

Exploratory analysis of association models using loglinear models and singular value decompositions

Mark P. Becker

Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109-2029, USA

Abstract: A method of approximating maximum likelihood fits of multivariate association models is explored. The proposed procedure is based on singular value decompositions of matrices of interaction terms from hierarchical loglinear models estimated by the method of maximum likelihood. Potential applications include use in exploratory data analyses, in screening procedures to select models to be estimated by maximum likelihood, and in generating initial estimates for algorithms which compute maximum likelihood estimates. Two examples are used to illustrate features of the estimation procedure and its application to model screening.

Keywords: Association models, Correspondence analysis, Exploratory data analysis, Loglinear models, Model screening, Singular value decomposition, Three-factor interaction.

1. Introduction

During the past two decades loglinear modelling has become the cornerstone of contingency table analysis as practiced by statisticians in much of the English speaking world. More recently, there has been much interest in broadening the class of models used for contingency table analysis by employing multi-linear terms to explicitly model interaction; see e.g., Goodman (1979, 1985, 1986). Such models are generally referred to as association models, and much of their popularity is due to the fact that they frequently provide a more parsimonious and meaningful description of the association in data sets than is possible with the usual loglinear models.

In this paper we present a simple estimation procedure for use with the family of multivariate association models described in Becker (1989a). This family of models is reviewed in Section 2. Related work on multivariate association models includes Goodman (1979, 1986), Clogg (1982a,b), Agresti and Kezouh (1983), Gilula and Haberman (1988), Khoudraji (1988), Becker and Clogg (1989a) and Green (1989a,b), among others. The estimation procedure, which is based on singular value decompositions (SVDs) of matrices that have maximum likelihood (ML) estimates of interaction terms from loglinear models as their entries, is outlined in Section 3. Readers not familiar with the SVD are

referred to Greenacre (1984, appendix). The proposed estimation procedure is not suggested as a replacement for maximum likelihood estimation, but rather as a means of obtaining approximations of ML estimates that may be used in

Table 1
(source: Andersen 1980, Table 5.22)

Social Rank	Education	Income (kr)			
		0-10.000	10.000-20.000	20.000-30.000	over 30.000
I and II	1	9.00	21.00	62.00	104.00
		6.16	25.34	64.12	96.28
		6.03	25.73	66.46	96.59
	2	2.00	8.00	15.00	59.00
		1.41	6.18	15.18	62.69
		1.53	5.93	15.21	63.34
	3	1.00	20.00	45.00	48.00
		3.58	15.05	41.22	51.47
		3.60	15.58	40.53	51.31
	4	2.00	8.00	13.00	16.00
		3.32	10.20	14.22	16.30
		3.01	9.39	13.08	15.65
III	1	46.00	228.00	126.00	39.00
		53.12	221.93	131.00	45.00
		52.08	219.76	136.21	44.73
	2	8.00	24.00	19.00	14.00
		5.74	25.59	14.66	13.91
		6.59	25.29	15.55	14.62
	3	49.00	186.00	132.00	43.00
		47.35	202.14	129.13	36.88
		48.20	206.58	128.94	36.88
	4	47.00	162.00	45.00	12.00
		46.92	146.52	47.62	12.49
		45.30	139.88	46.75	12.63
IV	1	30.00	141.00	86.00	10.00
		32.21	139.14	78.33	12.41
		31.65	138.35	78.11	12.06
	2	2.00	13.00	7.00	6.00
		3.29	15.16	8.28	3.62
		3.52	14.01	7.84	3.47
	3	45.00	171.00	104.00	13.00
		39.11	172.57	105.15	13.85
		39.44	175.13	99.57	13.39
	4	38.00	102.00	29.00	5.00
		33.44	107.94	33.46	4.05
		35.01	112.00	34.10	4.33

Table 1 (continued)

Social Rank	Education	Income (kr)			
		0–10.000	10.000–20.000	20.000–30.000	over 30.000
V	1	33.00	170.00	67.00	4.00
		35.70	159.85	71.19	4.21
		34.49	157.01	72.50	4.22
	2	1.00	5.00	0.00	2.00
		1.10	5.27	2.28	0.37
		1.10	4.55	2.08	0.35
	3	40.00	176.00	81.00	3.00
		37.96	173.70	83.72	4.11
		37.78	174.70	81.23	4.12
	4	54.00	181.00	55.00	2.00
		56.59	189.41	46.44	2.10
		57.66	192.09	47.82	2.29

exploratory analyses and as initial estimates for algorithms that compute ML estimates. Other authors have also suggested applying the SVD to matrices of parameters estimated by the method of maximum likelihood. Caussinus and Falguerolles (1987), for example, have done so in the context of quasi-symmetry

Table 2
(Source: Haberman 1978, Table 3.14)

Sex of respondent	Husband's highest degree	Wife's highest degree		
		Less than High School diploma	High School diploma or junior college	At least bachelor's degree
Male	Less than High School diploma	135	60	1
	High School diploma or junior college degree	43	151	19
	Bachelor's degree	4	35	12
	Graduate Degree	2	24	23
Female	Less than High School diploma	124	63	1
	High School diploma or junior college degree	39	219	18
	Bachelor's degree	1	24	26
	Graduate Degree	0	17	14

The quality of the approximation is investigated using Table 1; a $4 \times 4 \times 4$ cross-classification of 3,289 Danish residents between the ages of 14 and 30 according to father's income, level of education, and father's social rank. The top entry in each cell is the observed count; the middle and bottom entries are defined in Section 3.2. Estimates of parameters pertaining to association are also compared, both graphically and numerically, with those obtained using maximum likelihood. In Section 4, we demonstrate how the proposed methodology can be applied to model screening. Some features of this application are illustrated using Table 2; a cross-classification of 1,055 married couples from the 1974 General Social Survey according to sex of respondent, highest degree attained by the husband, and highest degree attained by the wife.

There has been much interest in comparing and contrasting the correspondence analysis approach to contingency table analysis with the association models based approach (see Goodman 1986, van der Heijden, de Falguerolles, and de Leeuw 1989, and the discussions that follow these two papers). There are a number of similarities between procedures considered in the present paper and an approach to contingency table analysis, first proposed in van der Heijden and de Leeuw (1985) and then further elaborated in van der Heijden et al. (1989), which uses both correspondence analysis and loglinear models. A brief comparison of the procedures proposed in this paper with the combined correspondence analysis and loglinear analysis approach is made in Section 5.

2. Multivariate association models

For the $I \times J \times K$ contingency table, let F_{ijk} denote the expected frequency under some model for cell (i, j, k) . All models considered in this paper include the "main effects" $\mu_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C$, and maximum likelihood estimates are computed assuming that the sampling model is either multinomial, product-multinomial, or independent Poisson. The most general loglinear model for three-way contingency tables is

$$\log F_{ijk} = \mu_{ijk} + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}, \quad (1)$$

and the restrictions

$$\begin{aligned} \sum_{i=1}^I \lambda_i^A &= \sum_{j=1}^J \lambda_j^B = \sum_{k=1}^K \lambda_k^C = \sum_{i=1}^I \lambda_{ij}^{AB} = \sum_{j=1}^J \lambda_{ij}^{AB} = \sum_{i=1}^I \lambda_{ik}^{AC} = \sum_{k=1}^K \lambda_{ik}^{AC} \\ &= \sum_{j=1}^J \lambda_{jk}^{BC} = \sum_{k=1}^K \lambda_{jk}^{BC} = \sum_{i=1}^I \lambda_{ijk}^{ABC} = \sum_{j=1}^J \lambda_{ijk}^{ABC} = \sum_{k=1}^K \lambda_{ijk}^{ABC} = 0 \end{aligned} \quad (2)$$

may be used to uniquely identify the parameters in this model. The family of hierarchical loglinear models follows by considering special cases of model (1) where one or more sets of interaction terms is set to zero (i.e., removed from the

model); all lower order relatives of terms present in the model are automatically included in the model.

A very general family of association models may be derived by replacing the non-null interaction terms in hierarchical loglinear models with one or more dimensions of multiplicative terms. For example, an alternative parameterization of the hierarchical loglinear model of no-three-factor interaction is

$$\log F_{ijk} = \mu_{ijk} + \sum_{m=1}^{M_1} \phi_m^{AB} \nu_{1im}^A \nu_{1jm}^B + \sum_{m=1}^{M_2} \phi_m^{AC} \nu_{2im}^A \nu_{1km}^C + \sum_{m=1}^{M_3} \phi_m^{BC} \nu_{2jm}^B \nu_{2km}^C, \tag{3}$$

where $M_1 = \min(I - 1, J - 1)$, $M_2 = \min(I - 1, K - 1)$, and $M_3 = \min(J - 1, K - 1)$. The ν parameters are standardized “scores” and the ϕ parameters provide measures of “intrinsic association” (Goodman 1986). A variety of restrictions may be used to derive interesting and useful special cases of this model. These include (a) setting M_1 , M_2 and M_3 to values less than their respective upper bounds, (b) using fixed scores in place of scores that are estimated from the data, (c) requiring sets of scores to be homogeneous (e.g., $\nu_{1im}^A = \nu_{2im}^A$ for all i and m), and (d) requiring sets of scores to be symmetric (e.g., $\nu_{1im}^A = \nu_{1jm}^B$ for all $i = j$ and m) if there is a one-to-one correspondence between the categories of two or more variables.

In deriving association models exhibiting three-factor interaction it is useful to think of three-factor interaction as describing how the association between two variables changes according to the levels of the third variable. It is then apparent that alternative parameterizations of the λ_{ijk}^{ABC} may be based on three different sets of SVDs. That is, one may examine three-factor interaction in terms of studying how the association between A and B varies according to the levels of C , in terms of studying how the association between A and C varies according to levels of B , or in terms of how the association between B and C varies according to the levels of A . Let placement of a subscript in parentheses indicate that the subscript is held constant. The corresponding reparameterizations of the λ_{ijk}^{ABC} using multiplicative structures are

$$\begin{aligned} \lambda_{ijk}^{ABC} &= \sum_{m=1}^{M_1} \phi_{m(k)}^{AB(C)} \nu_{im(k)}^{A(C)} \nu_{jm(k)}^{B(C)} = \sum_{m=1}^{M_2} \phi_{m(j)}^{AC(B)} \nu_{im(j)}^{A(B)} \nu_{km(j)}^{C(B)} \\ &= \sum_{m=1}^{M_3} \phi_{m(i)}^{BC(A)} \nu_{jm(i)}^{B(A)} \nu_{km(i)}^{C(A)}. \end{aligned} \tag{4}$$

Restrictions, such as those mentioned in the preceding paragraph for use in simplifying two-factor associations, may be used to derive special cases of model (1) that exhibit three-factor interaction but are not saturated. For example, models replacing the λ_{ijk}^{ABC} with tri-linear terms of the form

$$\lambda_{ijk} = \sum_{m=1}^M \phi_m^{ABC} \nu_{3im}^A \nu_{3jm}^B \nu_{3km}^C$$

are easily derived from the first equation in (4) by (i) restricting the $\nu_{im(k)}^{A(C)}$ to be homogeneous (i.e., $\nu_{im(k)}^{A(C)} = \nu_{3im}^A$, for all k), (ii) restricting the $\nu_{jm(k)}^{B(C)}$ to be homogeneous (i.e., $\nu_{jm(k)}^{B(C)} = \nu_{3jm}^B$, for all k), and (iii) rewriting $\phi_{m(k)}^{AB(C)}$ as $\phi_m^{ABC} \nu_{3km}^C$.

The association modelling approach to contingency table analysis facilitates direct examination of many important substantive questions. Take, for example, the situation where one is interested in comparing bivariate associations (e.g., association between education and income) across groups (e.g., gender). Three basic questions that are of immediate interest in such an analysis are (1) does the association differ across the groups?, (2) if there are differences, are these differences only in terms of the strength of association, or are there also differences in the patterns of association?, and (3) if there are differences in the patterns of association, what are they? All of these questions are easily addressed using multivariate association models. The interested reader is referred to Clogg (1982a) and Becker and Clogg (1989a) for further discussion of using association models for group comparisons. Other applications of multivariate association models are described in Becker (1989a) and references cited therein.

3. Approximating maximum likelihood estimates

3.1. Outline of the procedure

One criticism of the association models approach to contingency table analysis has been that estimating the models by the method of maximum likelihood can require a substantial computer programming effort. There are algorithms available for ML estimation of a wide variety of association models (see, e.g., Goodman 1979, Clogg 1982a, Gilula and Haberman 1986, Becker and Clogg 1989b, Green 1988, 1989a, and Becker 1989b, 1990), but there is a need for still further development. The method of estimation described in this section generalizes a procedure used in Becker (1990) for computing initial parameter estimates for an algorithm that fits the so-called RC(M) association model by the method of maximum likelihood. A comparable set of estimates can be obtained directly by SVD of suitably transformed data (Khoudraji (1988)).

The basic idea is to fit loglinear models by the method of maximum likelihood, and then apply SVD to matrices having estimated interaction parameters as their entries. Consider, for example, the usual loglinear model of no-three-factor interaction:

$$\log F_{ijk} = \mu_{ijk} + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}. \quad (5)$$

Let circumflexes (“ $\hat{}$ ”) denote maximum likelihood estimates, and let tildes (“ $\tilde{}$ ”) denote estimates obtained using the procedure being described here. The three matrices to which we apply the SVD are the $I \times J$ matrix $[AB] = [\hat{\lambda}_{ij}^{AB}]$, the

$I \times K$ matrix $[AC] = [\hat{\lambda}_{ik}^{AC}]$, and the $J \times K$ matrix $[BC] = [\hat{\lambda}_{jk}^{BC}]$. The resulting estimates for scores and intrinsic association parameters in model (3) are

$$\begin{aligned} [\hat{\lambda}_{ij}^{AB}] &= \left[\sum_{m=1}^{M_1} \tilde{\phi}_m^{AB} \tilde{\nu}_{1im}^A \tilde{\nu}_{1jm}^B \right], [\hat{\lambda}_{ik}^{AC}] = \left[\sum_{m=1}^{M_2} \tilde{\phi}_m^{AC} \tilde{\nu}_{2im}^A \tilde{\nu}_{1km}^C \right], \\ \text{and } [\hat{\lambda}_{jk}^{BC}] &= \left[\sum_{m=1}^{M_3} \tilde{\phi}_m^{BC} \tilde{\nu}_{2jm}^B \tilde{\nu}_{2km}^C \right]. \end{aligned} \quad (6)$$

It is important to note that the scores and intrinsic association parameters in (6) are indeed maximum likelihood estimates for model (3), but not otherwise. For example, the first dimension of the SVD of $[AB]$ (i.e., $\tilde{\phi}_1^{AB}$, $\tilde{\nu}_{1i1}^A$, and $\tilde{\nu}_{1j1}^B$) generally will not be maximum likelihood estimates for a model where the λ_{ij}^{AB} are replaced by $\phi^{AB} \nu_{1i}^A \nu_{1j}^B$.

There are a number of matrices to which the SVD can be applied when studying three-factor-interaction; i.e., model (1). One can work with the three matrices corresponding to the three sets of two-factor interaction terms, the $K(I \times J)$ matrices having $\hat{\lambda}_{ij(k)}^{ABC}$ as the ij th entry, and the similarly defined $J(I \times K)$ matrices and $I(J \times K)$ matrices. For some applications (e.g., analysing how association varies across groups) matrices such as

$$\left[\sum_{m=1}^{M_1} \tilde{\phi}_{m(k)}^{AB(C)} \tilde{\nu}_{im(k)}^{A(C)} \tilde{\nu}_{jm(k)}^{B(C)} \right] = [\hat{\lambda}_{ij}^{AB} + \hat{\lambda}_{ij(k)}^{ABC}] \quad (7)$$

are of particular interest. There are other applications where it is reasonable to concatenate several matrices before applying the SVD (Gilula and Haberman 1988, and van der Heijden, de Falguerolles, and de Leeuw 1989).

3.2. Application of the procedure to the data in Table 1

Applying the usual hierarchical loglinear models, the Pearson and likelihood-ratio statistics given in Becker (1989a) indicate that the model of no-three-factor interaction provides a satisfactory fit of the data and that none of the simpler models are tenable. Hence the analysis given here shall proceed using all three matrices given in (6); Becker (1989a) found that the model where $M_1 = 2$, $M_2 = 2$, and $M_3 = 2$ provides a reasonable summary of the data (likelihood-ratio statistic (L^2) = 35.78, Pearson statistic (X^2) = 37.33, on 30 degrees of freedom (d.f.)). The ML estimates of the association parameters (i.e., $\hat{\phi}$'s and $\hat{\nu}$'s) and their corresponding approximations are given in Table 3. The approximations to the ML estimates can be computed as follows: (a) fit the loglinear model of no-three-factor interaction using, for example, GLIM, (b) center the ML estimates of the interaction parameters in this model (i.e., λ_{ij}^{AB} , λ_{ik}^{AC} , and λ_{jk}^{BC}) so that they satisfy equation (2), (c) use the centered estimates of the λ_{ij}^{AB} to form an $I \times J$ matrix, use the centered estimates of the λ_{ik}^{AC} to form an $I \times K$ matrix, and use the centered estimates of the λ_{jk}^{BC} to form a $J \times K$ matrix, and (d) apply the SVD to each of the three matrices formed in step (c), the resulting

Table 3
 Association parameters estimated from the data in Table 1

<i>Education by income partial association</i>			
	$\hat{\phi}_1^{EI} = 1.18$		$\hat{\phi}_2^{EI} = 0.39$
	$\tilde{\phi}_1^{EI} = 1.14$		$\tilde{\phi}_2^{EI} = 0.42$
Education	$\hat{\underline{v}}_{11}^E = (0.02, 0.71, -0.04, -0.70)$		$\hat{\underline{v}}_{12}^E = (0.35, -0.47, 0.63, -0.51)$
	$\tilde{\underline{v}}_{11}^E = (0.03, 0.72, -0.06, -0.69)$		$\tilde{\underline{v}}_{12}^E = (0.43, -0.46, 0.56, -0.53)$
Income	$\hat{\underline{v}}_{11}^I = (-0.52, -0.31, 0.02, 0.80)$		$\hat{\underline{v}}_{12}^I = (-0.42, -0.09, 0.84, -0.33)$
	$\tilde{\underline{v}}_{11}^I = (-0.48, -0.35, 0.02, 0.81)$		$\tilde{\underline{v}}_{12}^I = (-0.51, 0.01, 0.81, -0.31)$
<i>Education by social rank partial association</i>			
	$\hat{\phi}_1^{ES} = 1.67$		$\hat{\phi}_2^{ES} = 0.26$
	$\tilde{\phi}_1^{ES} = 1.77$		$\tilde{\phi}_2^{ES} = 0.25$
Education	$\hat{\underline{v}}_{21}^E = (0.03, -0.82, 0.29, 0.49)$		$\hat{\underline{v}}_{22}^E = (0.59, -0.09, -0.76, 0.26)$
	$\tilde{\underline{v}}_{21}^E = (0.03, -0.81, 0.27, 0.52)$		$\tilde{\underline{v}}_{22}^E = (0.69, -0.12, -0.70, 0.13)$
Social Rank	$\hat{\underline{v}}_{11}^S = (0.68, 0.09, -0.04, -0.73)$		$\hat{\underline{v}}_{12}^S = (0.42, -0.07, -0.79, 0.44)$
	$\tilde{\underline{v}}_{11}^S = (0.67, 0.12, -0.06, -0.73)$		$\tilde{\underline{v}}_{12}^S = (0.45, -0.12, -0.76, 0.45)$
<i>Income by social rank partial association</i>			
	$\hat{\phi}_1^{IS} = 2.97$		$\hat{\phi}_2^{IS} = 0.31$
	$\tilde{\phi}_1^{IS} = 2.96$		$\tilde{\phi}_2^{IS} = 0.31$
Income	$\hat{\underline{v}}_{21}^I = (-0.41, -0.43, 0.02, 0.81)$		$\hat{\underline{v}}_{22}^I = (-0.32, -0.23, 0.86, -0.31)$
	$\tilde{\underline{v}}_{21}^I = (-0.41, -0.42, 0.02, 0.81)$		$\tilde{\underline{v}}_{22}^I = (-0.34, -0.21, 0.86, -0.31)$
Social Rank	$\hat{\underline{v}}_{21}^S = (0.80, -0.01, -0.23, -0.56)$		$\hat{\underline{v}}_{22}^S = (0.33, -0.78, -0.07, 0.52)$
	$\tilde{\underline{v}}_{21}^S = (0.80, -0.01, -0.24, -0.55)$		$\tilde{\underline{v}}_{22}^S = (0.33, -0.75, -0.12, 0.55)$

eigenvalues and eigenvectors from the first two dimensions of the SVDs are the approximations to the ML estimates of the association parameters. The ML estimates were computed using a uni-dimensional Newton algorithm similar to the ones described in Goodman (1979), Clogg (1982a), and Becker (1990). The approximate estimates appear to be most suitable for use as initial estimates in algorithms that compute maximum likelihood estimates. We suggest using weighted averages of the approximate estimates when computing initial estimates for algorithms that fit models having homogeneous scores by ML; see, e.g., Becker and Clogg (1989b).

The ML fit of the cell counts can be approximated by using the approximate estimates of the association parameters to form an OFFSET (in the GLIM parlance) for the model $\log F_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C$, which is then estimated by the method of maximum likelihood. The ML estimates of the cell counts under

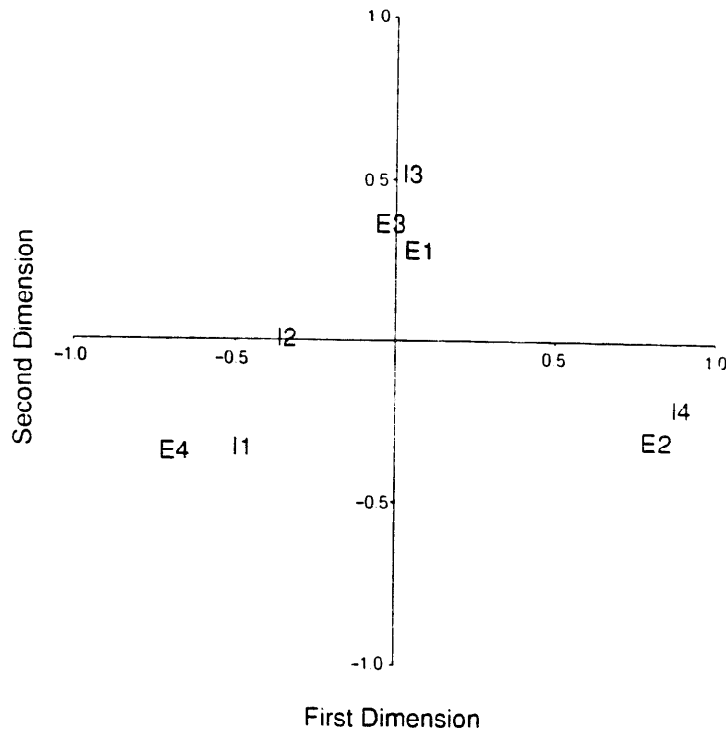


Fig. 1. Approximate estimates.

the $M_1 = 2$, $M_2 = 2$, and $M_3 = 2$ model are the middle entries in Table 1, and their corresponding approximations are the bottom entries. The Kullback–Leibler distance (see, e.g., Whittaker 1990, Ch. 4) between the observed data and the approximations, scaled for comparison to the likelihood-ratio statistic for this model, is 38.34, and $X^2 = 40.27$. The approximation appears to be quite adequate for exploratory purposes. It should be noted that, rather than fitting the main effects model (with OFFSET) in GLIM, the main effect parameters from the fit of the original loglinear model could be used in computing estimated cell frequencies. The reasons for preferring one method over the other are not compelling, particularly if the original loglinear model was estimated using software that identifies parameters by setting some of them equal to zero; as does GLIM. In this case one must make appropriate adjustments to the main effect parameters since the interaction parameters are centered prior to the SVD calculations. One can either adjust the main effects from the fit of the original loglinear model or they can estimate a new set of main effects as described above. Fitting the main effects model with the association parameters forming an OFFSET should result in a slightly better fit in terms of Kullback–Leibler distance.

Graphical displays of the approximate estimates can be quite useful in exploratory analyses. All plots presented in this paper are in the spirit of the so-called biplots (see, e.g., Gabriel 1972). For example, Figure 1 is a plot of the education by income partial association based on the approximate estimates $\tilde{\phi}$'s and $\tilde{\nu}$'s given in Table 3, and Figure 2 is based on the corre-

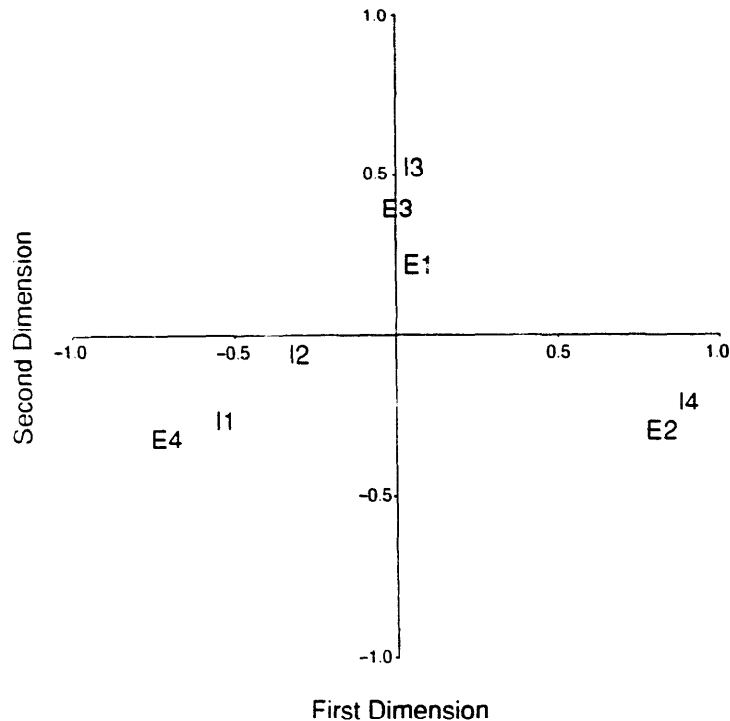


Fig. 2. Maximum likelihood estimates.

spending ML estimates. The coordinates of the E_i are estimates of the points $(\sqrt{\phi_1^{EI}} \nu_{i1}^E, \sqrt{\phi_2^{EI}} \nu_{i2}^E)$, and the coordinates of the I_j are estimates of the points $(\sqrt{\phi_1^{EI}} \nu_{j1}^I, \sqrt{\phi_2^{EI}} \nu_{j2}^I)$. Figure 2 corrects a minor error in Figure 1 in Becker (1989a); the interpretation given there is unaffected by the error. There are obvious differences between Figures 1 and 2, but both plots convey the same basic information.

4. A simple procedure for screening models

4.1. Outline of the procedure

The family of models considered in Becker (1989a) is far too general for all possible models to be estimated by the method of maximum likelihood in a particular data analysis. There are, of course, situations where the context of the problem will suggest a subset of models to be examined, but in some analyses a very broad range of models is of interest. The estimation and plotting procedures outlined in Section 3 may be used to screen models and suggest ones for further analysis. A simple approach to model screening is to (a) fit hierarchical loglinear models to arrive at one or more models that provide an adequate description of the data, (b) obtain SVDs of matrices of estimated interaction terms from the model(s) selected in (a), and (c) plot the resulting scores to serve as a guide in determine which restrictions should be considered. Agresti (1990,

pp. 275) also suggests using loglinear models to guide the model search process, but in reference to a smaller set of association models.

Plots of scores from SVDs of matrices which include the $\hat{\lambda}_{ijk}^{ABC}$ are especially important since the usual loglinear model is saturated for this case. Consider, for example, the SVDs of the $K(I \times J)$ matrices with the $\hat{\lambda}_{ij(k)}^{ABC}$ as entries. If a one dimensional approximation of each matrix is adequate, some potentially interesting parameterizations of three-factor interaction are

$$\lambda_{ijk}^{ABC} = \phi_k^{ABC} \nu_{ik}^{AC} \nu_{jk}^{BC}, \quad \lambda_{ijk}^{ABC} = \phi_k^{ABC} \nu_i^A \nu_{jk}^{BC},$$

$$\lambda_{ijk}^{ABC} = \phi_k^{ABC} \nu_{ik}^{AC} \nu_j^B \quad \text{and} \quad \lambda_{ijk}^{ABC} = \phi_k^{ABC} \nu_i^A \nu_j^B \nu_k^C.$$

The second parameterization is suggested if the $\tilde{\nu}_{ik}^{AC}$ ($k = 1, \dots, K$) only differ (approximately) by a scalar multiple, the third parameterization is suggested if the $\tilde{\nu}_{jk}^{BC}$ ($k = 1, \dots, K$) differ by a scalar multiple, and the fourth parameterization is suggested if both of these conditions appear to be met.

4.2. Application of the procedure to the data in Table 2

The simplest hierarchical loglinear model that appears to provide a reasonable description of the association in Table 2 is the model of no-three-factor interaction ($L^2 = 12.07$, $X^2 = 11.33$, d.f. = 6). However, as was shown in Becker (1989a), there exists a simple three-factor association model that provides a much better fit ($L^2 = 2.47$, $X^2 = 2.64$, d.f. = 6) to these data. The association model replaces λ_{ij}^{HW} and λ_{ijk}^{HWS} (H : husband's highest degree, W : wife's highest degree, S : sex of the respondent) in the saturated loglinear model with a single term, $\phi_k^{HWS} \nu_{ik}^{HS} \nu_j^W$, where the ν_{ik}^{HS} are parameters that are estimated from the data and the ν_j^W are equally spaced fixed scores. This association model is actually a loglinear model, but not a hierarchical loglinear model.

Application of the estimation and plotting procedures described in Section 3 does indeed lead one to consider simple three-factor association models. The following analysis is based on parameter estimates obtained using GLIM 3.77. The parameter estimates have been centered so that they satisfy (2). First we fit the model of no-three-factor interaction and focus on the 4×3 matrix $[\hat{\lambda}_{ij}^{HW}]$. The first two dimensions of the SVD of the $[\hat{\lambda}_{ij}^{HW}]$ are plotted in Figure 3. The coordinates have been scaled by the square roots of the $\hat{\phi}$ s, as before. This plot suggests that equally spaced scores for wife's highest degree should be considered, and that the first dimension accounts for most of the association. The ML fit of the model which replaces λ_{ij}^{HW} with $\phi^{HW} \nu_i^H \nu_j^W$ ($L^2 = 14.67$, $X^2 = 15.27$, d.f. = 9) confirms that most of the HW association can be accounted for using a one-dimensional multiplicative structure with equally spaced scores for the wife's highest degree variable. There is, however, evidence of lack of fit with this model. The likelihood-ratio based comparison of the two no-three-factor interaction models (i.e., $14.67 - 12.07 = 2.60$, on 3 d.f.) suggests that the lack of fit is not the result of oversimplifying the HW association, and hence the presence of three-factor interaction is suggested.

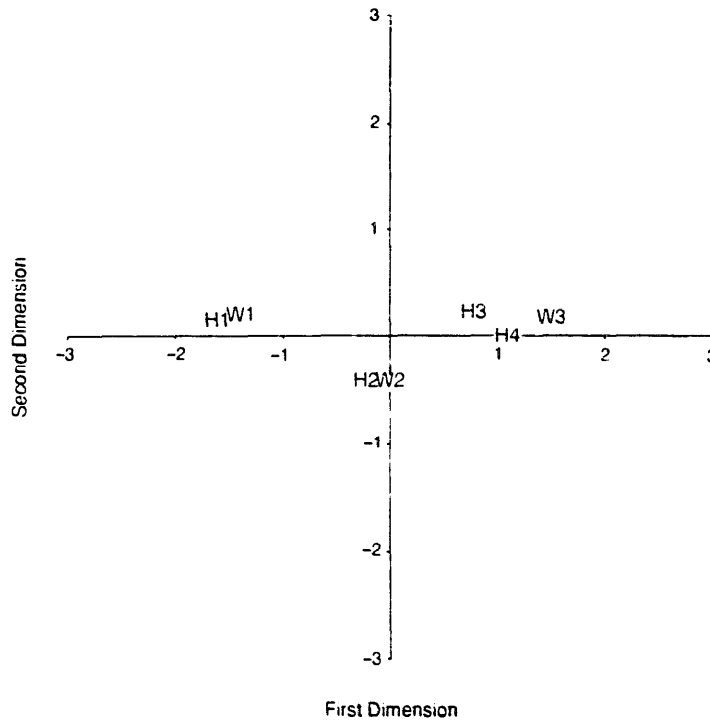


Fig. 3. Table 2. Husband by wife association, no 3-factor integration.

In studying the saturated loglinear model we focus our attention on the $2(4 \times 3)$ matrices $[\lambda_{ij}^{HW} + \lambda_{ij(k)}^{HWS}]$. Figure 4 is a plot of the first two dimensions from the SVD of the matrix for male respondents, and Figure 5 is the

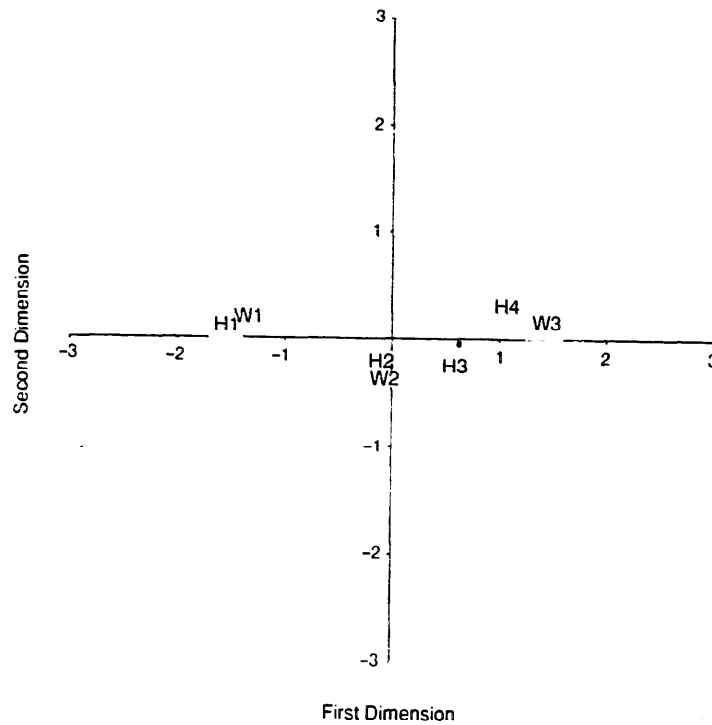


Fig. 4. Table 2. HW association, male respondents.

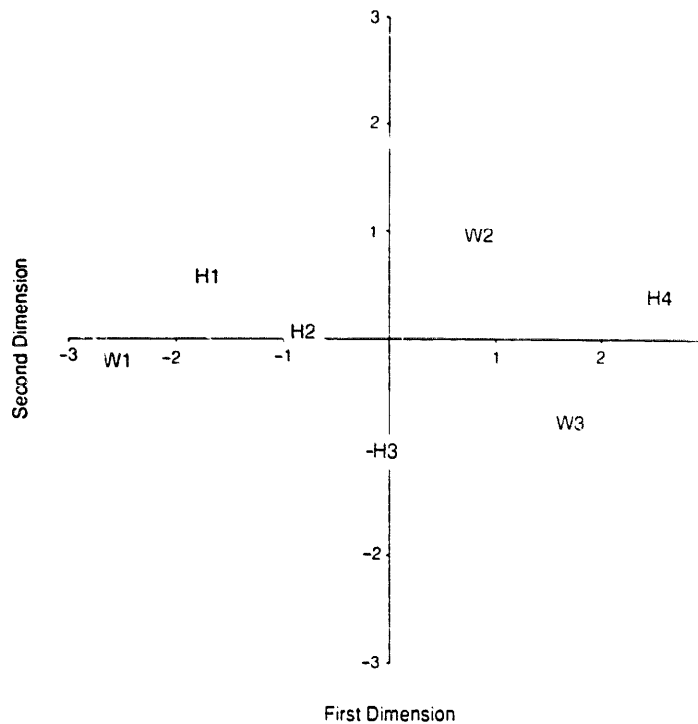


Fig. 5. Table 2. HW association. female respondents.

corresponding plot for female respondents. The variances associated with many of the $\hat{\lambda}_{ijk}^{HWS}$ terms were quite large, and this needs to be considered when interpreting these two plots. Figure 4 is quite similar to Figure 3, but Figure 5 suggests that both sets of scores may depend on the levels of S . It appears, once again, that a one-dimensional multiplicative structure will suffice. Standard likelihood based comparisons of models suggested by these observations leads one to conclude that the scores for husband's highest degree do indeed differ in some ways across the levels of S , but that the wife's highest degree scores are essentially homogeneous and may be taken to be equally spaced. This is in fact the model described by Becker (1989a). A few points to be noted from this analysis of Table 2 are that (a) the simplest loglinear model which "fits the data" may lead one to models that fail to account for important associations in the data, (b) parameter estimates from models that "fit too well" will have inflated variances and hence one should be careful not to overinterpret graphical displays of such estimates, and (c) likelihood based methodology continues to play an important role in the model search process.

5. Discussion

The combined use of loglinear models and the SVD is common both to the present paper and to the "combined approach to contingency table analysis using correspondence analysis and log-linear analysis" developed in van der

Heijden and de Leeuw (1985) and van der Heijden et al. (1989). Where we have emphasized use of the SVD in decomposing matrices of interaction parameters from loglinear models, van der Heijden, de Leeuw, and de Falguerolles emphasize use of the SVD in decomposing residuals from loglinear models. We see then that the principal difference between the two approaches is that the one motivated by associated models focuses on studying the structure of terms included in a loglinear model while the correspondence analysis (CA) based methodology focuses on studying the structure of terms not included in the model. Both approaches have merit and should be expected to provide compatible answers when used to answer questions that can be addressed from either perspective. There are, however, situations where they may well provide somewhat different answers. The previous example is a case in point. It is doubtful that the CA approach, as described in the two papers cited above, would have arrived at the same model, and hence the same interpretation, as was obtained above. We suspect that the CA approach would lead one to focus on the interpretation of the model of no-three-factor interaction since this model provides a reasonable fit of the data. This, however, is pure speculation. The end result depends both on the tools being used and on the person(s) using them. Differences in end results are more likely to arise because of differences in the perspectives of the analysts, which will in turn play a role in how they use their tools, rather than because of differences in the tools themselves.

References

- A. Agresti, *Categorical Data Analysis* (Wiley, New York, 1990).
- A. Agresti and A. Kezouh, Association models for multidimensional cross-classifications of ordinal variables, *Comm. in Stat., A - Theory and Methods* **12** (1983) 1261-1276.
- E.B. Andersen, *Discrete Statistical Models with Social Science Applications* (North Holland, Amsterdam, 1980).
- M.P. Becker, Models for the analysis of association in multivariate contingency tables, *J. Amer. Stat. Assoc.* **84** (1989a) 1014-1019.
- M.P. Becker, Square contingency tables having ordered categories and GLIM, *GLIM Newsletter* **19** (1989b) 22-31.
- M.P. Becker, Maximum likelihood estimation of the $RC(M)$ association model, *Applied Statistics* **39** (1) (1990) 152-166.
- M.P. Becker and C.C. Clogg, Analysis of sets of two-way contingency tables using association models, *J. Amer. Stat. Assoc.* **84** (1989a) 142-151.
- M.P. Becker and C.C. Clogg, Maximum likelihood estimation of association models for sets of two-way contingency tables, Technical Report 89-1, Department of Biostatistics, University of Michigan (1989b).
- H. Caussinus and A. de Falguerolles, Tableaux carrés: modélisation et méthodes factorielles, *Revue de Statistique Appliquée* **35** (1987) 35-52.
- C.C. Clogg, Some models for the analysis of association in multiway cross-classification having ordered categories, *J. Amer. Stat. Assoc.* **77** (1982a) 80-815.
- C.C. Clogg, Using association models in sociological research: Some examples, *Amer. J. of Soc.* **88** (1982b) 114-134.

- Z. Gilula and S.J. Haberman, Canonical analysis of two-way contingency tables by maximum likelihood, *J. Amer. Stat. Assoc.* **81** (1986) 780–788.
- Z. Gilula and S.J. Haberman, The analysis of multivariate contingency tables by restricted canonical and restricted association models, *J. Amer. Stat. Assoc.* **83** (1988) 760–771.
- L.A. Goodman, Simple models for the analysis of association in cross-classifications having ordered categories, *J. Amer. Stat. Assoc.* **74** (1979) 537–552.
- L.A. Goodman, The analysis of cross-classified data having ordered categories and/or unordered categories: Association models, correlation models, and asymmetric models for contingency tables with or without missing entries, *The Annals of Statistics* **13** (1985) 10–69.
- L.A. Goodman, Some useful extensions of the usual correspondence analysis approach and the usual loglinear models approach in the analysis of contingency tables, *Int. Stat. Rev.* **54** (1986) 243–309.
- M. Green, A modelling approach to multiple correspondence analysis in: *COMPSTAT88* (Physica-Verlag, New York, 1988) 317–322.
- M. Green, Generalisations of the Goodman association model for the analysis of multi-dimensional contingency tables, in: A. Decarli et al. (Eds.) *Lecture Notes in Statistics 57: Statistical Modelling* (Springer-Verlag, New York, 1989a) 165–171.
- M. Green, in: Discussion of the paper by van der Heijden, de Falguerolles and de Leeuw, *Applied Statistics* **38** (2) (1989b) 277–278.
- M. Greenacre, *Theory and Applications of Correspondence Analysis* (Academic Press, New York, 1984).
- S.J. Haberman, *Analysis of Qualitative Data, Volume 1: Introductory Topics* (Academic Press, New York, 1978).
- P.G.M. van der Heijden and J. de Leeuw, Correspondence analysis used complementary to log-linear analysis, *Psychometrika* **50** (1985) 429–447.
- P.G.M. van der Heijden, A. de Falguerolles and J. de Leeuw, A combined approach to contingency table analysis using correspondence analysis and log-linear analysis (with discussion), *Applied Statistics* **38** (2) (1989) 249–292.
- A. Khoudraji, Analyse des correspondances et mise en oeuvre du modèle de Goodman, These, Université Paul Sabatier de Toulouse (Sciences) (1988).
- J. Whittaker, *Graphical Models in Applied Multivariate Statistics* (John Wiley & Sons, New York, 1990).