– 0.313	0.646			
4	– 0.121	– 0.398	MBO	0.301	– 1.880			
5	0.398	0.924	HAVO	0.613	0.702			
6	1.139	– 0.712	VWO	1.757	0.502			
7	1.851	1.197						

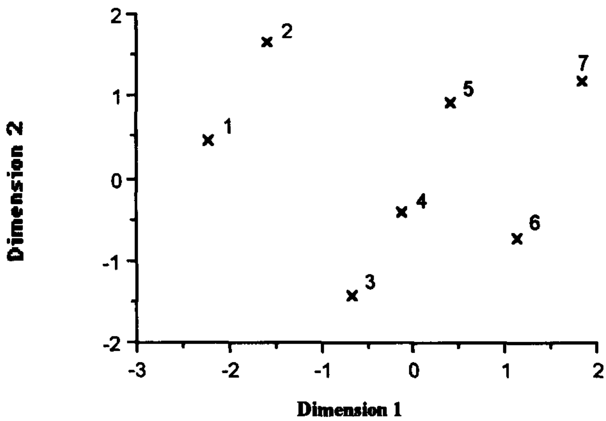

Fig. 1. TIC scores.

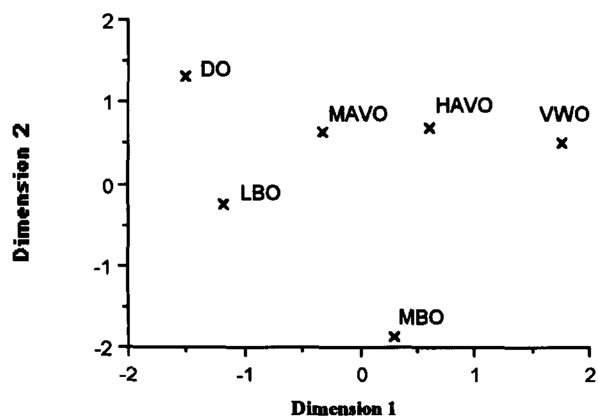

Fig. 2. Final educational level.

In Table 6 we give the three-factor interaction scores of the variables for the first dimension. Because the interaction parameters for the boys are set equal to zero, we give the interaction scores for the girls only.

Table 6
Three-factor interactions for dimension 1: girls only

TIC Education	1	2	3	4	5	6	7
DO	0.32	0.23	0.10	0.02	– 0.06	– 0.17	– 0.27
LBO	0.25	0.18	0.08	0.01	– 0.05	– 0.13	– 0.21
MAVO	0.07	0.05	0.02	0.00	– 0.01	– 0.03	– 0.06
MBO	– 0.07	– 0.05	– 0.02	0.00	0.01	0.03	0.05
HAVO	– 0.13	– 0.09	– 0.04	– 0.01	0.02	0.07	0.11
VWO	– 0.38	– 0.27	– 0.12	– 0.02	0.07	0.19	0.31

The scores in this table should be interpreted as follows: a positive score means that relatively more girls than boys fall in the specific combination of row and column category. Remember that the scores of boys are zero. On the other hand a negative score means just the opposite. The most interesting property of the table is the nice ordering of the scores over the columns and the rows. This can be interpreted as follows.

In the lowest intelligence categories we see that relatively more girls than boys fall in the categories Drop Out, LBO and MAVO, and on the other hand we see, for this intelligence group, that relatively less girls than boys fall in the categories MBO, HAVO and VWO.

In the highest intelligence categories we see that relatively less girls than boys fall in the categories Drop Out, LBO and MAVO, and on the other hand we see, for this intelligence group, that relatively more girls than boys fall in the categories MBO, HAVO and VWO.

In Table 7 we give the three-factor interaction scores of the variables for the second dimension. Also here we give the interaction scores for the girls only.

Obviously, because the second dimension is less important than the first dimension, the scores in Table 7 are closer to zero than the scores in Table 6. From Table 7 we see a less nice pattern of the scores over the columns and rows. The interpretation should be based here on specific combination of categories of the two variables Intelligence and Final Educational Level. For instance, in category 2 of Intelligence, more girls than boys fall in the category MBO, whereas in category 3 of Intelligence, more boys than girls fall in the category MBO.

By our analysis we have shown that a simple model, in which the three-factor interactions differ from the two-factor interactions by just one parameter for each dimension, results in a nice solution which fits the data well.

5.2. *Analysis of the data in Table 2*

We analyse the data in Table 2. As a matter of fact, this is a four-way contingency table. From the four-way contingency table we can pass to a three-way table either by adding up with respect to one variable or by cross-classifying the categories of

Table 7
Three-factor interactions for dimension 2: girls only

TIC Education	1	2	3	4	5	6	7
DO	− 0.02	− 0.09	0.08	0.02	− 0.05	0.04	− 0.06
LBO	0.00	0.02	− 0.01	0.00	0.01	− 0.01	0.01
MAVO	− 0.01	− 0.04	0.04	0.01	− 0.02	0.02	− 0.03
MBO	0.04	0.13	− 0.11	− 0.03	0.07	− 0.05	0.09
HAVO	− 0.01	− 0.05	0.04	0.01	− 0.03	0.02	− 0.03
VWO	− 0.01	− 0.03	0.03	0.01	− 0.02	0.01	− 0.02

two variables to form a "compound variable". Alternatively, we can analyse separately three-way tables conditional to different levels of one variable. We have analysed the data in table 2 considering all these approaches. In Table 8 the first column denotes the table to be analysed with a capital letter, the remaining columns denote the different types of models with their corresponding test statistics.

The contingency Table A describes the interrelations among type of housing, feeling of influence on apartment management and satisfaction with housing conditions. We have obtained Table A from the data in Table 2 by adding up with respect to the variable degree of contact with other residents. For this table both the standard log-linear model M_1 of conditional independence and the model M_2 with no-three-factor interaction do not fit to the data. Instead, the model M_3 with unrestricted two-factor interactions and one component for the three-factor interactions, i.e. $S = 1$, does fit. Model M_3 can be restricted further by imposing restrictions on the two-factor interactions. In particular, we consider the model of conditional association between variables 2 (Housing) and 3 (Influence) given variable 1 (Satisfaction). The model M_4 that refers to the model of conditional association with $S = 1$ does fit to the data. The solutions of the parameters are given in Table 9. The λ parameter is equal to 0.168. Note that the model M_5 in which we consider a reduced-rank decomposition of the two-factor interactions does not fit to the data. As a result, the model M_4 is the most parsimonious model that fits adequately to the data. Table 8 shows the ordering of the categories of both the variables Influence and Satisfaction from low to high. Table 8 also shows that there is a difference in living in tower blocks and atrium houses with respect to living in apartments and terraced houses. In Table 10 we give the three-factor interaction scores of the variables. Because the interaction parameters for the residents with low satisfaction with housing conditions are set equal to zero, we give the interaction scores for medium and high satisfaction only.

The scores in this table should be interpreted as follows: a positive score means that relatively more residents with medium or high satisfaction fall in the specific combination of row and column category. On the other hand a negative score means just the opposite. As an example, among all residents with low influence on

Table 8
Models and test statistics for the data in Table 2

Contingency table	Model	Interaction				G^2	d.f.	Signif.
		$\mu_{12(ij)}$	$\mu_{13(ik)}$	$\mu_{23(jk)}$	$\mu_{123(ijk)}$			
A	M_1	u	u	0	0	34.10	18	s
Housing × Influence × Satisfaction	M_2	u	u	u	0	25.89	12	s
Dimensions: 4 × 3 × 3	M_3	u	u	u	1	8.64	5	ns
	M_4	u	u	1*	1*	15.73	10	ns
	M_5	1	1	1*	1*	31.34	13	s
B	M_1	u	u	0	0	20.23	18	ns
Housing × Influence × Satisfaction	M_2	u	u	u	0	14.04	12	ns
Low contact	M_3	u	u	u	1	6.78	5	ns
Dimensions: 4 × 3 × 3	M_4	u	u	1*	1*	9.97	10	ns
	M_5	1	1	1*	1*	11.54	13	ns
Housing × Influence × Satisfaction	M_1	u	u	0	0	29.22	18	s
High contact	M_2	u	u	u	0	21.76	12	s
Dimensions: 4 × 3 × 3	M_3	u	u	u	1	6.62	5	ns
	M_4	u	u	1*	1*	13.42	10	ns
	M_5	1	1	1*	1*	34.28	13	s
D : Housing × Influence ×	M_2	u	u	u	0	51.91	30	s
(Satisfaction * contact)	M_6	u	u	2*	2*	29.95	23	ns
Dimensions: 4 × 3 × 6								

Table 9
Scores of the categories using model M_4 for the contingency table A of Table 8

Satisfaction		Housing		Influence	
Low	0.000	Tower blocks	1.606	Low	− 1.269
Medium	− 0.550	Apartments	− 0.780	Medium	0.745
High	− 1.507	Atrium houses	0.438	High	0.857
		Terraced houses	− 0.661		

apartment management we see that relatively more resident with medium or high satisfaction live in tower blocks and atrium houses; on the other hand, we see, for residents with low influence, that relatively less residents with medium or high satisfaction live in apartments and terraced houses. Instead, among all residents with medium or high influence on apartment management we see that relatively more residents with medium or high satisfaction live in apartments and terraced houses; on the other hand, we see, for residents with medium or high influence that relatively less residents with medium or high satisfaction live in tower blocks and atrium houses.

Table 10
Three-factor interactions using model M_4 for the contingency table A of Table 8

Influence Housing	Low	Medium	High
Satisfaction: medium			
Tower blocks	0.188	− 0.111	− 0.127
Apartments	− 0.092	0.054	0.062
Atrium houses	0.051	− 0.030	− 0.035
Terraced houses	− 0.078	0.045	0.052
Satisfaction: high			
Tower blocks	0.516	− 0.303	− 0.348
Apartments	− 0.251	0.147	0.169
Atrium houses	0.141	− 0.083	− 0.095
Terraced houses	− 0.212	0.125	0.143

We have also analysed the interrelations among the variables Housing, Influence and Satisfaction given the degree of contact, respectively, low and high degree, with other residents. In Table 8 we give the summary of some models fitted to such tables denoted by B and C. In particular, we see that for residents with low degree of contact all models fit to the data whereas for residents with high degree of contact only models M_3 and M_4 fit adequately to the data. This shows that the hypothesis of conditional independence between the variables Housing and Influence given the variable Satisfaction cannot be rejected for residents with both low and high degree of contact with other residents.

The last way to analyse the data in Table 2 is to consider the three-way table that cross-classify the variables Housing, Influence and the compound variable formed with the variables Satisfaction and Contact. For this, denoted by C in Table 8 the model M_6 that refers to the model of conditional association with $S = 2$ does fit to the data. Table C has a different dimension with respect to the previous tables, in particular the stratifying variable has now more categories; therefore we need at least two dimensions to explain the interrelations among the variables. We do not give the solutions of this model because the final interpretation does not change with respect to the results that we have previously found for Table A.

Acknowledgements

The authors are indebted to the unknown referees for helpful comments on a previous draft of this paper. The research of the first author was supported by MURST60% and CNR n.92.1872.P research funds.

References

Agresti, A., *Analysis of Ordinal Categorical Data* (Wiley, New York; 1984).

Agresti, A., *Categorical Data Analysis* (Wiley, New York, 1990).

Agresti, A. and A. Kezouh, Association models for multidimensional cross-classifications of ordinal variables. *Communication in Statistics, Part A – Theory and Method,* **12** (1983) 1261–1276.

Andersen E.B., *Discrete Statistical Models with Social Science Application* (North-Holland, Amsterdam, 1980).

Becker, M.P., Models for the analysis of association in multivariate contingency tables, *J. Amer. Statist. Assoc.,* **84**, (1989) 1014–1019.

Becker, M.P., Maximum likelihood estimation of the RC(M) association model, *Appl. Statist.,* **39** (1990) 152–166.

Becker, M.P., Exploratory analysis of association models using loglinear models and singular value decompositions, *Comput. Statist. Data Anal.,* **13** (1992) 253–267.

Becker, M.P. and C.C. Clogg, Analysis of sets of two-way contingency tables using association models. *J. Amer. Statist. Assoc.,* **84** (1989) 142–151.

Bishop, Y.M.M., S.E. Fienberg, and P.W. Holland, *Discrete Multivariate Analysis* (MIT Press, Cambridge, MA, 1975).

Carroll, J.D. and J.J. Chang, Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart–Young" decomposition, *Psychometrika,* **35** (1970) 283–319.

Choulakian, V., Exploratory analysis of contingency tables by log-linear formulation and generalizations of correspondence analysis, *Psychometrika,* **53** (1988) 235–250 (Errata p. 593).

D'Ambra, L. and H. Kiers, Analysis of log-trilinear models for a three-way contingency table using Parafac/Candecomp. in: *Atti delle Giornate di Studio: Classificazione e Analisi dei Dati,* Pescara, 11–12 ottobre (1990) 101–113.

Gilula, Z. S.J. and Haberman, The analysis of multivariate contingency tables by restricted canonical and restricted association models, *J. Amer. Statist. Assoc.,* **83** (1988) 760–771.

Gilula, Z., A.M. Krieger and Y. Ritov, Ordinal association in contingency tables: some interpretative aspects, *J. Amer. Statist. Assoc.,* **83** (1988) 540–545.

Goodman, L.A., Simple models for the analysis of association in cross-classifications having ordered categories, *J. Amer. Statist. Assoc.,* **74** (1979) 537–552.

Goodman, L.A., The analysis of cross-classified data having ordered and/or unordered categories: association models, correlation models, and asymmetry models for contingency tables with or without missing entries, *Ann. Statist.,* **13** (1985) 1069.

Goodman, L.A., Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables, *Internat. Statist. Rev.,* **54** (1986) 243–309.

Greenacre, M.J., *Theory and Applications of Correspondence Analysis*: (Academic Press, New York, 1984).

Haberman, S.J., *Analysis of Qualitative Data. Vol. 1: Introductory Topics.* (Academic Press, New York, 1978).

Haberman, S.J., *Analysis of Qualitative Data. Vol. 2: New Developments* (Academic Press, New York, 1979).

Harshman, R.A., Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis, *UCLA Working Papers in Phonetics,* **16** (1970) 1–84.

Harshman, R.A. and M.E. Lundy., The PARAFAC model for three-way factor analysis and multidimensional scaling, in: H.G. Law, C.W. Snyder, J.A. Hattie, and R.P. McDonald (Eds.), *Research Methods for Multimode Data Analysis* (Praeger, New York, (1984a) 122–215.

Harshman, R.A. and M.E. Lundy, Data preprocessing and the extended PARAFAC model, in: H.G. Law, C.W. Snyder, J.A. Hattie, and R.P. McDonald (Eds.), *Research Methods for Multimode Data Analysis* (Praeger, New York, 1984b) 216–284.

Kroonenberg, P.M., *Three-Mode Principal Component Analysis* (DSWO Press, Leiden, 1983).

Kruskal, J.B., Three-way arrays: rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics, *Linear Algebra Appl.*, **18** (1977) 95–138.

Kruskal, J.B., Rank, decompositions, and uniqueness for 3-way and *N*-way array, in: R. Coppi and S. Bolasco (Eds), *Multiway Data Analysis* (North-Holland, Amsterdam, 1989) 7–18.

Kruskal, J.B., R.A. Harshman and M.E. Lundy, How 3-MFA data can cause degenerate PARAFAC solutions, among other relationships. in: R. Coppi and S. Bolasco (Eds), *Multiway Data Analysis.* (North-Holland, Amsterdam, 1989).

Lauro, N.C. and R. Siciliano, Exploratory methods and modelling for contingency tables analysis: an integrated approach, *Statistica Applicata. Ital. J. Appl. Statist.*, **1** (1989) 532.

Meester, A. and J. de Leeuw, *Intelligence, Social Milieu and the School Career* Department of Data Theory, Leiden, 1983 (in Dutch).

Mooijaart, A. (1992). Three-mode association models, *Italian J. Appl. Statist.*, **4**.

Mooijaart, A. and P.G.M. van der Heijden, Log-trilinear models for three-way contingency tables, Presented at SMABS-92, Nijmegen, 1992.

Siciliano, R., N.C. Lauro and A. Mooijaart, Exploratory approach and maximum likelihood estimation of models for non-symmetrical analysis of two-way multiple contingency tables, *Compstat'90* (1990) 157–162.

Siciliano, R., A. Mooijaart and P.G.M. van der Heijden, A probabilistic model for nonsymmetric correspondence analysis and prediction in contingency tables, *J. Ital. Statist. Soc.*, **2** (1993) 106.

Siciliano, R. and P.G.M. van der Heijden, Simultaneous latent budget analysis of a set of two-way tables with constant-row-sum data. *Metron*, **53** (1994) 1–2.

ten Berge, J.M.F., H.A.L. Kiers, and J. de Leeuw, Explicit CANDECOMP/PARAFAC solutions for a contrived $2 \times 2 \times 2$ array of rank three, *Psychometrika*, **53** (1988) 579–584.

ten Berge, J.M.F., Kruskal's polynomial for $2 \times 2 \times 2$ arrays and a generalization to $2 \times N \times N$ array, *Psychometrika*, **56** (1991) 631–636.

Tucker, L.R., The extension of factor analysis to three-dimensional matrices, in: H. Gullikson, and N. Frederiksen, (Eds.), *Contributions to Mathematical Psychology*, (Holt, Rinehart and Winston, New York, 1964) 110–119.

Tucker, L.R., Some mathematical notes on three-mode factor analysis, *Psychometrika*, **31** (1966) 279–311.

van der Heijden, P.G.M., A. Mooijaart and J. de Leeuw, Constrained latent budget analysis, in: P. Marsden (Ed.), *Sociological Methodology*, Vol. 22, (Blackwell, Cambridge, 1992) 279–320.

Wong, R.R.S.K., Extensions in the use of log-multiplicative scaled association models in multiway contingency tables, *Sociol. Methods and Res.*, **23** (1995) 507–538.

Xie, Y., The log-multiplicative layer effect model for comparing mobility tables, *Amer. Socio. Rev.*, **57** (1992) 380–395.