

## LEAST-SQUARES APPROXIMATION OF AN IMPROPER CORRELATION MATRIX BY A PROPER ONE

DIRK L. KNOL

UNIVERSITY OF TWENTE

JOS M. F. TEN BERGE

UNIVERSITY OF GRONINGEN

An algorithm is presented for the best least-squares fitting correlation matrix approximating a given missing value or improper correlation matrix. The proposed algorithm is based upon a solution for Mosier's oblique Procrustes rotation problem offered by ten Berge and Nevels. A necessary and sufficient condition is given for a solution to yield the unique global minimum of the least-squares function. Empirical verification of the condition indicates that the occurrence of non-optimal solutions with the proposed algorithm is very unlikely. A possible drawback of the optimal solution is that it is a singular matrix of necessity. In cases where singularity is undesirable, one may impose the additional nonsingularity constraint that the smallest eigenvalue of the solution be  $\delta$ , where  $\delta$  is an arbitrary small positive constant. Finally, it may be desirable to weight the squared errors of estimation differentially. A generalized solution is derived which satisfies the additional nonsingularity constraint and also allows for weighting. The generalized solution can readily be obtained from the standard "unweighted singular" solution by transforming the observed improper correlation matrix in a suitable way.

Key words: missing value correlation, tetrachoric correlation, indefinite correlation matrix, constrained least-squares approximation.

When product-moment correlations of a set of  $n$  variables are computed by any of the missing value correlation methods described by Frane (1978), then it may happen that the resulting missing value correlation matrix is indefinite, and hence improper. This can be a serious problem in various multivariate data analysis techniques, for example, in regression and factor analysis.

One possible approach to this problem consists of avoiding an (indefinite) improper correlation matrix entirely by estimating the missing data themselves. Missing data can be estimated by maximum likelihood estimation from incomplete data (Beale & Little, 1975; Dempster, Laird & Rubin, 1977; Orchard & Woodbury, 1972) and by pragmatic procedures (Frane, 1976, 1978; Gleason & Staelin, 1975; Timm, 1970).

Another possible approach to the problem is to render the improper correlation matrix nonnegative definite by some smoothing procedure (Devlin, Gnanadesikan & Kettenring, 1975, p. 543; Dong, 1985; Frane, 1978).

The purpose of the present paper is to offer a least-squares smoothing procedure. That is, one may seek the best fitting (in the sense of least-squares) symmetric, unit-diagonal, nonnegative definite matrix  $G$  to the given improper correlation matrix  $R$ . Specifically, the function

$$e(G) \equiv \frac{1}{2} \text{tr} (G - R)^2 \quad (1)$$

Requests for reprints should be sent to Dirk Knol, University of Twente, Department of Education, PO Box 217, 7500 AE Enschede, THE NETHERLANDS or to Jos ten Berge, University of Groningen, Vakgroep Psychologie, Grote Markt 31-32, 9712 HV Groningen, THE NETHERLANDS.

can be minimized subject to the constraints  $G = G'$ ,  $\text{Diag}(G) = I_n$  and  $G \geq 0$ . For convenience we write  $\underline{Y} \geq 0$  and  $Y > 0$  to denote that a symmetric matrix  $Y$  is nonnegative definite and positive definite, respectively.

The minimization problem (1) can be generalized in three ways. Firstly, the problem can be applied to any improper correlation matrix, for example, an indefinite tetrachoric correlation matrix or a correlation matrix obtained by element-wise robust estimation (Devlin, Gnanadesikan, & Kettenring, 1975, 1981; Gnanadesikan & Kettenring, 1972). Secondly, the problem can be generalized to handle indefinite matrices with fixed diagonal elements not necessary equal to one. For example, the scope of the problem can be extended to missing value covariance matrices with known variances or to product-moment correlation matrices with known communalities. Thirdly, it is possible to exclude those product-moment correlations or covariances which are computed between complete variables (no missing values) from the minimization procedure. That is, the excluded elements of  $R$  can be held constant in (1). Without loss of generality these elements can be collected in the  $n_1 \times n_1$  ( $0 \leq n_1 < n$ ) submatrix  $\underline{R}_{11} \geq 0$  of  $R$ , where  $R$  is partitioned as

$$R = \left[ \begin{array}{c|c} R_{11} & R_{12} \\ \hline R_{21} & R_{22} \end{array} \right].$$

In order to incorporate these three generalizations, we shall address the generalized problem of minimizing (1) subject to the constraints

$$G = G', \quad (2)$$

$$G \geq 0, \quad (3)$$

$$G_{11} = R_{11} \geq 0 \quad (4)$$

and

$$\text{Diag}(G_{22}) = \text{Diag}(R_{22}) \geq 0, \quad (5)$$

where  $G$  is partitioned as

$$G = \left[ \begin{array}{c|c} G_{11} & G_{12} \\ \hline G_{21} & G_{22} \end{array} \right],$$

and  $G_{11}$  is of order  $n_1 \times n_1$ . Note that the constraints (4) and (5) for the problems with  $n_1 = 0$  and  $n_1 = 1$  are equivalent. In the next section a computational solution will be offered for the generalized problem of minimizing (1) subject to the constraints (2) through (5).

#### An Algorithm

The constraints  $G = G'$  (2) and  $G \geq 0$  (3) can equivalently be expressed by the constraint

$$G = AA', \quad (6)$$

for some  $n \times m$  ( $\underline{n}_1 \leq \underline{m} \leq n$ ) matrix  $A$ . Consider the partitioning

$$A = \left[ \begin{array}{c} A_1 \\ A_2 \end{array} \right] = \left[ \begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right],$$

where  $A_1$  is of order  $n_1 \times m$ ,  $A_{11}$  is of order  $n_1 \times n_1$ , and  $A_1$  is fixed in advance as

$$A_1 = [R_{11}^{1/2} \mid 0]. \quad (7)$$

This choice of  $A_1$  satisfies the constraint  $G_{11} = R_{11}$  (4) and can be adopted without loss of generality, because every matrix  $A$  satisfying (6) is determined up to an orthogonal rotation. Upon substitution of (6) and (7) for  $G$  in (1), the problem of minimizing (1), subject to the constraints (2) through (5), can be reduced to the problem of minimizing the function

$$f(A_2) \equiv \frac{1}{2} \text{tr} (A_2 A_2' - R_{22})^2 + \text{tr} (A_1 A_2' - R_{12})(A_1 A_2' - R_{12}), \quad (8)$$

subject to the constraint  $\text{Diag} (A_2 A_2') = \text{Diag} (R_{22})$ .

In order to simplify the notation, let for any positive integer  $\ell$  the index set  $\mathbb{N}_\ell^2$  be defined by the Cartesian product

$$\mathbb{N}_\ell^2 \equiv \{1, \dots, \ell\} \times \{1, \dots, \ell\},$$

and let  $\tau$  be the symmetric subset of  $\mathbb{N}_n^2$  defined by

$$\tau \equiv \{(i, j): i \neq j \ \& \ (i, j) \in \mathbb{N}_n^2 - \mathbb{N}_{n_1}^2\}.$$

Then the minimization problem (8) can be written as minimizing

$$f(A_2) = \frac{1}{2} \sum_{(i, j) \in \tau} (a'_i a_j - r_{ij})^2, \quad (9)$$

subject to the constraints  $a'_k a_k = r_{kk}$  ( $k = n_1 + 1, \dots, n$ ), where  $R \equiv [r_{ij}]$  and  $a'_i$  is row  $i$  ( $i = 1, \dots, n$ ) of  $A$ . For each  $k$  ( $k = n_1 + 1, \dots, n$ ), (9) can be written as

$$\begin{aligned} f(A_2) &= \frac{1}{2} \sum_{(i, k) \in \tau} (a'_i a_k - r_{ik})^2 + \frac{1}{2} \sum_{(k, j) \in \tau} (a'_k a_j - r_{kj})^2 + \frac{1}{2} \sum_{\substack{(i, j) \in \tau \\ i, j \neq k}} (a'_i a_j - r_{ij})^2 \\ &= \sum_{(i, k) \in \tau} (a'_i a_k - r_{ik})^2 + L_k \\ &= \sum_{i \neq k} (a'_i a_k - r_{ik})^2 + L_k \\ &= (A_k^{(0)} a_k - r_k^{(0)})'(A_k^{(0)} a_k - r_k^{(0)}) + L_k \\ &\equiv f_k(a_k) + L_k, \end{aligned} \quad (10)$$

where  $L_k$  is a constant with respect to  $a_k$ ,  $A_k^{(0)}$  is the matrix  $A$  with row  $k$  replaced by zeroes, and  $r_k^{(0)}$  is column  $k$  of  $[R - \text{Diag} (R)]$ .

In the context of Mosier's (1939) oblique Procrustes problem, ten Berge and Nevels (1977) have given a solution for the global minimum of  $f_k(a_k)$  subject to the constraint  $a'_k a_k = 1$ . With some minor adjustments, their solution can be generalized to minimize  $f_k(a_k)$  subject to any arbitrary constraint  $a'_k a_k = r_{kk} \geq 0$ . After taking a suitable initial choice for  $A_2$ , and row-wise minimization of (10) for  $k = n_1 + 1, \dots, n$  with the adjusted ten Berge and Nevels solution, an algorithm for solving (8) is obtained. For each  $k$  ( $k = n_1 + 1, \dots, n$ ),  $f(A_2)$  decreases with the row-wise minimization, affecting only elements of row  $k$  and column  $k$  of  $AA'$ . The  $n_2 \equiv n - n_1$  minimization steps can be repeated until no significant decrease of  $f(A_2)$  between two succeeding iteration cycles occurs. Because  $f(A_2)$  decreases monotonically and  $f(A_2)$  is bounded below, convergence of the algorithm is guaranteed. In the next section we shall describe a necessary and sufficient condition for a global minimum of  $f(A_2)$ .

### A Necessary and Sufficient Condition for a Global Minimum

After minimizing  $f_k(a_k)$  with the adjusted ten Berge and Nevels algorithm, there exists a Lagrange multiplier  $\theta_k$  such that

$$A_k^{(0)'} A_k^{(0)} a_k - \theta_k a_k = A_k^{(0)'} r_k^{(0)} \quad (11)$$

(Mulaik, 1972, p. 505). The Lagrange multiplier  $\theta_k$  can be evaluated directly from (11), (12) and (13) in ten Berge and Nevels (1977, p. 595) for their Cases 1, 2 and 3 respectively. Rewriting (11) yields

$$(A_k^{(0)'} A_k^{(0)} + a_k a_k') a_k - (A_k^{(0)'} r_k^{(0)} + a_k r_{kk}) - \theta_k a_k = 0,$$

and hence

$$A' A a_k - A' r_k - \theta_k a_k = 0, \quad (12)$$

where  $r_k$  is column  $k$  of  $R$ . It should be noted that during the iteration process, (12) holds for the index  $k$  only immediately after the minimization of row  $k - n_1$  of  $A_2$ . However, after convergence of the proposed algorithm, (12) holds simultaneously for all  $k$  ( $k = n_1 + 1, \dots, n$ ). Denote for convenience a solution of the proposed algorithm by  $A$ . Then the  $n_2$  equations (12) can be collected in the matrix equation

$$A' A A_2' - A' R_2' - A_2' \Theta_{22} = 0, \quad (13)$$

where  $R_2 \equiv [R_{21} | R_{22}]$  and  $\Theta_{22} \equiv \text{Diag}(\theta_{n_1+1}, \dots, \theta_n)$ .

It should be noted that the first-order necessary condition (13) for a minimum of (8) has been obtained from standard partial differentiation of a constrained function (Luenberger, 1984, chap. 10). Additional results can be obtained from a reformulation of the problem in terms of a semi-infinite convex program (Shapiro, 1985). Details of this have been given by Knol and ten Berge (1987). The most important results can be summarized in two lemmas.

*Lemma 1.* If  $R_{11} > 0$ , then a necessary and sufficient condition for a solution  $G^* = AA'$  to yield the global minimum of (1) subject to the constraints (2) through (5) is that

$$B_{22} \equiv A_2 A_2' - R_{22} - \underline{\Theta}_{22} \geq 0.$$

*Proof.* See Knol and ten Berge (1987).

It should be noted that, after convergence of the proposed algorithm,  $\Theta_{22}$  can be evaluated hence the condition  $B_{22} \geq 0$  can be verified. Moreover, Knol and ten Berge (1987) proved that the global minimum is unique.

In the case of singular  $R_{11}$ , Alexander Shapiro (personal communication, August 11, 1986) has shown that the problem of minimizing (1) subject to the constraints (2) through (5) can be transformed to a problem of (lower) dimensionality [ $\text{rank}(R_{11}) + n_2$ ], with a (transformed) fixed submatrix  $R_{11}^* > 0$ . For details see Knol and ten Berge (1987).

In order to verify the necessary and sufficient condition  $B_{22} \geq 0$  for a solution  $G^* = AA'$  to yield the (unique) global minimum of  $e(G)$  subject to the constraints (2) through (5), a computer program yielding the solution of the minimization problem with the proposed algorithm and evaluating the smallest eigenvalue of  $B_{22}$  has been implemented. The computer program was run on 100 symmetric unit-diagonal indefinite matrices, where  $n$  ranged from 5 to 25,  $n_1$  ranged from 0 to  $\min(10, n - 2)$  and the column order  $m$  of  $A$  was set equal to  $n$ . With changes in each (free) element of  $G$  between

TABLE 1

De Leeuw's target matrix  $R$  of polychoric correlations with the fifth and sixth variable interchanged

		R					
var		1	2	3	4	5	6
1		1.000					
2		.477	1.000				
3		.644	.516	1.000			
4		.478	.233	.599	1.000		
5		.651	.682	.581	.741	1.000	
6		.826	.750	.742	.800	.798	1.000

two succeeding iteration cycles less than  $10^{-4}$  as convergence criterion, the algorithm never took more than 10 iteration cycles until convergence. Computation time never exceeded 1 minute CPU time on a VAX8650 computer. In all cases, the obtained solution satisfied the condition  $B_{22} \geq 0$  within accuracy limits. From these results, it can be concluded that the proposed algorithm tends to produce the (unique) globally optimal solution.

In the following lemma, another important property of the solution is stated.

*Lemma 2.* The rank of  $G^*$  equals  $n$  if and only if  $R > 0$ .

*Proof.* See Knol and ten Berge (1987).

#### A Numerical Example

As an illustration and for reasons of possible checks, an indefinite  $6 \times 6$  matrix  $R$  of polychoric correlations (smallest eigenvalue  $-.063$ ) published by de Leeuw (1983, p. 121) has been analyzed with various values of  $n_1$  ( $R_{11} > 0$  for  $n_1 \leq 4$ ). In order to have  $R_{11} > 0$  for  $n_1 = 5$  too, the fifth and sixth variable have been interchanged. The matrix  $R$  is given in Table 1. Table 2 gives the residual matrices ( $G^* - R$ ) for various values of  $n_1$ , together with the values of  $e(G^*)$ . Because the constraints (2) through (5) for the problems with  $n_1 = 0$  and  $n_1 = 1$  are equivalent, the solutions are equal. In all cases, the solution satisfies the condition  $B_{22} \geq 0$  within accuracy limits. It can be verified that the value of  $e(G^*)$  increases as  $n_1$  increases, as is to be expected.

#### A Generalized Solution

It is clear from Lemma 2 that the best fitting proper correlation matrix  $G^*$  is singular. Also, it is clear from (1) that  $e(G)$  weights the errors of estimation equally for all variables. In the present section a generalization of the previous solution is derived by introducing the nonsingularity constraint that the smallest eigenvalue of the approximating symmetric matrix  $P$  be at least  $\delta$ , where  $\delta$  is an arbitrarily (small) positive scalar. For convenience, this will be denoted as  $P \geq \delta I$ . Additionally, the elements of the matrix

TABLE 2

The values of  $e(G^*)$  and the lower-triangular parts of the residual matrices  $(G^* - R)$  using De Leeuw's target matrix, for various values of  $n_1$  (structural zeroes omitted)

$n_1$	$e(G^*)$	var	$(G^* - R)$				
			1	2	3	4	5
0,1	.002760	2	.011				
		3	-.001	-.001			
		4	.012	.017	-.002		
		5	-.006	-.009	.001	-.010	
		6	-.018	-.025	.002	-.029	.014
2	.002884	3	-.001	-.002			
		4	.013	.018	-.002		
		5	-.007	-.009	.001	-.010	
		6	-.019	-.026	.003	-.030	.015
3	.002888	4	.013	.018	-.002		
		5	-.007	-.009	.001	-.011	
		6	-.019	-.026	.003	-.030	.015
4	.003515	5	-.008	-.012	.002	-.013	
		6	-.022	-.031	.004	-.036	.018
5	.004062	6	-.024	-.035	.006	-.040	.025

$P \equiv [p_{ij}]$  may be weighted differentially per variable with positive weights  $w_k$  ( $k = 1, \dots, n$ ). Let the matrix  $P$  be partitioned as

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix},$$

where  $P_{11}$  is of order  $n_1 \times n_1$  and  $P_{22}$  is of order  $n_2 \times n_2$ . Throughout this section, all symmetric matrices of order  $n$  will be partitioned similarly. Then, with  $D_w \equiv [\text{Diag}(w_1, \dots, w_n)]^{1/2} > 0$ , the generalized problem can be stated formally as minimizing the function

$$\begin{aligned} h(P) &\equiv \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_i w_j (p_{ij} - r_{ij})^2 \\ &= \frac{1}{2} \text{tr} [D_w(P - R)D_w]^2, \end{aligned} \quad (14)$$

subject to the constraints

$$P = P', \quad (15)$$

$$P \geq \delta I, \quad (16)$$

$$P_{11} = R_{11} \quad (17)$$

and

$$\text{Diag } (P_{22}) = \text{Diag } (R_{22}). \quad (18)$$

In order to apply the algorithm derived above to the generalized minimization problem (14) subject to the constraints (15) through (18), the problem will be rewritten using the identity

$$D_w(P - R)D_w = D_w(P - \delta I)D_w - D_w(R - \delta I)D_w. \quad (19)$$

Firstly, with (19) the function (14) can be written as

$$h(P) = \frac{1}{2} \text{tr } [D_w(P - \delta I)D_w - D_w(R - \delta I)D_w]^2. \quad (20)$$

Secondly, the constraint (15) is equivalent to

$$D_w(P - \delta I)D_w = [D_w(P - \delta I)D_w]'. \quad (21)$$

Thirdly, the constraint (16) is equivalent to

$$(P - \delta I) \geq 0,$$

and hence, since  $D_w > 0$ , equivalent to

$$D_w(P - \delta I)D_w \geq 0. \quad (22)$$

Fourthly, it is clear that the constraints (17) and (18) can be written as

$$[D_w(P - \delta I)D_w]_{11} = [D_w(R - \delta I)D_w]_{11} \quad (23)$$

and

$$\text{Diag } [D_w(P - \delta I)D_w]_{22} = \text{Diag } [D_w(R - \delta I)D_w]_{22}, \quad (24)$$

respectively.

If we define  $G \equiv D_w(P - \delta I)D_w$  and  $e^+(G) \equiv h(P)$ , then the problem of minimizing (20) subject to the constraints (21) through (24) is equivalent to the problem of minimizing the function

$$e^+(G) = \frac{1}{2} \text{tr } [G - D_w(R - \delta I)D_w]^2 \quad (25)$$

subject to the constraints

$$G = G', \quad (26)$$

$$G \geq 0, \quad (27)$$

$$G_{11} = [D_w(R - \delta I)D_w]_{11}, \quad (28)$$

and

$$\text{Diag } (G_{22}) = \text{Diag } [D_w(R - \delta I)D_w]_{22}. \quad (29)$$

Hence we have shown that the problem of minimizing (14) subject to the constraints (15) through (18) are equivalent to the problem of minimizing (25) subject to the constraints (26) through (29). The latter problem can be solved directly by the algorithm derived above, where the target matrix is now  $D_w(R - \delta I)D_w$  instead of  $R$ . If we denote the solution by  $G^*$ , then the solution to problem (14) subject to the constraints (15) through (18) is given by  $P^* = (D_w^{-1}G^*D_w^{-1} + \delta I)$ , with smallest eigenvalue  $\delta$ .

It should be noted that by the nature of the algorithm, restrictions on the choice of the smallest eigenvalue  $\delta$  of the desired approximating matrix  $P$  have to be observed. In order to make the problem solvable the conditions  $(R_{11} - \delta I_{n_1}) \geq 0$  (see (23)) and  $\text{Diag } (R_{22} - \delta I_{n_2}) \geq 0$  (see (24)) have to be satisfied. In most applications,  $\delta$  will be chosen as

TABLE 3

The values of the unweighted  $h(P^*)$  with  $\delta = .05$  and the lower-triangular parts of the residual matrices  $(P^* - R)$  using De Leeuw's target matrix, for various values of  $n_1$  (structural zeroes omitted)

$n_1$	$h(P^*)$	var	$(P^* - R)$				
			1	2	3	4	5
0,1	.009065	2	.019				
		3	-.001	-.002			
		4	.022	.030	-.002		
		5	-.010	-.014	.001	-.016	
		6	-.034	-.046	.003	-.054	.024
2	.009462	3	-.002	-.002			
		4	.024	.032	-.002		
		5	-.011	-.014	.001	-.016	
		6	-.036	-.048	.004	-.055	.025
3	.009469	4	.024	.032	-.003		
		5	-.011	-.014	.001	-.016	
		6	-.036	-.048	.004	-.055	.025
4	.011439	5	-.014	-.019	.003	-.022	
		6	-.042	-.057	.008	-.066	.031
5	.013502	6	-.041	-.062	.015	-.072	.052

the smallest positive scalar that avoids ill-conditionedness of the matrix  $P^*$  with respect to inversion.

Clearly, if we set  $D_w = I$  then we obtain an unweighted best-fitting solution with  $\delta$  as smallest eigenvalue. If we set  $\delta = 0$  then we obtain a weighted singular solution. If both  $D_w = I$  and  $\delta = 0$  then the previously derived solution is recovered.

For the purpose of illustration, the unweighted solution with  $\delta = .05$  obtained for the data of Table 1 is given in Table 3. It can be seen that the residual elements of Table 3 closely resemble those of Table 2. In fact all signs are equal. For each value of  $n_1$  we have  $h(P^*) > e(G^*)$  as is to be expected because  $h$  is more heavily constrained. Specifically,  $h(P^*)$  seems to be about three times as high as  $e(G^*)$ , in this example. Nevertheless, even these values of  $h(P^*)$  do not seem to be excessively high, for  $\delta = .05$ . If smaller residuals are needed, then a smaller value of  $\delta$  must be chosen.

#### Discussion

Above, a monotonely convergent algorithm has been constructed for the best least-squares proper approximation of an improper correlation or covariance matrix, preserv-



ing the diagonal elements. Also, a verifiable necessary and sufficient condition for a solution to yield the (unique) global minimum of the least-squares function has been given. Practical experience indicates that this condition tends to be satisfied. This makes it possible not only to detect but, in fact, to avoid local minima.

In addition, the method is highly flexible, for three reasons: It can handle fixed covariances or correlations for specific variables, it can handle weighted loss functions, and it can be adjusted to satisfy a nonsingularity constraint, at the cost of some loss of fit.

In view of these properties, it seems that an attractive alternative to existing smoothing procedures for improper correlation and covariance matrices has been obtained.

#### References

- Beale, E. M. L., & Little, R. J. A. (1975). Missing values in multivariate analysis. *Journal of the Royal Statistical Society, Series B*, 37, 129–145.
- de Leeuw, J. (1983). Models and methods for the analysis of correlation coefficients. *Journal of Econometrics*, 22, 113–137.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B*, 39, 1–38.
- Devlin, S. J., Gnanadesikan, R., & Kettenring, J. R. (1975). Robust estimation and outlier detection with correlation coefficients. *Biometrika*, 62, 531–545.
- Devlin, S. J., Gnanadesikan, R., & Kettenring, J. R. (1981). Robust estimation of dispersion matrices and principal components. *Journal of the American Statistical Association*, 76, 354–362.
- Dong, H. K. (1985). Non-Gramian and singular matrices in maximum likelihood factor analysis. *Applied Psychological Measurement*, 9, 363–366.
- Frane, J. W. (1976). Some simple procedures for handling missing data in multivariate analysis. *Psychometrika*, 41, 409–415.
- Frane, J. W. (1978). Missing data and BMDP: Some pragmatic approaches. *Proceedings of the Statistical Computing Section* (pp. 27–33). Washington, DC: American Statistical Association.
- Gleason, T. C., & Staelin, R. (1975). A proposal for handling missing data. *Psychometrika*, 40, 229–252.
- Gnanadesikan, R., & Kettenring, J. R. (1972). Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics*, 28, 81–124.
- Knol, D. L., & ten Berge, J. M. F. (1987). *Least-squares approximation of an improper by a proper correlation matrix using a semi-infinite convex program* (Research Report 87-7). Enschede, The Netherlands: University of Twente, Department of Education.
- Luenberger, D. G. (1984). *Introduction to linear and nonlinear programming* (2nd ed.). Reading, MA: Addison-Wesley.
- Mosier, C. I. (1939). Determining a simple structure when loadings for certain tests are known. *Psychometrika*, 4, 149–162.
- Mulaik, S. A. (1972). *The foundations of factor analysis*. New York: McGraw-Hill.
- Orchard, T., & Woodbury, M. A. (1972). A missing information principle: Theory and applications. *Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1), 6, 697–715.
- Shapiro, A. (1985). Extremal problems on the set of nonnegative definite matrices. *Linear Algebra and its Applications*, 67, 7–18.
- ten Berge, J. M. F., & Nevels, K. (1977). A general solution to Mosier's oblique Procrustes problem. *Psychometrika*, 42, 593–600.
- Timm, N. H. (1970). The estimation of variance-covariance and correlation matrices from incomplete data. *Psychometrika*, 35, 417–437.

Manuscript received 6/5/86

Final version received 10/14/87