PSYCHOMETRIKA—VOL. 58, NO. 4, 567–574
DECEMBER 1993

A MONOTONICALLY CONVERGENT ALGORITHM FOR FACTALS

HENK A. L. KIERS

UNIVERSITY OF GRONINGEN

YOSHIO TAKANE

MCGILL UNIVERSITY

AB MOOIJAART

LEIDEN UNIVERSITY

Takane, Young, and de Leeuw proposed a procedure called FACTALS for the analysis of variables of mixed measurement levels (numerical, ordinal, or nominal). Mooijaart pointed out that their algorithm does not necessarily converge, and Nevels proposed a new algorithm for the case of nominal variables. In the present paper it is shown that Nevels' procedure is incorrect, and a new procedure for handling nominal variables is proposed. In addition, a procedure for handling ordinal variables is proposed. Using these results, a monotonically convergent algorithm is constructed for FACTALS of any mixture of variables.

Key worlds: FACTALS, alternating least squares, monotone regression, majorization.

FACTALS has been proposed by Takane, Young, and de Leeuw (1979) as a method for common factor analysis of variables of mixed measurement level. FACTALS minimizes the loss function

$$\sigma(A, D, Z) = \|Z'Z/\mathrm{N} - AA' - D^2\|^2, \tag{1}$$

where Z is an $\mathrm{N} \times n$ matrix of scores on n variables, A is an $n \times r$ loading matrix, and D^2 is an $n \times n$ diagonal matrix of unique variances. The function in (1) is minimized over A, D^2, and over those columns of Z that correspond to nominal or ordinal variables. If the variables are all numerical, Z contains standard-scores, and $Z'Z/\mathrm{N}$ contains correlations between the variables. If some of the variables are nominal or ordinal, only categorical scores are available, and these are transformed into quantitative scores by means of optimal scaling. That is, if n_j denotes the number of categories of variable j and G_j denotes the $\mathrm{N} \times n_j$ indicator matrix (with, if variable j is ordinal, the columns ordered in accordance to the ordering of the categories), then column j of Z is computed as $G_j\mathbf{y}_j$, where $\mathbf{y}_j$ is determined such that the total loss in (1) is minimized, subject to the constraints that $G_j\mathbf{y}_j$ is standardized, and, if variable j is ordinal, $\mathbf{y}_j(1) \geq \mathbf{y}_j(2) \geq \cdots \mathbf{y}_j(n_j)$.

Takane et al. (1979) proposed an algorithm for this method based on iteratively updating A, D^2, and Z subject to the constraints at hand. This algorithm was based on the equivalent problem of minimizing

The authors are obliged to Jos ten Berge for stimulating comments on an earlier version of this paper. The research of H. A. L. Kiers has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences. The research of Y. Takane has been supported by the Natural Sciences and Engineering Research Council of Canada, grant number A6394, and by the McGill-IBM Cooperative Grant.

Request for reprints should be sent to Henk A. L. Kiers, Department of Psychology, Grote Kruisstraat 2/1, 9712 TS Groningen, THE NETHERLANDS.

0033-3123/93/1200-92439$00.75/0

$$\sigma^*(A, D, Y) = \text{tr}\,[D_s^{-1/2}(Y'Y - AA' - D^2)D_s^{-1/2}]^2, \tag{2}$$

where D_s denotes the diagonal of $Y'Y$, and Y is an N × n matrix with centered but nonnormalized (quantified) scores on the n variables. The pre- and postmultiplication by $D_s^{-1/2}$ is used to avoid the explicit normalization of the quantified variables. Finding the optimal quantifications of a variable, say j, then reduces to minimizing

$$\sigma_j(\mathbf{y}_j) = \frac{(Y_j^{*\prime}G_j\mathbf{y}_j - \hat{\mathbf{r}}_j)'(Y_j^{*\prime}G_j\mathbf{y}_j - \hat{\mathbf{r}}_j)}{(\mathbf{y}_j'G_j'G_j\mathbf{y}_j)}, \tag{3}$$

where Y_j^* denotes the N × $(n - 1)$ matrix with quantified variables except the j-th, and $\hat{\mathbf{r}}_j$ is the j-th column of $AA' + D^2$, with the j-th element eliminated. Mooijaart (1984) pointed out that the procedure by Takane et al. for minimizing (3) was incorrect. Nevels (1989) derived a different solution, but as will be shown below, this procedure is also incorrect. In the present paper we derive a correct procedure for handling nominal variables, and a procedure for handling ordinal variables (for which no alternatives since Takane et al. seem to have been given). With these two procedures, and the ones for updating A and D^2 discussed by Takane et al., we are in a position to construct a monotonically convergent FACTALS algorithm for handling mixtures of numerical, nominal, and ordinal variables. We will first discuss why Nevels' procedure is incorrect, and then present correct procedures for handling nominal and ordinal variables, respectively.

Why Nevels' Procedure Fails

Nevels (1989) approached the problem of minimizing (3) over $\mathbf{y}_j$ by first defining new parameters in terms of the old ones, and then minimizing the function over these new parameters. For convenience we drop the constant N, as can be done without loss of generality. In terms of our interpretation of his procedure, he used the singular value decomposition (SVD) $Y_j^* \equiv U\Sigma^{1/2}V'$, with U an orthonormal N × N matrix, $\Sigma^{1/2}$ an N × $(n - 1)$ matrix composed of an $(n - 1) \times (n - 1)$ diagonal upper part, and zeros in the lower part of the matrix (where it is tacitly assumed that N ≥ $(n - 1)$), and V an orthonormal matrix of order $(n - 1) \times (n - 1)$. Furthermore, the definitions $\Sigma \equiv \Sigma^{1/2}\Sigma^{1/2\prime}$, $\mathbf{w} \equiv U'G_j\mathbf{y}_j$, and $\mathbf{x} \equiv \Sigma^{1/2}V'\hat{\mathbf{r}}_j$ are used. Here Y_j^* is an N × $(n - 1)$ matrix of optimally scaled data (Nevels seems to have taken the SVD of the full N × n matrix Y^*, including the j-th column, but must have meant Y_j^*, as is clear from the context). With these definitions, he elaborated (3) as

$$\sigma(\mathbf{w}) = \frac{\mathbf{w}'\Sigma\mathbf{w} - 2\mathbf{x}'\mathbf{w} + \hat{\mathbf{r}}_j'\hat{\mathbf{r}}_j}{\mathbf{w}'\mathbf{w}}, \tag{4}$$

and minimized this function over *arbitrary* $\mathbf{w}$. However, the reparametrization $\mathbf{w} \equiv U'G_j\mathbf{y}_j$ is admissible only if one can make sure that the solution for $\mathbf{w}$ can be written as $\mathbf{w} \equiv U'G_j\mathbf{y}_j$. Nevels made no such provision, and, in fact, there are many possible situations in which $\mathbf{w}$ cannot be written as $\mathbf{w} \equiv U'G_j\mathbf{y}_j$. For instance, let $n = 6$, and $n_j = 3$, which is a very natural situation. Then the solution for $\mathbf{w}$ can be written as $\mathbf{w} \equiv U'G_j\mathbf{y}_j$ only if $\mathbf{w}$ is in the column space of the 6 × 3 matrix $U'G_j$. However, the vector $\mathbf{w}$ that minimizes (4) can be anywhere in $\mathbb{R}^6$. Nothing constrains $\mathbf{w}$ to be in the column space of $U'G_j$. Hence, Nevels' procedure gives a solution for $\mathbf{w}$, but there may be no $\mathbf{y}_j$ that corresponds to this $\mathbf{w}$ (as illustrated in the appendix), thus leaving the basic problem of finding the $\mathbf{y}_j$ that minimizes (3) unsolved. In the next section we propose an alternative procedure for finding the optimal scaling of nominal variables. To do so, we use the original minimization problem (of minimizing (1)) instead of using the derived problem of minimizing (2).

A Monotonically Convergent FACTALS Algorithm

A monotonically convergent FACTALS algorithm can be constructed by alternately updating A, D^2, and Z, such that each of these steps decreases (or at least not increases) the loss function (1). For updating A and D^2, we can use the same procedure as Takane et al. (1979), or Harman and Jones' (1966) MINRES procedure supplemented with a procedure for avoiding Heywood cases (see Harman & Fukuda, 1966; Mulaik, 1972, pp. 152–153; ten Berge & Nevels, 1977). Hence we focus on updating Z. The matrix Z can be updated column by column (considering the other columns fixed). A column of Z has to be updated only if a variable j, $j = 1, \ldots, n$, is nominal or ordinal. Before considering how a column of Z is to be updated in those cases, we simplify the notation by constraining the columns of Z to have unit sums of squares instead of unit variances, and consequently drop the N in loss function (1).

If variable j is nominal, we want to update $\mathbf{y}_j$ such that (1) is minimal, considering A, D^2, and all other columns of Z, collected in Z_j^*, fixed. The problem then reduces to minimizing

$$\mathrm{f}(\mathbf{y}_j) = \|Z_j^{*\prime} G_j \mathbf{y}_j - \hat{\mathbf{r}}_j\|^2, \tag{5}$$

subject to the constraints that $G_j\mathbf{y}_j$ be centered, and that $\mathbf{y}_j'G_j'G_j\mathbf{y}_j = 1$. First note that Z_j^* is centered columnwise, and hence G_j in (5) can be replaced by JG_j, where $J \equiv I - \mathbf{1}\mathbf{1}'/\mathrm{N}$ is the centering operator, and $\mathbf{1}$ a vector with unit elements. Now we can express $JG_j\mathbf{y}_j$ in terms of an orthonormal basis B for JG_j as $B\mathbf{t}$, and hence we have to minimize

$$\mathrm{g}(\mathbf{t}) = \|Z_j^{*\prime} B\mathbf{t} - \hat{\mathbf{r}}_j\|^2, \tag{6}$$

over $\mathbf{t}$, subject to the constraint $\mathbf{y}_j'G_j'G_j\mathbf{y}_j = \mathbf{y}_j'G_j'JG_j\mathbf{y}_j = \mathbf{t}'B'B\mathbf{t} = \mathbf{t}'\mathbf{t} = 1$. This problem is equivalent to Mosier's oblique Procrustes problem, and has been solved by ten Berge and Nevels (1977). To obtain the updated category quantifications (in $\mathbf{y}_j$), we have to solve $\mathbf{y}_j$ from

$$JG_j\mathbf{y}_j = G_j\mathbf{y}_j = B\mathbf{t}_0, \tag{7}$$

where $\mathbf{t}_0$ denotes the $\mathbf{t}$ obtained from the ten Berge and Nevels procedure applied to the problem of minimizing (6). From the regression of $B\mathbf{t}_0$ on G_j, we obtain the unique solution $\mathbf{y}_j^0 = (G_j'G_j)^{-1}G_j'B\mathbf{t}_0$. To check if $\mathbf{y}_j^0$ indeed satisfies (7), we first write $B = JG_jT_j$ for a certain matrix T_j. Then the second equality in (7) follows from $G_j\mathbf{y}_j^0 = G_j(G_j'G_j)^{-1}G_j'JG_jT_j\mathbf{t}_0 = JG_jT_j\mathbf{t}_0 = B\mathbf{t}_0$, where it is used that the projection of JG_j on G_j yields JG_j, because JG_j lies in the subspace spanned by G_j. From this result, it follows at once that $G_j\mathbf{y}_j^0$ is centered, thus establishing the first equality in (7).

The procedure for updating $\mathbf{y}_j$ in case variable j is ordinal is more complicated. We now minimize (5) over $\mathbf{y}_j$, subject to the constraints that $G_j\mathbf{y}_j$ be centered, $\mathbf{y}_j'G_j'G_j\mathbf{y}_j = 1$, and the elements of $\mathbf{y}_j$ are weakly ordered. Because of the ordering constraint it is no longer useful to reparametrize $\mathbf{y}_j$ by a vector $\mathbf{t}$. Instead, we propose a different approach in which centering will be maintained automatically. In analogy to a procedure by Meulman (1986, pp. 147–149), we use the fact that we can update $\mathbf{y}_j$ so that $\mathrm{f}(\mathbf{y}_j)$ *decreases* (or at least does not increase) by minimizing a function that majorizes $\mathrm{f}(\mathbf{y}_j)$. To find such a majorizing function, we view f as a function of $\mathbf{q} = G_j\mathbf{y}_j$, and expand it as

$$\begin{aligned} f(\mathbf{q}) &= \|Z_j^{*\prime}\mathbf{q} - \hat{\mathbf{r}}_j\|^2 \\ &= \hat{\mathbf{r}}_j'\hat{\mathbf{r}}_j - 2\hat{\mathbf{r}}_j'Z_j^{*\prime}\mathbf{q} + \text{tr } Z_j^*Z_j^{*\prime}\mathbf{q}\mathbf{q}', \end{aligned} \tag{8}$$

which is a special case of Kiers' (1990) general function f(X) (with $\mathbf{q}$ instead of X). A function that majorizes f($\mathbf{q}$) is (see Kiers, p. 421)

$$g(\mathbf{q}) = c_1 + \alpha(\|\mathbf{q}^0 - (2\alpha)^{-1}(-2Z_j^*\hat{\mathbf{r}}_j + 2Z_j^*Z_j^{*\prime}\mathbf{q}^0) - \mathbf{q}\|^2 + c_2), \tag{9}$$

where c_1 and c_2 are constants for $\mathbf{q}$, α is the first eigenvalue of $Z_j^*Z_j^{*\prime}$, and $\mathbf{q}^0$ denotes the current (or "old") value for $\mathbf{q}$. Because g($\mathbf{q}$) majorizes f($\mathbf{q}$), and $g(\mathbf{q}^0) = f(\mathbf{q}^0)$, minimizing g($\mathbf{q}$) will yield an update for $\mathbf{q}$ that decreases (or at least does not increase) f($\mathbf{q}$). Re-expressing $\mathbf{q}$ and $\mathbf{q}^0$ in $\mathbf{y}_j$ and $\mathbf{y}_j^0$, respectively, we end with the problem of minimizing

$$\begin{aligned} h(\mathbf{y}_j) &= \|G_j\mathbf{y}_j^0 - (2\alpha)^{-1}(-2Z_j^*\hat{\mathbf{r}}_j + 2Z_j^*Z_j^{*\prime}G_j\mathbf{y}_j^0) - G_j\mathbf{y}_j\|^2 \\ &= \|(G_j\mathbf{y}_j^0 + \alpha^{-1}Z_j^*\hat{\mathbf{r}}_j - \alpha^{-1}Z_j^*Z_j^{*\prime}G_j\mathbf{y}_j^0) - G_j\mathbf{y}_j\|^2 \\ &= \|\mathbf{z} - G_j\mathbf{y}_j\|^2, \end{aligned} \tag{10}$$

where $\mathbf{z} \equiv (G_j\mathbf{y}_j^0 + \alpha^{-1}Z_j^*\hat{\mathbf{r}}_j - \alpha^{-1}Z_j^*Z_j^{*\prime}G_j\mathbf{y}_j^0)$, over $\mathbf{y}_j$ subject to the constraints that $G_j\mathbf{y}_j$ be centered, $\mathbf{y}_j'G_j'G_j\mathbf{y}_j = 1$, and the elements of $\mathbf{y}_j$ are weakly ordered. Noting that $\mathbf{z}$ is centered, we can observe that this problem is equivalent to the normalized monotone regression problem encountered in PRINCIPALS (Young, Takane, & de Leeuw, 1978; also, see de Leeuw, Young, & Takane, 1976). Applying their procedure to the problem of minimizing (10), we find an update for $\mathbf{y}_j$ for the case where variable j is ordinal.

Implementing the above procedures for updating category quantifications for nominal and ordinal variables in the FACTALS algorithm described by Takane et al. (1979), we obtain an algorithm that monotonically decreases the FACTALS loss function. Because the FACTALS loss function is bounded below by zero, the algorithm must converge to a stable function value.

Exemplary Analysis

The algorithm has been programmed in the matrix language PCMATLAB. Specifically, the algorithm is started at prespecified values for the scalings in $\mathbf{y}_j$, $j = 1, \ldots, m$; then the modified MINRES approach is used for updating A; next D^2 is computed as the diagonal of $R - AA'$, where R is the correlation matrix for the variables (based on the current scalings); finally, the scalings of the variables are updated by the procedures described in the previous section. In an attempt to accelerate the program, the update of the ordinal variables is repeated until the total absolute difference between consecutive solutions for $\mathbf{y}_j$ becomes smaller than some prespecified value (here .000001), or a maximum of 20 such inner iterations has been performed. The complete cycle of updating A, D^2 and the scalings is repeated until the function value decreases by less than a prespecified proportion of the function value (here .0001), or the function becomes smaller than .01 times this proportion.

As an example, we analyzed Hartigan's (1975, p. 228) hardware data, consisting of scores of 24 objects on five nominal variables and one ordinal variable (the fifth). We obtained solutions with one and two factors. Both solutions were replicated by using different starts. To give an impression of the convergence rate, Table 1 lists some important aspects of the iteration history of the first one-dimensional FACTALS anal-

TABLE 1

Iteration History of The One-Dimensional FACTALS Analysis of The Hartigan Data

iteration number	number of inner iterations for the ordinal variable	function value
0	–	.7107
1	20 (maximum)	.3197
2	20 (maximum)	.2216
3	20 (maximum)	.1832
4	20 (maximum)	.1558
5	13	.1272
6	1	.1213
7	1	.1202
8	1	.1193
9	1	.1186
10	1	.1179
11	1	.1174
12	1	.1170
13	1	.1168
14	1	.1167
15	1	.1166
16	1	.1166

ysis. It can be seen that in the first steps, repeated updating of the scale values of the ordinal variable is necessary, but quite soon one inner iteration per major cycle suffices. This phenomenon of relatively quick convergence with few inner iterations after a few main cycles was also observed in one-dimensional FACTALS analyses of other data sets. For the Hartigan data (as well as for other data), higher dimensional solutions with nearly perfect fit took many more major iterations. For example, a two-dimensional FACTALS analysis of the Hartigan data required 819 iterations before the function value dropped below 10^{-6}.

Without attempting to give a full account of the FACTALS analysis of the Hartigan data, in Table 2 we briefly report the matrices A and D^2 obtained from the one- and two-dimensional analyses, as well as the optimal correlation matrices. Clearly, the two solutions differ considerably. The one-dimensional solution seems the most reasonable one, because the obtained factor reflects a strong association between the first three variables, a result which is similar to that of other analyses of these data (e.g., Gifi, 1990, pp. 128–135). In the two-dimensional solution the first three variables are scaled in such a way that they hardly correlate with each other, contradicting the considerable association between these variables. Apparently, FACTALS obtained a good fit here by finding scalings such that the variables were correlated only mildly, and by modeling

TABLE 2

Matrices A, D^2 and R From The One- and Two-Dimensional FACTALS Analyses of The Hartigan Data

$r = 1$

A	D^2	R					
0.98	0.03	1.00	0.98	1.00	0.58	0.00	-0.22
1.00	0.00	0.99	1.00	0.99	0.61	-0.05	-0.25
0.99	0.01	1.00	0.99	1.00	0.60	-0.06	-0.22
0.61	0.63	0.58	0.61	0.60	1.00	-0.17	-0.26
-0.05	1.00	0.00	-0.05	-0.06	-0.17	1.00	-0.13
-0.25	0.94	-0.23	-0.25	-0.22	-0.26	-0.13	1.00

$r = 2$

A		D^2	R					
0.63	0.08	0.60	1.00	0.16	0.19	0.58	0.00	-0.22
0.21	0.45	0.76	0.16	1.00	-0.07	0.15	0.42	-0.23
0.34	-0.31	0.78	0.19	-0.07	1.00	0.35	-0.35	0.00
0.94	-0.10	0.11	0.58	0.15	0.35	1.00	-0.21	-0.26
-0.12	0.99	0.01	0.00	0.42	-0.35	-0.21	1.00	-0.31
-0.31	-0.36	0.77	-0.22	-0.23	0.00	-0.26	-0.31	1.00

most variance as unique variance, which is reflected by the relatively high unique variances in the two-dimensional solution.

Discussion

As the results of the FACTALS analysis of the Hartigan data indicate, one has to be very careful in interpreting FACTALS solutions. Especially with nominal variables, the large amount of freedom in scaling the variables may cause the method to capitalize on finding unique factors rather than common, even though strong common factors may clearly be present. More research will be needed to understand these phenomena and to assess the usefulness of FACTALS in practice. The main purpose of the present paper was to make such research possible by providing a working algorithm for FACTALS of mixtures of nominal, ordinal, and numerical variables.

Appendix

The following example shows that Nevels' procedure cannot always solve the problem of minimizing (3). Suppose, at a particular stage of his procedure,

$$Y_j^* = 2^{-1/2}\begin{pmatrix} 1 & 1 \\ 0 & -1 \\ -1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \hat{\mathbf{r}}_j = 2^{-1/2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad G_j = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then, Nevels' solution for $\mathbf{w}$ must be

$$\mathbf{w}_0 = \begin{pmatrix} 6^{-1/2} \\ 2^{-1/2} \\ 0 \\ 2/3^{1/2} \end{pmatrix},$$

because $\mathbf{w}_0$ globally minimizes (4), since it yields a function value of 0, as can be verified as follows. First we calculate the singular value decomposition of $Y_j^* = U\Sigma^{1/2}V'$, which yields

$$\tilde{U} = \begin{pmatrix} 2/6^{1/2} & 0 & 0 & 3^{-1/2} \\ -1/6^{1/2} & 2^{-1/2} & 0 & 3^{-1/2} \\ -1/6^{1/2} & -2^{-1/2} & 0 & 3^{-1/2} \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad \Sigma^{1/2} = \begin{pmatrix} 6^{1/2}/2 & 0 \\ 0 & 2^{-1/2} \\ 0 & 0 \\ 0 & 0 \end{pmatrix},$$

and

$$V = \begin{pmatrix} 2^{-1/2} & 2^{-1/2} \\ 2^{-1/2} & -2^{-1/2} \end{pmatrix},$$

where U can be taken equal to the first two columns of $\tilde{U}$ supplemented by any rotation of the last two columns of $\tilde{U}$; hence,

$$U = \tilde{U}\begin{pmatrix} I & 0 \\ 0 & T \end{pmatrix},$$

for an arbitrary orthonormal matrix T. With these expressions, we can compute

$$\Sigma \equiv \Sigma^{1/2}\Sigma^{1/2\prime} = \begin{pmatrix} 3/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{and} \quad \mathbf{x} \equiv \Sigma^{1/2}V'\hat{\mathbf{r}}_j = \begin{pmatrix} 6^{1/2}/4 \\ 2^{1/2}/4 \\ 0 \\ 0 \end{pmatrix},$$

and it can now be computed that $\mathbf{w}_0'\Sigma\mathbf{w}_0 = 1/2$, $\mathbf{x}'\mathbf{w}_0 = 1/2$, and $\hat{\mathbf{r}}_j'\hat{\mathbf{r}}_j = 1/2$; hence, $\sigma(\mathbf{w}_0) = 0$. This shows that $\mathbf{w}_0$ globally minimizes (4). Now, we should find the vector $\mathbf{y}_j$ such that

$$\mathbf{w}_0 = \begin{pmatrix} 6^{-1/2} \\ 2^{-1/2} \\ 0 \\ 2/3^{1/2} \end{pmatrix} = U'G_j\mathbf{y}_j.$$

We can calculate

$$\tilde{U}'G = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 3^{1/2} & 0 \end{pmatrix},$$

and hence, with

$$U = \tilde{U}\begin{pmatrix} I & 0 \\ 0 & T \end{pmatrix},$$

find that the first two rows of $U'G$ are zero. It follows that we can never find a $\mathbf{y}_j$ such that $\mathbf{w}_0 = U'G_j\mathbf{y}_j$, demonstrating that Nevels' procedure, which if it minimizes (4) should give the unique solution $\mathbf{w}_0$ (see Nevels, 1989, pp. 343–344), fails to solve for the $\mathbf{y}_0$ that minimizes (3).

References

de Leeuw, J., Young, F. W., & Takane, Y. (1976). Additive structure in qualitative data: An alternating least squares method with optimal scaling features. *Psychometrika, 41*, 471–503.

Gifi, A. (1990). *Nonlinear multivariate analysis*. Chichester: Wiley.

Harman, H. H., & Fukuda, Y. (1966). Resolution of the Heywood case in the MINRES solution. *Psychometrika, 31*, 563–571.

Harman, H. H., & Jones, W. H. (1966). Factor analysis by minimizing residuals (MINRES). *Psychometrika, 31*, 351–368.

Hartigan, J. A. (1975). *Clustering algorithms*. New York: Wiley.

Kiers, H. A. L. (1990). Majorization as a tool for optimizing a class of matrix functions. *Psychometrika, 55*, 417–428.

Meulman, J. J. (1968). *A distance approach to nonlinear multivariate analysis*. Leiden: DSWO Press.

Mooijaart, A. (1984). The nonconvergence of FACTALS: A nonmetric common factor analysis. *Psychometrika, 49*, 143–145.

Mulaik, S. A. (1972). *The foundations of factor analysis*. New York: McGraw-Hill.

Nevels, K. (1989). An improved solution for FACTALS: a nonmetric common factor analysis. *Psychometrika, 54*, 339–343.

Takane, Y., Young, F. W., & de Leeuw, J. (1979). Nonmetric common factor analysis: An alternating least squares method with optimal scaling features. *Behaviormetrika, 6*, 45–56.

ten Berge, J. M. F., & Nevels, K. (1977). A general solution to Mosier's oblique Procrustes problem. *Psychometrika, 42*, 593–600.

Young, F. W., Takane, Y., & de Leeuw, J. (1978). The principal components of mixed measurement level multivariate data: An alternating least squares method with optimal scaling features. *Psychometrika, 43*, 279–281.

Manuscript received 3/23/92
Final version received 9/22/92