PSYCHOMETRIKA—VOL. 54, NO. 3, 515–521
SEPTEMBER 1989

AN ALTERNATING LEAST SQUARES ALGORITHM FOR FITTING THE TWO- AND THREE-WAY DEDICOM MODEL AND THE IDIOSCAL MODEL

HENK A. L. KIERS

UNIVERSITY OF GRONINGEN

The DEDICOM model is a model for representing asymmetric relations among a set of objects by means of a set of coordinates for the objects on a limited number of dimensions. The present paper offers an alternating least squares algorithm for fitting the DEDICOM model. The model can be generalized to represent any number of sets of relations among the same set of objects. An algorithm for fitting this three-way DEDICOM model is provided as well. Based on the algorithm for the three-way DEDICOM model an algorithm is developed for fitting the IDIOSCAL model in the least squares sense.

Key words: DEDICOM, three mode data analysis, IDIOSCAL.

Harshman (1978) has proposed a family of models for analyzing data matrices that are intrinsically asymmetric. The family of models is called DEcomposition into DIrectional COMponents (DEDICOM). The simplest member of this family is the single-domain DEDICOM model. Because this is the only member of the family that will be considered here it will be referred to as DEDICOM for short. For an extensive description of this model we refer to Harshman, Green, Wind and Lundy (1982). A brief description of the model will be given here.

According to the DEDICOM model a square data matrix X, containing entries x_{ij} representing the (asymmetric) relation of object i to object j, is decomposed as

$$X = ARA' + E, \tag{1}$$

where A is an n by p $(p < n)$ matrix of weights (or "loadings") for the n objects on p dimensions or aspects, R is a square matrix of order p, representing (asymmetric) relations among the p dimensions, and E is a matrix with entries e_{ij} representing the part of the relation of object i to object j that is not explained by the model (the error-part). The objective of fitting this model to the data is to explain the data by means of relations among as small a number of dimensions as possible. These dimensions can be considered as "aspects" of the objects. The loadings of the objects on these aspects are given by matrix A. The entries in matrix A indicate the importance of the aspects for the objects. The dimensionality of R and A, and hence the number of aspects to be determined, is to be based on some external criterion, defined by the user.

Several methods have been proposed for fitting data to the model (1) in the least squares sense. These algorithms have been discussed in detail by Harshman and Kiers (1987). Some of these algorithms are alternating least squares (ALS) algorithms and fit a model in which the left and the right hand matrix A in ARA' are treated independently. Mostly, upon convergence of the algorithm the left and right hand A are equal so that the proper model has been fitted. However, equality of left and right hand A is by no

The author is obliged to Jos ten Berge and Richard Harshman.

Requests for reprints should be sent to Henk A. L. Kiers, Department of Psychology, Grote Markt 31/32, 9712 HV Groningen, THE NETHERLANDS.

0033-3123/89/0900-9099$00.75/0

means guaranteed. Other algorithms, especially gradient algorithms, may be problematic because they do not necessarily decrease a function monotonically, and hence convergence is not guaranteed for them. An example where such an algorithm failed to converge, and oscillated between two solutions is available from the author upon request. In the present paper an ALS algorithm is described that does not differentiate between left and right hand matrix A and is proven to converge monotonically.

Harshman and Lundy (1984, p. 187ff) have described a generalization of the DEDICOM model for three-way data. That is, a model has been described for representing m matrices ($m > 1$) originating from m instances (e.g., observers or occasions) with asymmetric relations among the same set of objects by means of relations among a single set of dimensions. This three-way DEDICOM model is a special case of the PARAFAC-2 model (Harshman & Lundy). According to this generalized DEDICOM model the k-th data matrix X_k is represented by

$$X_k = AD_kRD_kA' + E_k, \tag{2}$$

where A and R are defined as in (1), while D_k is a diagonal matrix with weights to represent the dimensions in the k-th instance.

The model (2) described above is a rather heavily constrained model, because it specifies both the A and the R matrices to be equal for every instance, except for row- and column-scalings. However, a less constrained three-way generalization of DEDICOM is feasible in such a way that this model is an asymmetric variant of the IDIOSCAL model in its usual scalar product form (Carroll & Wish, 1974). That is, the positive semi-definite matrix in IDIOSCAL defining the metric for the space in which scalar products between objects are computed for the k-th instance is replaced by an asymmetric matrix. The latter matrix cannot be considered to define a metric anymore. It merely provides relations among the dimensions at the k-th instance. This three-way DEDICOM model specifies X_k as

$$X_k = AR_kA' + E_k, \tag{3}$$

which is considerably less constrained than model (2). In the present paper an algorithm for the least squares fitting of the model specified by (3) is given as well by means of a simple extension of the algorithm for fitting the DEDICOM model (1).

When Carroll and Chang (1970, 1972) proposed the IDIOSCAL model, they described methods for obtaining approximate solutions for fitting the model. They remarked (Carroll & Chang, 1970, p. 309) that a procedure similar to their CANDECOMP procedure for fitting the INDSCAL model could be developed, but did not elaborate this. Since then not very much attention seems to have been paid to fitting the IDIOSCAL model in the least squares sense. All techniques reviewed by de Leeuw and Pruzansky (1978, p. 483ff) yield approximate solutions only.

Kroonenberg and de Leeuw (1980, p. 78) suggest that their TUCKALS2 algorithm for fitting the Tucker-2 model can be used to test the appropriateness of the IDIOSCAL model. Although they do not elaborate this approach, it is readily verified that an ALS algorithm for fitting the IDIOSCAL model can be constructed by means of a slight modification of the TUCKALS2 algorithm. However, the latter procedure again is based on differentiating between left and right hand matrices A, while it is not guaranteed that the left and right hand matrices A will be equal upon convergence. (An example where the left and right hand matrices A differ at convergence of the TUCKALS2 algorithm is available from the author upon request.) Therefore, it seems that no straightforward techniques for fitting the IDIOSCAL model in the least squares sense have been developed yet. This gap is filled by means of a slight modification of the algorithm for fitting model (3). First, however, the algorithm for fitting the DEDICOM model will be described.

An ALS Algorithm for Fitting the DEDICOM Model

In order to fit the DEDICOM model in the least squares sense the function

$$f(A, R) = \|X - ARA'\|^2 \tag{4}$$

has to be minimized over A and R. Because every linear one-to-one transformation of A can be undone by an inverse transformation of R, matrix A can be constrained to be column-wise orthonormal, that is, $A'A = I_p$, without loss of generality. Then (4) can be elaborated as

$$f(A, R) = \text{tr } X'X - 2 \text{ tr } X'ARA' + \text{tr } R'R. \tag{5}$$

Function $f(A, R)$ can be minimized by alternatingly minimizing f over R while A is fixed, and decreasing f over A while R is fixed, until f is not decreased anymore by means of this procedure.

The step of minimizing $f(A, R)$ over R is accomplished by choosing $R = A'XA$. That this choice of R minimizes $f(A, R)$ for fixed A can be seen as follows. Function $f(A, R)$ can be elaborated as

$$f(A, R) = \text{tr } X'X - \text{tr } A'XAA'X'A + \|A'XA - R\|^2. \tag{6}$$

Clearly, $f(A, R)$ is minimized over R by choosing $R = A'XA$.

In the sequel the procedure for decreasing function $f(A, R)$, for fixed R, over A, subject to $A'A = I_p$, is described. Decreasing $f(A, R)$ over A while R is fixed is equivalent to decreasing

$$g(A) = -2 \text{ tr } X'ARA' = -2 \text{ tr } A'X'AR \tag{7}$$

over A. Let $\mathbf{a}_j$ be column j of matrix A, then (7) can be rewritten as

$$g(A) = -2 \sum_{j=1}^{p} \sum_{l=1}^{p} \mathbf{a}_j'X'\mathbf{a}_l r_{lj}. \tag{8}$$

The constraint $A'A = I_p$ can be reformulated as $\mathbf{a}_j'\mathbf{a}_l = \delta_{jl}$, for $j, l = 1, \ldots, p$, in which δ_{jl} is the Kronecker symbol, that is, $\delta_{jl} = 0$ when $j \neq l$ and $\delta_{jl} = 1$ when $j = l$.

Decreasing (8) is achieved by successively minimizing $g(A)$ over the columns of A separately, subject to the constraint $\mathbf{a}_j'\mathbf{a}_l = \delta_{jl}$, for $j, l = 1, \ldots, p$. Minimizing $g(A)$ over $\mathbf{a}_i$ subject to $\mathbf{a}_i'\mathbf{a}_j = \delta_{ij}$, for $j = 1, \ldots, p$, is equivalent to minimizing

$$\begin{aligned} g(\mathbf{a}_i) &= -2\left(\mathbf{a}_i' \sum_{l \neq i} X'\mathbf{a}_l r_{li} + \sum_{j \neq i} \mathbf{a}_j'X'\mathbf{a}_i r_{ij} + \mathbf{a}_i'X'\mathbf{a}_i r_{ii} \right) + c \\ &= -2\left(\mathbf{a}_i' \sum_{l \neq i} X'\mathbf{a}_l r_{li} + \mathbf{a}_i' \sum_{j \neq i} X\mathbf{a}_j r_{ij} + \mathbf{a}_i'X'\mathbf{a}_i r_{ii} \right) + c \\ &= -2\left(\mathbf{a}_i' \sum_{j \neq i} (X'\mathbf{a}_j r_{ji} + X\mathbf{a}_j r_{ij}) + \mathbf{a}_i'X'\mathbf{a}_i r_{ii} \right) + c, \end{aligned} \tag{9}$$

where c is a constant with respect to $\mathbf{a}_i$. Obviously, $\mathbf{a}_i'X'\mathbf{a}_i = \mathbf{a}_i'X\mathbf{a}_i = \mathbf{a}_i'\{1/2(X + X')\}\mathbf{a}_i$. Hence

$$g(\mathbf{a}_i) = -2\ \mathbf{a}_i' \sum_{j \neq i} (X'\mathbf{a}_j r_{ji} + X\mathbf{a}_j r_{ij}) + \mathbf{a}_i'(-r_{ii}(X + X'))\mathbf{a}_i + c. \tag{10}$$

The constraint $\mathbf{a}_i'\mathbf{a}_j = \delta_{ij}$, for $j = 1, \ldots, p$, forces $\mathbf{a}_i$ to be a length one vector in the orthogonal complement space of the remaining columns of A. Let A_{-i} denote the matrix with the elements of A except for column i, which contains zero elements. The columns of $(I - A_{-i}A_{-i}')$ span the orthogonal complement space of $\mathbf{a}_i$. Let the n by $(n - p + 1)$ matrix B contain columns that span an orthonormal basis for this space. The restriction $\mathbf{a}_i'\mathbf{a}_j = \delta_{ij}$, for $j = 1, \ldots, p$, is tantamount to $\mathbf{a}_i = B\mathbf{v}$ for some vector $\mathbf{v}$ of length one and order $(n - p + 1)$.

The problem of minimizing (10) subject to the constraint $\mathbf{a}_i'\mathbf{a}_j = \delta_{ij}, j = 1, \ldots, p$, can be reformulated as minimizing the function

$$h(\mathbf{v}) = -2\ \mathbf{v}'B' \sum_{j \neq i} (X'\mathbf{a}_j r_{ji} + X\mathbf{a}_j r_{ij}) + \mathbf{v}'B'(-r_{ii}(X + X'))B\mathbf{v}, \tag{11}$$

over $\mathbf{v}$, subject to $\mathbf{v}'\mathbf{v} = 1$. Let an eigendecomposition of $B'(-r_{ii}(X + X'))B$ be given by $B'(-r_{ii}(X + X'))B = UDU'$, let vector $\mathbf{z}$ be defined by $\mathbf{z} \equiv U'B'\Sigma_{j \neq i}(X'\mathbf{a}_j r_{ji} + X\mathbf{a}_j r_{ij})$ and let $\mathbf{w} \equiv U'\mathbf{v}$, then minimizing $h(\mathbf{v})$ subject to $\overline{\mathbf{v}}'\mathbf{v} = 1$ is tantamount to minimizing

$$k(\mathbf{w}) = -2\ \mathbf{w}'\mathbf{z} + \mathbf{w}'D\mathbf{w}, \tag{12}$$

over $\mathbf{w}$, subject to $\mathbf{w}'\mathbf{w} = 1$.

The problem of minimizing $k(\mathbf{w})$ subject to $\mathbf{w}'\mathbf{w} = 1$ has been solved by ten Berge and Nevels (1977, p. 594–597). Although in their problem matrix D is positive semi-definite, exactly the same algorithm can be used for solving the problem when D is an arbitrary diagonal matrix with elements in weakly descending order. The reason for this is that in deriving their algorithm, ten Berge and Nevels never used the restriction that the elements of D be nonnegative.

The problem of minimizing $g(\mathbf{a}_i)$ subject to the constraint $\mathbf{a}_i'\mathbf{a}_j = \delta_{ij}, j = 1, \ldots, p$, is now solved as follows. Let $\mathbf{w}_0$ be the vector found by the ten Berge and Nevels procedure that minimizes $k(\mathbf{w})$ subject to $\mathbf{w}'\mathbf{w} = 1$. Then $\mathbf{a}_{i0} = BU\mathbf{w}_0$ minimizes $g(\mathbf{a}_i)$ subject to $\mathbf{a}_i'\mathbf{a}_j = \delta_{ij}, j = 1, \ldots, p$. When all the columns of A are updated in this way, $f(A, R)$ is decreased over A while R is fixed.

The procedures for updating R and for updating A both decrease function $f(A, R)$, and $f(A, R)$ is bounded from below by zero. Therefore, alternating these procedures will finally converge to an A and R such that $f(A, R)$ cannot be improved any further by either of these procedures. In this way convergence of this algorithm is guaranteed.

The algorithm described here is an alternating least squares algorithm, because it alternates the least squares procedures of minimizing f over R while A is fixed, and minimizing f over each column of A, while R and the other columns of A are fixed.

The constraint that A be column-wise orthonormal has been imposed in order to simplify the problem. A different constraint on A could also simplify the problem (although to a lesser extent) and obviate the use of orthogonal complement spaces. This is the constraint $\text{Diag}(A'A) = I_p$, which may be imposed without loss of generality, also. This leads to a different algorithm which is beyond the scope of the present paper.

Performance of the Column-Wise DEDICOM Algorithm

The algorithm described above for fitting the DEDICOM model has been programmed on a CDC Cyber. As a starting configuration for the matrix A we have chosen the matrix containing the p eigenvectors that are associated with the eigenvalues of the symmetric part of X, that are the largest in absolute value.

Harshman and Kiers (1987) compared the above described algorithm with four other algorithms on twenty random data sets. The latter algorithms were based on differentiating between left and right hand matrices A. The column-wise algorithm might be expected to be slower than the other algorithms, because each step of updating a column of A requires the computation of a complete eigendecomposition of a matrix of order $(n - p + 1)$ by $(n - p + 1)$ and minimization of the function k by the ten Berge and Nevels procedure. However, the results of Harshman and Kiers indicate that the column-wise algorithm is faster than one of the other four algorithms, and only slightly slower than the other three algorithms. The column-wise algorithm took an average of 5.0 cp seconds computation time, whereas the average computation times for the other four algorithms were 4.2, 4.0, 5.6 and 3.8 seconds.

It might be expected that the column-wise algorithm would converge to local minima more often than the other four algorithms, because updating A column-wise involves smaller changes of the A matrix than updating the complete matrix A, as the other algorithms do. However, the column-wise algorithm did not converge to local minima more often than the other four algorithms did on the average. Moreover, in those cases restarting with a different starting configuration led to a better solution for the column-wise algorithm.

Although only tentative conclusions can be made on the basis of this limited number of test runs, it seems safe to conclude that the DEDICOM algorithm developed here is not particularly problematic. That is, it is not particularly slow and it does not very often converge to a local minimum. Moreover, when it does converge to a local minimum, it is likely that in several new runs with different starting configurations the global minimum will be found. For this reason, it is advised to use more than one starting configuration.

Three-Way DEDICOM and IDIOSCAL

In order to fit the three-way DEDICOM model as formulated by (3) in the least squares sense, a procedure similar to the one for fitting the DEDICOM model can be used. The function to be minimized in order to fit the three-way DEDICOM model (3) is

$$f_1(A, R_1, \ldots, R_m) = \sum_{k=1}^{m} \|X_k - AR_kA'\|^2. \tag{13}$$

The constraint that A be column-wise orthonormal can again be imposed without loss of generality. The minimization of f_1 is accomplished by an ALS algorithm alternating over the columns of A, and the matrices $R_1, \ldots, R_m$. The update for R_k is given by $R_k = A'X_kA$. The derivation of the updates for the columns of the matrix A in the three-way DEDICOM model is analogous to that for the updates of the columns of matrix A in the DEDICOM model, as can be seen as follows.

Minimizing f_1 over A, for fixed $R_1, \ldots, R_m$, subject to $A'A = I_p$, is equivalent to minimizing

$$g_1(A) \underset{-}{=} -2 \sum_{k=1}^{m} \operatorname{tr} X_k'AR_kA' = -2 \sum_{k=1}^{m} \operatorname{tr} A'X_k'AR_k \tag{14}$$

over A. Let $\mathbf{a}_j$ be column j of matrix A, then (14) can be rewritten as

$$g_1(A) = -2 \sum_{k=1}^{m} \sum_{j=1}^{p} \sum_{l=1}^{p} \mathbf{a}_j' X_k' \mathbf{a}_l r_{ljk}, \tag{15}$$

where r_{ljk} refers to the entry in cell (l, j) of matrix R_k.

Analogously to decreasing $g(A)$, see (8), decreasing $g_1(A)$ is achieved by successively minimizing $g_1(A)$ over the columns of A separately, subject to the constraint $\mathbf{a}_j'\mathbf{a}_l = \delta_{jl}$, for $j, l = 1, \ldots, p$. Function $g_1(A)$ can be rewritten as

$$g_1(\mathbf{a}_i) = -2\left(\mathbf{a}_i' \sum_{l \neq i} \sum_{k=1}^{m} X_k' \mathbf{a}_l r_{lik} + \sum_{j \neq i} \mathbf{a}_j' \sum_{k=1}^{m} X_k' \mathbf{a}_i r_{ijk} + \sum_{k=1}^{m} \mathbf{a}_i' X_k' \mathbf{a}_i r_{iik} \right) + c_1$$

$$= -2\left(\mathbf{a}_i' \sum_{l \neq i} \sum_{k=1}^{m} X_k' \mathbf{a}_l r_{lik} + \mathbf{a}_i' \sum_{j \neq i} \sum_{k=1}^{m} X_k \mathbf{a}_j r_{ijk} + \sum_{k=1}^{m} \mathbf{a}_i' X_k' \mathbf{a}_i r_{iik} \right) + c_1$$

$$= -2\left(\mathbf{a}_i' \sum_{j \neq i} \sum_{k=1}^{m} (X_k' \mathbf{a}_j r_{jik} + X_k \mathbf{a}_j r_{ijk}) + \sum_{k=1}^{m} \mathbf{a}_i' X_k' \mathbf{a}_i r_{iik} \right) + c_1, \tag{16}$$

where c_1 is a constant with respect to $\mathbf{a}_i$. Obviously, $\mathbf{a}_i' X_k' \mathbf{a}_i = \mathbf{a}_i' X_k \mathbf{a}_i = \mathbf{a}_i' \{1/2(X_k + X_k')\} \mathbf{a}_i$. Hence

$$g_1(\mathbf{a}_i) = -2\, \mathbf{a}_i' \sum_{j \neq i} \sum_{k=1}^{m} (X_k' \mathbf{a}_j r_{jik} + X_k \mathbf{a}_j r_{ijk}) + \mathbf{a}_i' \left(-\sum_{k=1}^{m} r_{iik}(X_k + X_k') \right) \mathbf{a}_i + c_1. \tag{17}$$

From (17) it is clear that $g_1(\mathbf{a}_i)$ can be minimized in essentially the same way as $g(\mathbf{a}_i)$, see (10).

Alternating the procedures for minimizing f_1 over $R_1, \ldots, R_m$ while A is fixed and minimizing f_1 over each of the columns of A while the other columns of A and $R_1, \ldots, R_m$ are fixed yields an ALS algorithm for fitting the three-way DEDICOM model in the least squares sense. Because the ALS algorithm decreases f_1 monotonically and f_1 is bounded from below, the algorithm must converge.

The IDIOSCAL model differs from the three-way DEDICOM model mentioned above in two respects. Firstly, IDIOSCAL is a model for a number of symmetric data matrices, instead of asymmetric matrices. Secondly, the matrices R_k in the IDIOSCAL model are not arbitrary but are required to be positive semi-definite. The first difference does not affect the applicability of the three-way DEDICOM algorithm, because the algorithm is suitable for any set of square data matrices X_k. The second difference calls for a slight modification of the three-way DEDICOM algorithm in order to make it suitable for fitting the IDIOSCAL model. The steps for updating the matrices R_k, $k = 1, \ldots, m$, should be modified to the effect that the best positive semi-definite matrix R_k is found. This comes down to minimizing the function

$$f_2(R_k) = \|A' X_k A - R_k\|^2 \tag{18}$$

over positive semi-definite matrices R_k. When X_k is itself positive semi-definite the solution for R_k is the same as in three-way DEDICOM, that is, $R_k = A' X_k A$, because when X_k is positive semi-definite, the product $A' X_k A$ is positive semi-definite as well. However, if X_k is not positive semi-definite then a different choice for R_k should be made. Let q be the number of positive eigenvalues of $A'X_kA$, let D_q be the diagonal matrix containing these positive eigenvalues, and let K_q be the column-wise orthonormal matrix containing the eigenvectors associated with these positive eigenvalues. Then, in order to minimize $f_2(R_k)$ matrix R_k should be chosen as $R_k = K_qD_qK_q'$ (Keller, 1962). These steps combined yield an ALS algorithm for fitting the IDIOSCAL model in the least squares sense.

Discussion

In the present paper algorithms have been described for finding a set of coordinates on a limited number of dimensions in order to represent a set of objects for which measures of relations among the objects are given. These relations are not necessarily symmetric. The algorithms that had been developed for these purposes have been based on the assumption that when two sets of object coordinates are estimated independently, these will be equal upon convergence. However, it is by no means guaranteed that this assumption holds. The algorithms described in the present paper do not require this assumption to be made. This advantage is due to the fact that these algorithms are based on column-wise updating of the object coordinate matrices in the ALS algorithms described above. This feature turns out to be very useful for fitting the DEDICOM model, the three-way DEDICOM model (3), and the IDIOSCAL model. The column-wise updating procedure used here is in no way limited to solving the specific problems mentioned above. It might successfully be used as a step in ALS algorithms for a wide variety of least squares minimization problems.

References

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

Carroll, J. D., & Chang, J. J. (1972, March). *IDIOSCAL (Individual Differences In Orientation SCALing): A generalization of INDSCAL allowing IDIOsyncratic reference systems as well as an analytic approximation to INDSCAL.* Paper presented at the meeting of the Psychometric Society, Princeton, NJ.

Carroll, J. D., & Wish, M. (1974). Models and methods for three-way multidimensional scaling. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. II, pp. 57–105). San Francisco: Freeman & Company.

De Leeuw, J., & Pruzansky, S. (1978). A new computational method to fit the weighted euclidean distance model. *Psychometrika, 43*, 479–490.

Harshman, R. A. (1978, August). *Models for analysis of asymmetrical relationships among N objects or stimuli.* Paper presented at the First Joint Meeting of the Psychometric Society and the Society of Mathematical Psychology, Hamilton, Ontario.

Harshman, R. A., Green, P. E., Wind, Y., & Lundy, M. E. (1982). A model for the analysis of asymmetric data in marketing research. *Marketing Science, 1*, 205–242.

Harshman, R. A., & Kiers, H. A. L. (1987, June). *Algorithms for DEDICOM analysis of asymmetric data.* Paper presented at the European Meeting of the Psychometric Society, Enschede.

Harshman, R. A., & Lundy, M. E. (1984). The PARAFAC model for three-way factor analysis and multidimensional scaling. In H. G. Law, C. W. Snyder Jr., J. A. Hattie, & R. P. McDonald (Eds.), *Research methods for multimode data analysis* (pp. 122–215). New York: Praeger.

Keller, J. B. (1962). Factorization of matrices by least squares. *Biometrika, 49*, 239–242.

Kroonenberg, P. M., & de Leeuw, J. (1980). Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika, 45*, 69–97.

ten Berge, J. M. F., & Nevels, K. (1977). A general solution to Mosier's oblique procrustes problem, *Psychometrika, 42*, 593–599.

Manuscript received 4/24/87
Final version received 5/26/88