Multiple Correspondence Analysis

and

Ordered Latent Structure Models

B.F. Schriever
Nederlandse Philips Bedrijven B.V., Centre for Quantitative Methods
P.O. Box 218, 5600 MD Eindhoven

This paper discusses application of multiple correspondence analysis in ordered latent structure models. Such models are frequently used in psychological measurement theory to analyse ability (or attitude) tests, e.g. intelligence tests. The models considered are related to those of Mokken. It turns out that, under realistic assumptions, multiple correspondence analysis orders the questions (items) in the test according to their difficulty and orders individuals according to their ability (or attitude).

1. INTRODUCTION

Multiple correspondence analysis (abbreviated MCA) is a modern statistical technique for describing the association between categorical (i.e. discrete) variables. The technique is commonly used to analyse large data sets. It gives insight into the complex dependence structure of such data sets by making plots. MCA has proved to be an important and useful tool for analysing the association which is present in data sets with many variables. In this paper we discuss MCA in a more unusual situation and we need not consider its graphical representation. We apply MCA in the analysis of ordered latent structure models. Such models are developed for the following situation which frequently arises in e.g. psychology and medicine. In a population individuals must be ordered according to their value on an *unobservable* characteristic (e.g. intelligence, knowledge of a subject, attitude in a given context, a specific disease). For this purpose responses on a set of variables related to the characteristic are collected for each individual (e.g. an intelligence test). We restrict attention to the simple case in which there is only one such characteristic, called the *latent variable,* and in which the collected response variables, called *items,* are dichotomous (i.e. have only two response categories; for example 'correct' and 'wrong'). The set of all items is called the *test.* Since the characteristic of interest is often hard to separate from other characteristics, the assumption that responses on the items (i.e. the response variables) systematically depend on only one latent variable is for most applications more restrictive than the dichotomy assumption.

The paper is organized as follows. Section 2 reviews the definition of the technique MCA. In Section 3 we introduce latent structure models with ordered items. These ordered models are special cases of the models introduced by MOKKEN [17]. Our main result is given in Section 4. It demonstrates that the ordering of items is reflected in the MCA scores. This implies that MCA orders the individuals according to their latent value and orders the items according to their difficulty. In the last section it is shown that most well-known examples of latent structure models possess the orderings of Section 3. GIFI ([5], Chap. 9) already noted the ordering property of MCA for these specific examples, but proofs were not given.

The result presented in this paper also appeared in SCHRIEVER [23].

2. MULTIPLE CORRESPONDENCE ANALYSIS

Let $X_1, X_2, \ldots, X_k$ be categorical random variables. The technique MCA seeks k real valued functions $\phi_{11}, \phi_{21}, \ldots, \phi_{k1}$, defined on the categories (possible values) of $X_1, X_2, \ldots, X_k$ respectively, such that the first principal component of the correlation matrix of $\phi_{11}(X_1), \phi_{21}(X_2), \ldots, \phi_{k1}(X_k)$ has maximal variance. This principal component is called the first MCA component. It describes the most informative part of the variation between the categorical variables. Clearly, it is no restriction to assume that the derived variables $\phi_{l1}(X_l)$ have expectation zero and variance unity, for $l = 1, \ldots, k$. Subsequently, MCA seeks a second component which has maximal variance but which is uncorrelated with the first. This procedure is continued with a third component, a fourth component, etc. until no new component which is uncorrelated with the previous components can be found.

DEFINITION. *The $t-$th MCA component is the linear combination of transformed variables*

$$Y_t = \sum_{l=1}^{k} \alpha_{lt} \phi_{lt}(X_l)$$

for which $\mu_t = Var(Y_t)$ is maximal subject to

$$E\phi_{lt}(X_l) = 0, \quad Var(\phi_{lt}(X_l)) = 1 \textit{ for } l = 1, \ldots, k,$$

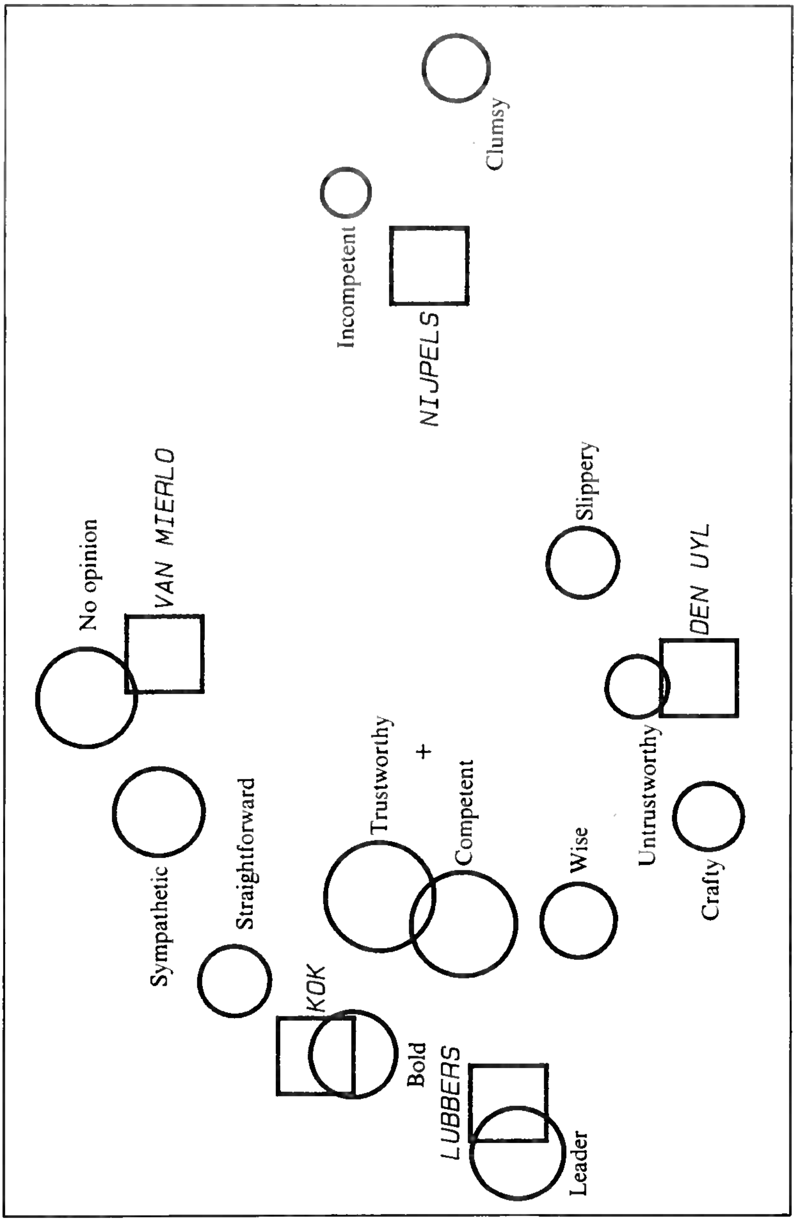
and the normalization constraint

$$\sum_{l=1}^{k} \alpha_{lt}^2 = 1, \tag{2.1}$$

$$Corr(Y_t, Y_s) = 0 \textit{ for } s = 1, \ldots, t-1.$$

The MCA *solution* consists of all $k+1$ tuples $(\mu_t, \alpha_{1t}\phi_{1t}(X_1), \ldots, \alpha_{kt}\phi_{kt}(X_k))$ for $t = 1, 2, \ldots$. The value $\alpha_{lt}\phi_{lt}(x)$ is called the *category score* on the t-th MCA component of the category x of X_l; $l = 1, \ldots, k$; $t = 1, 2, \ldots$.

The present definition of MCA may depart from other definitions given in the literature with respect to the normalization of the variable weights



The illustration shows the result of a classical correspondence analysis. Respondents were asked to characterize each of a number of Dutch politicians. The picture shows the embedding of politicians and characteristics in a two-dimensional space which preserves as closely as possible chi-square type measures of distance derived from the data.

© SPECTRAMAP, International Brainware, P.O. Box 423, 2270 CK Voorburg, telephone 070-809141/476280

$\alpha_{1t},...,\alpha_{kt}$. Also, different names for this technique are used in the literature, e.g. homogeneity analysis (GIFI [5]) and first order correspondence analysis (HILL [9]).

It follows directly from the definition that

$$Var(Y_t) = \sum_{j=1}^{k} \sum_{l=1}^{k} \alpha_{jt}\alpha_{lt} Corr(\phi_{jt}(X_j),\phi_{lt}(X_l))$$

which means that MCA only considers the bivariate marginals of the k-dimensional probability distribution of $X_1,...,X_k$. It is well-known (cf. GIFI [5]; GREENACRE [6]; HILL [9]; LEBART et al. [13]; SCHRIEVER [23]) that an MCA solution always exists and can be obtained by solving a generalized eigenvalue problem of the super matrix containing all bivariate marginal probability distributions.

MCA can be seen as a generalization of principal component analysis to nominal variables. Moreover, when $X_1,X_2,...,X_k$ are all dichotomous, e.g. $0-1$ variables, then by the normalization (2.1), $\phi_{lt}(1) = ((1-\pi_l)/\pi_l)^{\frac{1}{2}}$ and $\phi_{lt}(0) = -(\pi_l/(1-\pi_l))^{\frac{1}{2}}$, where $\pi_l = P\{X_l = 1\} = 1-P\{X_l = 0\}$ for $l = 1,2,...,k$ and $t = 1,2,...$. (Note that the signs of $\phi_{lt}(1)$ and $\phi_{lt}(0)$ may be taken arbitrary but opposite.) Hence the variance of Y_t is only maximized with respect to the variable weights $\alpha_{1t},\alpha_{2t},...,\alpha_{kt}$ for $t = 1,2,...$. Therefore, MCA in the dichotomous case is equivalent to finding the principal components of the covariance matrix of $\phi_{11}(X_1),\phi_{21}(X_2),...,\phi_{k1}(X_k)$, that is, of the correlation matrix of $X_1,X_2,...,X_k$.

For further properties, for different approaches and for applications of MCA consult DE LEEUW [2], GIFI [5], GREENACRE [6], LEBART et al. [13] and NISHISATO [20].

3. ORDERED LATENT STRUCTURE MODELS

The latent structure model we consider supposes that the responses of the individuals on the k dichotomous items (variables) $X_1,X_2,...,X_k$ can be accounted for, to a substantial degree, by one latent variable Z. It is assumed that conditionally on Z the items $X_1,X_2,...,X_k$ are stochastically independent. This assumption of *local independence* means that each individual responds independently on the items. This implies that *the (global) dependence structure between the items is caused and hence can completely be explained by variation in the latent variable.* Local independence is essential in latent structure models; if it does not hold then the latent variable cannot be distinguished from other interactions between the items.

Let the probability distribution function of Z be denoted by H; i.e. $H(z) = P\{Z \leqslant z\}$. Our results are not based on any assumption on H and thus hold for any (sampled) population. Let the response categories of each item be labeled with 1 ('correct') and 0 ('wrong'). The probability of a correct response on item X_l for an individual with latent value z is denoted by

$$\pi_l(z) = P\{X_l = 1 | Z = 1\} \text{ for } z \in \mathbb{R};\ l = 1,...,l.$$

It can be interpreted as the (local) difficulty of item X_l for this individual. The function $\pi_l(\cdot)$ is called the *trace line* of item X_l; $l = 1,\dots,k$. The unconditional probability of a correct response on item X_l,

$$\pi_l = P\{X_l = 1\} = \int_{\mathbf{R}} \pi_l(z)dH(z),$$

is the (global) difficulty of this item for the population. By local independence, the joint probability of correct responses to both item X_l and item X_j, $j \neq l$, for an individual with latent value z is given by $\pi_j(z)\pi_l(z)$. The unconditional joint probability of correct responses to both items is denoted by

$$\pi_{lj} = \int_{\mathbf{R}} \pi_l(z)\pi_j(z)dH(z) \quad \text{for } l,j = 1,\dots,k;\ l \neq j,$$

but we define $\pi_{ll} = \pi_l$ for $l = 1,\dots,k$. It is easily shown that the correlation between the items X_l and X_j equals

$$\sigma_{lj} = (\pi_{lj} - \pi_l\pi_j)/(\pi_j(1-\pi_j)\pi_l(1-\pi_l))^{\frac{1}{2}} \quad \text{for } l,j = 1,\dots,k.$$

The correlation matrix of the items $X_1, X_2, \dots, X_k$ is denoted by $\Sigma = (\sigma_{lj})$.

MOKKEN [17] imposes two natural conditions on the trace lines of the items in the test. First, he assumes for each item that the probability of a correct response increases as the individual has a higher latent value, i.e., for $l = 1,\dots,k$:

$$z_1 < z_2 \Rightarrow \pi_l(z_1) \leqslant \pi_l(z_2) \text{ and not almost everywhere } (dH)\text{—equality;} \qquad (3.1)$$

Secondly, Mokken assumes that if for one individual an item is more difficult than another item, then it must be more difficult for all individuals. In other words, the trace lines of the items may not cross each other. This means that the items in the test can be indexed such that

$$1 \leqslant l < j \leqslant k \Rightarrow \pi_l(z) \geqslant \pi_j(z) \text{ for all } z \text{ and not } dH - \text{a.e. equality.} \qquad (3.2)$$

The items in the test are then indexed from easy to difficult. Tests satisfying (3.1) and (3.2) are called *doubly monotone.* More about interpretation and examples of doubly monotone tests can be found in MOKKEN [17]. In many examples, see for instance those of Section 5, double monotonicity typically occurs in combination with the following stronger ordering of trace lines,

$$z_1 < z_2,\ 1 \leqslant l < j \leqslant k \Rightarrow \pi_l(z_1)\pi_j(z_2) \geqslant \pi_l(z_2)\pi_j(z_1). \qquad (3.3)$$

In this case the trace lines in the test are said to be *totally positive of order* 2 (TP_2). Also, a similar TP_2 property with respect to the wrong responses,

$$z_1 < z_2,\ 1 \leqslant l < j \leqslant k \Rightarrow (1-\pi_l(z_1))(1-\pi_j(z_2)) \geqslant (1-\pi_l(z_2))(1-\pi_j(z_1)) \qquad (3.4)$$

frequently holds.

The increasing property (3.1) implies that all items are positively correlated, because

$$\pi_{lj} - \pi_l\pi_j = \tfrac{1}{2}\iint_{\mathbf{R}^2} (\pi_l(z_2)-\pi_l(z_1))(\pi_j(z_2)-\pi_j(z_1))dH(z_2)dH(z_1) > 0 \qquad (3.5)$$

for $l,j = 1,...,k$. Thus the correlations σ_{lj} for $l,j = 1,...,k$ are even strictly positive. Moreover, it trivially follows from (3.2) that

$$1 \leqslant l < j \leqslant k \Rightarrow \pi_l > \pi_j \text{ and } \pi_{li} \geqslant \pi_{ji} \text{ for } i \neq j. \tag{3.6}$$

Large departures from double monotonicity violate (3.5) and (3.6) and might be detected by inspection of these properties. Notice that (3.5) and (3.6) only concern properties of the observable items.

All these assumptions concern the underlying (population) model. We make some remarks on the statistical problems which arise when one has only a finite sample of real data at the end of the next section.

4. ANALYSIS OF THE MODEL WITH MCA

Analysis with MCA of the latent structure model described in the previous section may be motivated by the interpretation of this technique and by the main result of this section. Recall that in the dichotomous case the first MCA component Y_1 equals the first principal component of the correlation matrix Σ of the items $X_1, X_2, ..., X_k$. Therefore, Y_1 'best explains' the dependence structure between the items among all linear combinations of items. Since the latent variable completely explains this dependence structure, Y_1 can be interpreted as the linear combination of items which best fits the latent variable Z in this sense. So the model will be analysed using the correct and wrong category scores, $\gamma_{l1} = \alpha_{l1}\phi_{l1}(1)$ and $\omega_{l1} = \alpha_{l1}\phi_{l1}(0)$ for $l = 1,...,k$, on the first MCA component. It follows from Section 2 that

$$\gamma_{l1} = ((1-\pi_l)/\pi_l)^{\frac{1}{2}}\,\alpha_{l1} \text{ and } \omega_{l1} = -(\pi_l/(1-\pi_l))^{\frac{1}{2}}\,\alpha_{l1}$$

for $l = 1,...,k$, where $\alpha_1 = (\alpha_{11},...,\alpha_{k1})^T$ is the eigenvector of Σ corresponding to the largest eigenvalue. (The superscript T denotes the transposition of a vector.)

Now suppose that a subset of items in the test satisfies the double monotonicity and TP_2 conditions. Then these items possess strong orderings with respect to their (local and global) difficulties. The next theorem shows that these orderings are reflected in the correct and wrong category scores of these items, even when the remaining items do not match the orderings of the items in the subset.

THEOREM. *Suppose the test consists of m items which all satisfy (3.1) with k replaced by m. Furthermore, suppose k of the items, which without loss of generality can be taken as the first k, can be indexed such that (3.2) and (3.3) hold. Then the correct scores of these k items satisfy*

$$0 < \gamma_{11} < \gamma_{21} < ... < \gamma_{k1}. \tag{4.1}$$

Similarly, if (3.1), (3.2) and (3.4) hold for these first k items, then

$$\omega_{11} < \omega_{21} < ... < \omega_{k1} < 0. \tag{4.2}$$

PROOF. Let S denote the $k \times k$ lower triangular matrix with unit elements on and below the diagonal and all other elements zero. Its inverse S^{-1} is the matrix with unit elements on the main diagonal, with elements -1 adjacent and below the diagonal (i.e. on the first sub diagonal) and all other elements zero. Denote by T the $m \times m$ block diagonal matrix with diagonal blocks S and the $(m-k) \times (m-k)$ identity matrix I. Then its inverse T^{-1} is a block diagonal matrix with diagonal blocks S^{-1} and I.

Note that the vector $\gamma_1 = (\gamma_{11},\dots,\gamma_{m1})^T$ is an eigenvector corresponding to the largest eigenvalue of the matrix C with elements $c_{lj} = (\pi_{lj} - \pi_l\pi_j)/(\pi_l(1-\pi_j))$ for $l,j = 1,\dots,m$. Since T is non-singular, γ_1 is an eigenvector corresponding to the largest eigenvalue of C iff $d = T^{-1}\gamma_1$ is an eigenvector corresponding to the largest eigenvalue of $D = T^{-1}CT$.

Under the conditions of the theorem, all elements of D turn out to be positive (i.e. larger than or equal to zero) and all elements of D^2 even turn out to be strictly positive. This can be verified as follows. The elements of the matrix $T^{-1}C = B = (b_{lj})$ equal

$$b_{lj} = (\pi_{lj} - \pi_l\pi_j)/(\pi_l(1-\pi_j)) \quad \text{for} \quad l = 1, k+1,\dots,m;\ j = 1,\dots,m,$$

$$b_{lj} = (\pi_{l-1}\pi_{lj} - \pi_l\pi_{l-1j})/(\pi_l\pi_{l-1}(1-\pi_j)) \quad \text{for} \quad l = 2,\dots,k;\ j = 1,\dots,m.$$

Since (3.1) holds for all m items, it follows that all correlations are strictly positive and hence $b_{lj} > 0$ for $l = 1, k+1,\dots,m$; $j = 1,\dots,m$. Furthermore, by (3.4), $\pi_l^{-1}\pi_l(z)$ can be interpreted as a density (with respect to the measure dH) which has the monotone likelihood ratio or TP_2 property. Therefore, since $\pi_j(z)$ is increasing in z for each $j = 1,\dots,m$, it follows from LEHMANN [14], p. 74) or KARLIN ([10], p. 22) that $\int \pi_l^{-1}\pi_l(z)\pi_j(z)dH(z)$ is increasing in l. Thus, $\pi_{l-1}\pi_{lj} - \pi_l\pi_{l-1j} \geq 0$ and hence $b_{lj} \geq 0$ for $l = 2,\dots,k$ and $j \neq l-1, l$. Obviously, $b_{ll} > 0$ and $b_{ll-1} < 0$ for $l = 2,\dots,k$. So the matrix B has positive elements except for b_{ll-1}; $l = 2,\dots,k$. But by (3.2), $\pi_{l-1} > \pi_l$ which implies that $b_{ll-1} + b_{ll} > 0$ for $l = 2,\dots,k$. Therefore, $D = BT$ has positive elements. Moreover, since $b_{lj} > 0$ for $l = 1, k+1,\dots,m$ and $j = 1,\dots,m$ and since $b_{ll-1} + b_{ll} > 0$ for $l = 2,\dots,k$, it follows that the elements in the first row of D and the elements in the first column of D are strictly positive. This implies that all elements of D^2 are strictly positive.

Application of the Perron-Frobenius theorem (cf. GANTMACHER [4], p. 53 or RAO [21], p. 46) yields that the eigenvector $d = (d_1,\dots,d_m)^T$ corresponding to the largest eigenvalue of D^2 (or of D) has strictly positive components. Since $d = T^{-1}\gamma_1$ or equivalently $d_l = \gamma_{l1}$ for $l = 1, k+1,\dots,m$ and $d_l = \gamma_{l1} - \gamma_{l-11}$ for $l = 2,\dots,k$, the result (4.1) follows. The proof of (4.2) is similar. □

The conditions (3.2), (3.3) and (3.4) of the theorem can be relaxed; see SCHRIEVER [23].

This result shows that the MCA correct and wrong category scores reflect the difficulties of the items. Since these scores do not depend on the order in which items are presented to MCA, this ordering property can be used, in combination with (3.5) and (3.6), for a first investigation of the model assumptions. Moreover, the theorem suggests that ordering the individuals according to their MCA test score Y_1 is reasonable: responding to a difficult item correctly yields a large contribution to this test score and responding to it wrongly does not cost much, whereas for an easy item it is the other way around. The test score Y_1 is a weighted sum of items with small weights for items which are less related to the latent variable than other items.

Analysis with MCA is an alternative to the method proposed by Mokken [17] in which individuals are ordered according to the unweighted sum of correct responses. It is unknown, however, in which cases which method actually works better. An advantage of the MCA approach is that our results can be generalized, in a natural fashion, to the case where items have three or more response categories. Such a generalization of Mokken's method, as discussed by Molenaar [19], is quite complicated and less natural.

In practice the probabilities π_l and π_{jl} for $j,l = 1,...,m$ have to be estimated by the relative frequencies of correct responses. Although the MCA scores based on these estimates need not reflect the difficulties of the items even when the underlying model satisfies the assumptions of the theorem, one would expect the total score Y_1 to reflect the appropriate ordering of the individuals quite well. It is, however, difficult to derive precise and useful statistical properties of such qualitative aspects.

5. EXAMPLES OF MODELS

Latent structure models for dichotomous variables studied in the literature (e.g. Andersen [1]; Fischer [3]; Lord and Novick [16] are commonly of parametric form, that is, the functional form of the trace lines is specified. Often there is, however, no evidence that the specificd functional form is actually present in the test at hand. The parametric examples below generally satisfy the double monotonicity and TP_2 conditions and, therefore, analysis both with MCA and with Mokken's method is legitimate. These examples are also discussed in Mokken [17] and Molenaar [18]; the ordering property of MCA under these models are mentioned, but not proved in Gifi [5].

In *Guttman's model* the responses on the items are deterministic functions of the latent variable. The trace lines are given by

$$\pi_l(z) = \begin{cases} 1 & \text{if } z \geqslant \delta_l \\ 0 & \text{if } z < \delta_l \end{cases} \quad \text{for } l = 1,...,k,$$

where the item parameters satisfy $\delta_1 < \delta_2 <... < \delta_k$. In this model an individual cannot respond correctly to a difficult item and wrongly to an easier item. Hence a perfect analysis is possible. Double monotonicity and TP_2 of trace lines are easily verified and thus by the theorem the correct and wrong category scores of the items (on the first MCA component) increase as the item becomes more difficult. Moreover, it is demonstrated in Schriever [23] that the correct and wrong category scores on higher MCA component are oscillating functions of the item difficulty. The practical relevance of these stronger ordering properties is, however, limited. (Slightly weaker oscillatory properties for the principal components $\alpha_1,...,\alpha_k$ of Σ are proved in GUTTMAN [7] and interpreted in GUTTMAN [8].)

A somewhat more realistic generalization of the previous model is the *latent distance model* of LAZARSFELD and HENRY [12]. The trace lines of this model satisfy

$$\pi_l(z) = \begin{cases} 1-\zeta_l & \text{if } z \geqslant \delta_l \\ \epsilon_l & \text{if } z < \delta_l \end{cases} \quad \text{for } l = 1,...,k,$$

where $\delta_1 < \delta_2 < ... < \delta_k$ and $\epsilon_l < 1-\zeta_l$ for $l = 1,...,k$. If $\epsilon_l > 0$ and $\zeta_l > 0$ for $l = 1,...,k$, then double monotonicity and TP_2 cannot hold simultaneously. But the weaker conditions for our main theorem are satisfied when $\pi_{l-1} > \pi_l$, $\pi_l\epsilon_{l-1} \geqslant \pi_{l-1}\epsilon_l$ and $\pi_l(1-\zeta_{l-1}) \geqslant \pi_{l-1}(1-\zeta_l)$ for $l = 2,...,k$ and hence (4.1) and (4.2) remain valid (cf. SCHRIEVER [23]).

In the *linear model* of Lazarsfeld the trace lines satisfy $\pi_l(z) = a_l z + b_l$ provided $0 \leqslant a_l z + b_l \leqslant 1$ for $l = 1,...,k$. The conditions (3.1) and (3.2) are for instance satisfied when $a_{l-1} \leqslant a_l$ and $b_{l-1} \geqslant b_l$ for $l = 2,...,k$. The TP_2 conditions (3.3) and (3.4) hold when $a_{l-1}b_l \leqslant a_l b_{l-1}$ and $a_{l-1}(1+b_l) \geqslant a_l(1+b_{l-1})$ for $l = 2,...,k$.

RASCH [22] developed a model in which the unweighted sum of all correct responses is sufficient for Z. The trace lines are given by $\pi_l(z) = z/(z+\delta_l)$ or $= 0$ as $z \geqslant 0$ or $z < 0$, where $0 < \delta_1 < ..., \delta_k$. This model is a special case of a model considered by Birnbaum (cf. Chap. 17-20, LORD and NOVICK [16]).). In Rasch's model the unweighted sum of correct responses 'uniformly best discriminates' the individuals (cf. MOKKEN [17], p. 141). Double monotonicity and TP_2 of trace lines is easily verified.

The last example consists of models based on shifts in distribution functions. For an univariate distribution function F the trace lines are defined by $\pi_l(z) = F(z-\delta_l)$ for $l = 1,...,k$ and $\delta_1 < \delta_2 < ... < \delta_k$. Double monotonicity is obvious. The TP_2 conditions hold when the density p of F (with respect to some measure) is log concave. Special choices of such distributions F yield well-known models, e.g. degenerate distribution (Guttman's model), logistic distribution (Rasch's model), normal distribution (models of LAWLEY [11] and LORD [15]). Other examples of such distribution functions are the gamma, Poisson and binomial distribution function.

REFERENCES

1. E.B. ANDERSEN (1980). *Discrete Statistical Models with Social Science Applications,* North-Holland Publ. Co; Amsterdam.
2. J. DE LEEUW (1984). Statistical properties of multiple correspondence analysis. Paper presented at the conference on New Multivariate Methods in Statistics. *The 1984 Joint Summer Research Conference Series in Mathematical Sciences,* June 10-16, Bowdoin College, Brunswick, Maine. (Report RR-84-06, Dept. of Data Theory, University of Leiden).
3. G.H. FISCHER (1974). *Einführung in die Theorie der Psychologischer Tests,* Huber, Bern.

4. F.R. Gantmacher (1977). *Matrix Theory, Vol II,* Chelsea, New York.
5. A. Gifi (1981). *Non-Linear Multivariate Analysis,* Dept. of Data Theory, University of Leiden.
6. M.J. Greenacre (1984). *Theory and Applications of Correspondence Analysis,* Academic Press, New York.
7. L. Guttman (1950). The principal components of scale analysis. S.A. Stouffer et al. *Measurement and Prediction,* Princeton University Press, Princeton, pp. 312-361.
8. L. Guttman (1954). The principal components of scalable attitudes. P.F. Lazarsfeld (ed.). *Mathematical Thinking in the Social Sciences,* Free Press, Glencoe, pp. 216-257.
9. M.O. Hill (1974). Correspondence analysis: a neglected multivariate method. *Appl. Statist. 23,* 340-354.
10. S. Karlin (1968). *Total Positivity I,* Stanford University Press, Stanford.
11. D.N. Lawley (1943). On problems connected with item selection and test construction. *Proc. Royal Society of Edinburgh 61,* 273-287.
12. P.F. Lazarsfeld, N.W. Henry (1968). *Latent Structure Analysis,* Houghton-Mifflin, Boston.
13. L. Lebart, A. Morineau, N. Tabard (1977). *Techniques de la Description Statistique,* Dunod, Paris. (English translation: L. Lebart, A. Morineau, K.M. Warwick (1984). *Multivariate Descriptive Statistical Analysis,* Wiley, New York.)
14. E.L. Lehmann (1959). *Testing Statistical Hypotheses,* Wiley, New York.
15. F.M. Lord (1952). *A Theory of Test Scores, Psychometric Monograph No. 7.*
16. F.M. Lord, M.R. Novick (1968). *Statistical Theories of Mental Test Scores,* Addison Wesley, Reading (Mass).
17. R.J. Mokken (1971). *A Theory and Procedure of Scale Analysis,* Mouton, The Hague.
18. I.W. Molenaar (1982). Mensen die het beter meten. *Kwantitatieve Methoden 5,* 3-29. (Dutch Statistical Society)
19. I.W. Molenaar (1982). Mokken scaling revisited. *Kwantitatieve Methoden 8,* 145-164. (Dutch Statistical Society)
20. S. Nishisato (1980). *Analysis of Categorical Data: Dual Scaling and its Applications,* University of Toronto Press, Toronto.
21. C.R. Rao (1973). *Linear Statistical Inference and its Applications,* Wiley, New York.
22. G. Rasch (1960). *Probabilistic Models for Some Intelligence and Attainment Tests.* Copenhagen: Danish Institue for Educational Research, (Reprinted 1980: University of Chicago Press.)
23. B.F. Schriever (1986). *Order Dependence, CWI Tract 20.* Centre for Mathematics and Computer Science, Amsterdam.