Reduced rank models for contingency tables

BY JAN DE LEEUW  
Departments of Psychology and Mathematics, University of California, Los Angeles,  
California 90024-1563, U.S.A.

AND PETER G. M. VAN DER HEIJDEN  
Department of Empirical and Theoretical Sociology, University of Utrecht,  
3508 TC Utrecht, The Netherlands

SUMMARY

Reduced rank models for the analysis of two-way contingency tables are introduced. Two classes of reduced rank models are discerned, with well-known exponents canonical analysis and latent class analysis. The relation between these two classes is discussed. Results on the subject mentioned earlier in the literature are shown to be either redundant or inaccurate.

Some key words: Canonical analysis; Correspondence analysis; Latent class analysis; Reduced rank models.

1. INTRODUCTION

In recent years much attention has been given to models for two-way contingency tables that can be formulated in terms of reduced rank of a matrix with probabilities. A well-known reduced rank model is the independence model, where the rank is one. For rank higher than one distinct classes of reduced rank models are possible. Each has the independence model as the special case for rank one. A first class of such models is closely related to what is known under names as canonical analysis or correspondence analysis. Recently much attention has been given to the maximum likelihood estimation of versions of these models by Goodman (1985, 1986, 1987) and Gilula & Haberman (1986, 1988). A second class of models that can be formulated in terms of reduced rank is latent class analysis, LCA, for two-way tables. Latent class analysis was proposed by Lazarsfeld (1950a,b). See Clogg (1981) for a more recent review.

In this paper we relate these classes of models to each other. The relation has been discussed earlier by Gilula (1979, 1983, 1984), Gilula & Haberman (1986), Goodman (1987), and van der Heijden, Mooijaart & de Leeuw (1989). We summarize existing results in a simple way using new proofs. Gilula (1979) provided conditions that had to hold for rank-2 correspondence analysis to imply rank-2 latent class analysis. We show here that rank-2 correspondence analysis always implies rank-2 latent class analysis. This implies that the theorem and the example given by Gilula (1979) are incorrect.

2. GENERAL REDUCED RANK MODELS

The basic model studied in this paper assumes that an $n \times m$ probability matrix $\Pi$ has rank $\rho$, where $\rho \leq \min(n,m)$. We call this model $R_\rho$. The probability matrix $\Pi$ has all elements nonnegative, while the sum of the $\pi_{ij}$ is equal to one. We suppose, unless indicated otherwise, that $\Pi$ is full, in the sense that its row sums $\pi_{i+}$ and its column sums $\pi_{+j}$ are all positive. Thus no row or column is equal to zero.

We compare this model with the canonical model $C_\rho$, in which at most $\rho - 1$ of the canonical correlations between the row and the column variables of the table are nonzero. These canonical correlations are the stationary values of the product moment correlation coefficient, seen as a function of scores for rows and scores for columns.
We also compare \( R_p \) with the model suggested by correspondence analysis, written as \( A_p \), in which

\[
\pi_{ij} = \omega_j \theta_i \left( 1 + \sum_{s=1}^{n-1} \lambda_s x_i y_{js} \right).
\]

Another way of formulating \( A_p \) is by saying that \( \Pi \) has a Fisher-decomposition of rank \( \rho - 1 \) (Lancaster, 1958).

**Theorem 1.** We have that \( C_p, A_p \) and \( R_p \) are equivalent.

**Proof.** If \( \rho - 1 \) canonical correlations are nonzero, then \( \Pi \) can be written in the form \( A_p \), with \( \omega_j \) and \( \theta_i \) equal to the marginals \( \pi_{i+} \) and \( \pi_{+j} \), with \( \lambda_s \) equal to the canonical correlations, and with \( x_{ij} \) and \( y_{ij} \) equal to the canonical scores (Lancaster, 1958). Thus \( C_p \) implies \( A_p \). It is obvious, moreover, that \( A_p \) implies \( R_p \). We now prove that \( R_p \) implies \( C_p \). Suppose rank \( \Pi = \rho \). By the Lagrange theorem, for instance Gutman (1944), we know that \( \pi_{ij} - \pi_{i+} \pi_{+j} \) has rank exactly equal to \( \rho - 1 \) and is doubly centred. The canonical correlations are computed from the singular value decomposition of the matrix of normalized residuals \( Z \), given by

\[
\pi_{ij}^* = \frac{\pi_{ij} - \pi_{i+} \pi_{+j}}{\sqrt{(\pi_{i+} \pi_{+j})}}.
\]

The matrix \( Z \) is of rank \( \rho - 1 \), and thus has \( \rho - 1 \) nonzero canonical correlations. \( \square \)

3. **Reduced rank models with nonnegativity constraints**

Let us now look at the model \( R^*_p \), which assumes that rank \( \Pi = \rho \), and moreover that there exists a full rank decomposition \( \Pi = AB^t \), with \( A \geq 0 \) and \( B \geq 0 \). Clearly \( R^*_p \) implies \( R_p \), but in general the reverse implication is not true, at least not obvious.

There are some interesting alternative ways to write \( R^*_p \). In the first place the latent class model \( \text{LCA}_p \), mentioned by Good (1965), is such that \( \Pi \) is a mixture of \( \rho \) bivariate distributions with independence. Thus

\[
\pi_{ij} = \sum_{s=1}^\rho \eta_s \alpha_i \beta_j,
\]

with \( \eta_s = \alpha_{s+} = \beta_{s+} = 1 \). Moreover all parameters are nonnegative. There is also the latent budget model \( \text{LBA}_p \) (van der Heijden et al., 1989), in which

\[
\frac{\pi_{ij}}{\pi_{i+}} = \sum_{s=1}^\rho \alpha_i \beta_j,
\]

with \( \alpha_{s+} = \beta_{s+} = 1 \), and again all parameters are nonnegative.

**Theorem 2.** We have that \( R^*_p \), \( \text{LBA}_p \) and \( \text{LCA}_p \) are equivalent.

**Proof.** Suppose \( \Pi \) satisfies \( R^*_p \). Thus \( \Pi = AB^t \), with \( A \geq 0 \) and \( B \geq 0 \). Suppose \( \Phi \) is a diagonal matrix of order \( \rho \), with the \( b_{ss} \) on the diagonal. Let \( \tilde{A} = A\Phi \) and \( \tilde{B} = B(\Phi^{-1})^t \). Then clearly \( \Pi = \tilde{A}\tilde{B}^t \). Moreover \( \tilde{b}_{ss} = 1 \) and \( \tilde{a}_{ss} = \pi_{ss} \). If we define \( \beta_{ss} = \tilde{b}_{ss} \) and \( \alpha_{ss} = \tilde{a}_{ss} / \tilde{a}_{ss} \), then we satisfy \( \text{LBA}_p \). Let \( \beta_{ss} = \tilde{b}_{ss} \) and \( \alpha_{ss} = \tilde{a}_{ss} / \tilde{a}_{ss} \), and \( \eta_s = \tilde{a}_{ss} \). These quantities satisfy \( \text{LCA}_p \). Thus \( R^*_p \) implies \( \text{LBA}_p \), and \( \text{LBA}_p \) and \( \text{LCA}_p \) imply each other. It is trivial that \( \text{LCA}_p \) implies \( R^*_p \). \( \square \)

4. **Existence of nonnegative decompositions**

As we said above, in general \( R^*_p \) implies \( R_p \), but the reverse implication is not necessarily true. The relationship between these models was already mentioned by Good (1965, p. 64), and studied by Gilula (1979, 1983, 1984).
THEOREM 3. We have that $R_2$ and $R^*_2$ are equivalent.

Proof. We know that $R^*_2$ implies $R_2$, so we merely have to prove the reverse. Because of $R_2$ the columns of $\Pi$ are $m$ vectors in a two-dimensional subspace of $R^n$. Because all columns are nonnegative they are actually in a pointed cone in this plane. Two-dimensional cones are simplicial, i.e. they have exactly two extreme rays. The bundle of rays corresponding with the columns of $\Pi$ has two extremes, all other columns are positive linear combinations of these two columns. But this means that $R^*_2$ is true, with $A$ equal to these two extreme columns. \square

This very simple geometric proof is due to Paul Bekker. It replaces a lengthy computational proof we first had, and a complicated algebraic proof by Thomas (1974) we subsequently discovered. Thomas (1974) also gave a necessary and sufficient condition for $R_2$ to imply $R^*_2$, which reformulates the problem in terms of the existence of certain polyhedral convex cones. He also provided the counterexample

$$
\begin{bmatrix}
0.125 & 0.125 & 0.0 & 0.0 \\
0.125 & 0.0 & 0.125 & 0.0 \\
0.0 & 0.125 & 0.0 & 0.125 \\
0.0 & 0.0 & 0.125 & 0.125
\end{bmatrix}
$$

This matrix satisfies $R_3$, but not $R^*_3$.

It follows from our result that the Theorem and Corollary 1 of Gilula (1979) are not correct. This result also shows that van der Heijden et al. (1989) are incorrect in stating that latent class analysis and correspondence analysis are always equivalent, i.e. for any rank $\rho$.

The example Gilula (1979) gives is supposed to satisfy $R_2$ and not $R^*_2$. The probability matrix is

$$
\begin{bmatrix}
0.165 & 0.005 & 0.030 \\
0.015 & 0.580 & 0.105 \\
0.020 & 0.065 & 0.015
\end{bmatrix}
$$

In this example the first two columns of $\Pi$ are the extreme columns, and thus columns $\pi_1$, $\pi_2$ and $\pi_3$ satisfy the relationship $\pi_3 = \frac{1}{17}(\pi_1 + \pi_2)$. Consequently

$$
\Pi = \begin{bmatrix}
0.165 & 0.005 \\
0.015 & 0.580 \\
0.020 & 0.065
\end{bmatrix} \begin{bmatrix}
1 & 0 & 0.1765... \\
0 & 1 & 0.1765...
\end{bmatrix},
$$

which counters Gilula's counterexample.

ACKNOWLEDGEMENTS

We are grateful to P. Bekker and J. F. B. M. Kraaij evanger for helpful comments.

REFERENCES


[Received August 1989. Revised June 1990]