DOCUMENT RESUME

ED 310 118 TM 013 691

AUTHOR	Knol, Dirk L.; ten Berge, Jos M. F.
TITLE	Least-Squares Approximation of an Improper by a Proper Correlation Matrix Using a Semi-Infinite Convex Program. Research Report 87-7.
INSTITUTION	Twente Univ., Enschede (Netherlands). Dept. of Education.
PUB DATE	Oct 87
NOTE	33p.; Also cited as Project Psychometric Aspects of Item Banking No. 22.
AVAILABLE FROM	Faculty Library, Department of Education, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Algorithms; *Computer Software; Correlation; *Estimation (Mathematics); *Least Squares Statistics; Statistical Analysis
IDENTIFIERS	Convex Program; *Correlation Matrices; *Semi Infinite Programs; Tetrachoric Correlation

ABSTRACT

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 on a solution for C. I. Mosier's oblique Procrustes rotation problem offered by J. M. F. ten Berge and K. Nevels (1977). It is shown that the minimization problem belongs to a certain class of convex programs in optimization theory. A necessary and sufficient condition for a solution to yield the unique global minimum of the least-squares function is derived from a theorem by A. Shapiro (1985). A computer program was implemented to yield the solution of the minimization problem with the proposed algorithm. This empirical verification of the condition indicates that the occurrence of non-optimal solutions with the proposed algorithm is very unlikely. Two tables present values using J. de Leeuw's target matrix. (Author/SLD)

Least-squares Approximation of an Improper by a Proper Correlation Matrix Using a Semi-infinite Convex Program

Research Report 87-7

Dirk L. Knol
Jos M.F. ten Berge

Department of EDUCATION

Division of Educational Measurement and Data Analysis

University of Twente

Project Psychometric Aspects of Item Banking No.22

Colofon:
Typing: Dirk L. Knol
Cover design: Audiovisuele Sectie TOLAB Toegepaste Onderwijskunde
Printed by: Centrale Reproductie-afdeling

Least-squares Approximation of an Improper by a Proper Correlation Matrix Using a Semi-infinite Convex Program

Dirk L. Knol

Jos M.F. ten Berge

Least-squares Approximation of an Improper by a Proper Correlation Matrix Using a Semi-infinite Convex Program / Dirk L. Knol and Jos M.F. ten Berge - Enschede : University of Twente, Department of Education, October, 1987. - 27 p.

Abstract

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. It is shown that the minimization problem belongs to a certain class of convex programs in optimization theory. A necessary and sufficient condition for a solution to yield the unique global minimum of the least-squares function is derived from a theorem by Shapiro. Empirical verification of the condition indicates that the occurrence of non-optimal solutions with the proposed algorithm is very unlikely.

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

Least-squares Approximation of an Improper by a Proper Correlation Matrix Using a Semi-infinite Program

When product-moment correlations of a set of n variables are computed by any of the missing value correlation methods described by Frane (1978), 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, e.g., 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 non-negative 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, non-negative definite matrix G to the given improper missing value correlation matrix R. Specifically, the function

(1) $e(G) = \frac{1}{2} \operatorname{tr} (G - R)^2$

can be minimized subject to the constraints $G = G'$, $\operatorname{Diag}(G) = I_n$ and $G \geq 0$. For convenience we write $Y \geq 0$ and $Y > 0$ to denote that a symmetric matrix Y is non-negative 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, e.g., 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 $R_{11} \geq 0$ 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 adress the generalized problem of minimizing (1) subject to the constraints

(2a) $G = G'$,

(2b) $G \geq 0$,

(2c) $G_{11} = R_{11} \geq 0$

and

(2d) $\text{Diag}\ (G_{22}) = \text{Diag}\ (R_{22}) \geq 0$.

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 (2c) and (2d) 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).

An algorithm

The constraints $G = G'$ (2a) and $G \geq C$ (2b) can equivalently be expressed by the constraint

(3) $\quad G = AA'$

for some $n \times m$ ($n_1 \leq m \leq n$) matrix A. Consider the partitioning

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

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

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

This choice of A_1 satisfies the constraint $G_{11} = R_{11}$ (2c) and can be adopted without loss of generality, because every matrix A satisfying (3) is determined up to an orthogonal rotation.

Upon substitution of (3) and (4) for G in (1), the problem of minimizing (1) subject to the constraints (2) can be reduced to the problem of minimizing the function

$$(5) \qquad f(A_2) = \tfrac{1}{4}\, \mathrm{tr}\, (A_2A_2' - R_{22})^2 + \mathrm{tr}\, (A_1A_2' - R_{12})'(A_1A_2' - R_{12})$$

subject to the constraint Diag (A_2A_2') = Diag (R_{22}).

In order to simplify the notation, let for any positive integer ℓ the index set N_ℓ^2 be defined by the Cartesian product

$$N_\ell^2 = \{1,\ldots,\ell\} \times \{1,\ldots,\ell\} \ ,$$

and let τ be the symmetric subset of N_n^2 defined by

$$\tau = \{(i,j):\ i \neq j \ \&\ (i,j) \in N_n^2 - N_{n_1}^2\} \ .$$

Then the minimization problem (5) can be written as minimizing

$$(6) \qquad f(A_2) = \tfrac{1}{4} \sum_{(i,j)\in\tau}\sum (a_i'a_j - r_{ij})^2$$

subject to the constraints $a_k'a_k = r_{kk}$ $(k = n_1+1,\ldots,n)$, where $R = [r_{ij}]$ and a_i' is row i $(i = 1,\ldots,n)$ of A. For each k $(k = n_1+1,\ldots,n)$, (6) can be written as

$$(7) \qquad f(A_2) = \tfrac{1}{2} \sum_{(i,k)\in T} (a_i'a_k - r_{ik})^2$$

$$+ \tfrac{1}{2} \sum_{(k,j)\in T} (a_k'a_j - r_{kj})^2$$

$$+ \tfrac{1}{2} \sum_{\substack{(i,j)\in T \\ i,j\neq k}} \sum (a_i'a_j - r_{ij})^2$$

$$= \sum_{(i,k)\in T} (a_i'a_k - r_{ik})^2 + L_k$$

$$= \sum_{i\neq k} (a_i'a_k - r_{ik})^2 + L_k$$

$$= (A_k^{(o)}a_k - r_k^{(o)})'(A_k^{(o)}a_k - r_k^{(o)}) + L_k$$

$$= f_k(a_k) + L_k ,$$

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

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 (7) for $k = n_1+1,\ldots,n$ with the adjusted Ten Berge and Nevels solution, an algorithm for solving (5) is obtained. For each k $(k = n_1+1,\ldots,n)$, $f(A_2)$ decreases with the row-wise minimization, affecting only elements of row k and column k of AA'. The $n_2 = n - n_1$ 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 θ_k such that

$$(8) \qquad A_k^{(o)\prime}A_k^{(o)}a_k - \theta_k a_k = A_k^{(o)\prime}r_k^{(o)}$$

(Mulaik, 1972, p. 505). The Lagrange multiplier θ_k can be evaluated directly from the equations (11), (12) and (13) in Ten Berge and Nevels (1977, p. 595) for their cases 1, 2 and 3 respectively. Rewriting (8) yields

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

and hence

(9) $$A'Aa_k - A'r_k - \theta_k a_k = 0 \; .$$

where r_k is column k of R. It should be noted that during the iteration process, (9) 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, (9) holds simultaneously for all k ($k = n_1+1,\ldots,n$). Denote for convenience a solution of the proposed algoritm by A. Then the n_2 equations (9) can be collected in the matrix equation

(10) $$A'AA_2' - A'R_2' - A_2'\Theta_{22} = 0 \; ,$$

where $R_2 = [R_{21}|R_{22}]$ and $\Theta_{22} = \text{Diag}\,(\theta_{n_1+1},\ldots,\theta_n)$. Transposing (10) and rewriting yields the first-order necessary conditions for a minimum of (5)

(11a) $$(A_2A_1' - R_{21})A_{11} + (A_2A_2' - R_{22} - \Theta_{22})A_{21} = 0$$

and

$$(11b)\quad (\Lambda_2\Lambda_2' - R_{22} - \Theta_{22})\Lambda_{22} = 0 \; .$$

It should be noted that the first-order necessary conditions (11) for a minimum of (5) have been obtained from standard partial differentiation of a constrained function (cf. 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). This will be pursued next.

Let $\Omega(\tau)$ denote the set of symmetric $n \times n$ matrices $X = [x_{ij}]$ such that $x_{ij} = 0$ whenever $(i,j) \in \tau$. Then the matrix G can be written as

$$(12)\quad G = C + X \; ,$$

where $X \in \Omega(\tau)$ and $C = [c_{ij}]$ such that $c_{ij} = 0$ whenever $(i,j) \in \tau$ and $c_{ij} = r_{ij}$ otherwise. In (12) G is decomposed as the sum of a matrix C containing the $(n_1)^2 + n_2$ known (fixed) elements of G, and a matrix X containing the unknown (free) elements of G. Inserting (12) in (1) leads to the restatement of the minimization problem

$$(13)\quad g(X) = e(G) = \tfrac{1}{2}\,\mathrm{tr}\,(C + X - R)^2$$

subject to the constraints $X \in \Omega(\tau)$ and $(C + X) \geq 0$.

Replacing the constraint $(C + X) \geq 0$ by the equivalent constraint

$$(14) \quad h(X,u) = u'(C + X)u \geq 0$$

for all $u \in \Psi = \{u \in \mathbb{R}^n: u'u = 1\}$ makes problem (13) a semi-infinite program.

Assuming that we have $R_{11} > 0$ it can be verified that the semi-infinite program defined by (13) and (14) has the following nine properties:

(P1) $\Omega(\tau)$ is convex.

(P2) $g(X)$ is convex.

(P3) $h(.,u)$ is concave for all $u \in \Psi$.

(P4) The Slater (1950) condition (cf. Stoer & Witzgall, 1970, p. 247) holds, i.e. there exists a matrix $X \in \Omega(\tau)$, viz., $X_0 = 0$, such that $h(X_0,u) > 0$ for all $u \in \Psi$.

(P5) Ψ is compact.

(P6) $g(X)$ is continuously differentiable.

(P7) $h(.,u)$ is continuously differentiable for all $u \in \Psi$.

(P8) $h(X,u)$ is continuous.

(P9) $\mathrm{grad}_X\, h(X,u)$ is continuous.

Properties (P1) through (P3) make the program a convex program and properties (P4) through (P9) are regularity conditions.

For semi-infinite programs satisfying the conditions (P1) through (P9), Theorem 2.2 of Shapiro (1985) is applicable, which states: A feasible $X^* \in \Omega(\tau)$, i.e. $(C + X^*) \geq 0$, is a solution of the minimization problem if and only if there exists an n x n matrix $B = [b_{ij}]$ satisfying

(i) $B = B'$,

(ii) $(C + X^*)B = 0$,

(iii) $\text{grad } g(X)|_{X=X^*} = P_\tau(B)$,

where $P_\tau(B)$ is the projection of B onto the space $\Omega(\tau)$ defined by

$$[P_\tau(B)]_{ij} = \begin{cases} b_{ij} & \text{whenever } (i,j) \in \tau , \\ 0 & \text{otherwise} . \end{cases}$$

(iv) $B \geq 0$.

In order to assess whether these necessary and sufficient conditions are satisfied after convergence of the proposed algorithm, we shall use the following lemma.

<u>Lemma 1</u>. For the matrix

(15) $B = W'B_{22}W$.

where $W = [-A_{21}A_{11}^{-1} | I_{n_2}]$ and $B_{22} = (A_2A_2' - R_{22} - \Theta_{22})$, the conditions (i) through (iii) are satisfied, and condition (iv) is equivalent to the condition $B_{22} \geq 0$.

Proof. Condition (i) is obviously satisfied.

To prove condition (ii), note that

$$(16) \qquad B_{22}WA = [\ 0 \mid B_{22}A_{22}\ ] \ .$$

Rewriting (11b) as $B_{22}A_{22} = 0$ and transposing (16) yields

$$A'W'B_{22} = 0$$

and hence

$$(17) \qquad AA'W'B_{22}W = 0 \ .$$

Substituting $(C + X^*) = AA'$ and (15) in (17) proves condition (ii).

To prove condition (iii), the matrix B is written out as

$$B = \left[\begin{array}{c|c} (B_{22}A_{21}A_{11}^{-1})'A_{21}A_{11}^{-1} & -(B_{22}A_{21}A_{11}^{-1})' \\ \hline -B_{22}A_{21}A_{11}^{-1} & B_{22} \end{array}\right] .$$

From (11a) it follows

(19) $$B_{22}A_{21}A_{11}^{-1} = -(A_2A_1' - R_{21}) .$$

Inserting (19) in (18) yields

$$B = \left[\begin{array}{c|c} -(A_1A_2' - R_{12})A_{21}A_{11}^{-1} & A_1A_2' - R_{12} \\ \hline A_2A_1' - R_{21} & A_2A_2' - R_{22} - \Theta_{22} \end{array}\right] .$$

which can be written as

(20) $$B = AA' - R - \Theta$$

$$= C + X^* - R - \Theta ,$$

where

$$\Theta = \left[\begin{array}{c|c} (A_1A_2' - R_{12})A_{21}A_{11}^{-1} & 0 \\ \hline 0 & \Theta_{22} \end{array}\right] .$$

From (20) it is easily shown that $P_T(B) = (C + X^* - R)$, which equals grad $g(X)|_{X=X^*}$. This proves condition (iii).

Regarding condition (iv), it is obvious from (15) that the condition $B \geq 0$ is equivalent to the condition $B_{22} \geq 0$. ▪

From our Lemma 1 and Shapiro's Theorem 2.2 it is obvious that $B_{22} \geq 0$ is a necessary and sufficient condition for a feasible $X^* \in \Omega(\tau)$ to be a solution of the minimization problem (13). It should be noted that, after convergence of the proposed algorithm, Θ_{22} can be evaluated hence the condition $B_{22} \geq 0$ can be verified. Moreover, when $B_{22} \geq 0$, it follows immediately from the strict convexity of $g(X)$ that X^* is the unique solution of the minimization problem (13), which means that if and only if $B_{22} \geq 0$ the unique global minimum of $e(G)$ subject to the constraints (2) has been attained for $G^* = (C + X^*)$.

In the derivation of the necessary and sufficient condition for a solution to yield the unique global minimum of $e(G)$, it has to be assumed that $R_{11} > 0$. In the case of singular R_{11} the Slater condition (P4) does not hold and ${A_{11}}^{-1}$ does not exist, hence it cannot be verified whether the obtained solution yields the unique global minimum of $e(G)$. However, for singular R_{11}, Alexander Shapiro (personal communication, August 11, 1986) has shown that the problem of minimizing (1) subject to the constraints (2) can be transformed to a problem of (lower) dimensionality [rank $(R_{11}) + n_2$], with a (transformed) fixed submatrix $R_{11}^* > 0$. For reasons of availability, we give the proof which is due to Shapiro.

Firstly, the function $e(G)$ can be written as

(21) $\quad e(G) = \frac{1}{2}\,\mathrm{tr}\,[P(G - R)P']^2 = \frac{1}{2}\,\mathrm{tr}\,(PGP' - PRP')^2$

for any orthogonal matrix P of order $n \times n$. Secondly, let us take P in the form

$$P = \left[\begin{array}{c|c} P_{11} & 0 \\ \hline 0 & I_{n_2} \end{array}\right],$$

such that

(22) $$P_{11}R_{11}P_{11}' = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & R_{11}^{*} \end{array}\right],$$

with $R_{11}^{*} > 0$. Then the constraints (2) become

(23a) $\quad PGP' = PG'P'$,

(23b) $$PGP' = \left[\begin{array}{c|c} P_{11}G_{11}P_{11}' & P_{11}G_{12} \\ \hline G_{21}P_{11}' & G_{22} \end{array}\right] \geq 0 \; .$$

(23c) $P_{11}G_{11}P_{11}' = P_{11}R_{11}P_{11}' \geq 0$

and

(23d) $\text{Diag}\ (G_{22}) = \text{Diag}\ (R_{22}) \geq 0$.

From (22) and (23c) it follows that the first $[n_1 - \text{rank}\ (R_{11})]$ diagonal elements of PGP' are zero. From this and (23b) it follows that the first $[n_1 - \text{rank}\ (R_{11})]$ rows and columns of PGP' are zeroes. Hence the problem of minimizing (21) subject to the constraints (23) is reduced to a problem of dimensionality $[\text{rank}\ (R_{11}) + n_2] < n$.

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), a computer program has been implemented yielding the solution of the minimization problem with the proposed algorithm and evaluating the smallest eigenvalue of B_{22}. 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 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. Suppose first that $R > 0$. Then, $G^* = R > 0$, and hence the rank of G^* equals n.

Conversely, let the rank of $G^* = (C + X^*)$ equal n. From condition (ii) it follows that $B = 0$, and from condition (iii) that grad $g(X)|_{X=X^*} = 0$. From this it follows that $P_T(B) = (C + X^* - R) = (G^* - R) = 0$ and hence $G^* = R > 0$. •

In practice it seems to be true that the rank of G^* is always less than or equal to the number p of positive eigenvalues of R. Since computation time heavily depends upon the column order m of A, it is advised to take $m = p$. A further reduction of computation time can be accomplished by setting a suitable initial value for A. A reasonable initial value $A^{(o)}$ can be based upon an eigen-decomposition of R

$$R = K\Delta K'\ .$$

where K is an orthogonal matrix of order n x n containing as columns the normalized eigenvectors of R and Δ is a diagonal matrix containing the n eigenvalues of R. Let Δ_p be the diagonal matrix of order p with diagonal elements the p positive eigenvalues of R, and let K_p be the matrix of order $n \times p$ with columns the p corresponding (normalized) eigenvectors. In the case $n_1 = 0$, it is advised to take as initial value for A

$$(24) \qquad A^{(o)} = [\mathrm{Diag}\ (R)]^{½}[\mathrm{Diag}\ (K_p \Delta_p K_p')]^{-½} K_p \Delta_p^{½} T$$

where T is an arbitrary orthogonal matrix of order $p \times p$. In the case $n_1 > 0$ one can take T such that the upper $p \times p$ submatrix of $A^{(o)}$ is in lower triangular form and replace the submatrix $A_1^{(o)}$ by (4). For the 100 least-squares problems used above, total computation time could be reduced by more than 50% using (24) as initial value, and the condition B_{22} was again satisfied in all cases after convergence of the algorithm.

A numerical example

As an illustration and for reasons of possible checks, an indefinite 6×6 matrix R of polychoric correlations (smallest eigenvalue $-.0626$) 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.

Insert Table 1 about here

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) 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.

Insert Table 2 about here

It can be verified that the value of $e(G^*)$ increases as n_1 increases, as is to be expected.

Discussion

A monotone convergent algorithm has been constructed for the best least-squares non-negative definite approximation of an improper correlation or covariance matrix, preserving the diagonal elements. Additionally a verifiable necessary and sufficient condition for a solution to yield the unique global minimum of the least-quares function has been derived.

Moreover, this condition tends to be satisfied in practice. Thus a possibly useful alternative to existing smoothing procedures has been found.

However, the solution G^* is singular except in the trivial case $R > 0$ (cf. Lemma 2) hence the inverse of G^* does not exist. When inversion of G^* is required for a particular subsequent multivariate analysis, one may impose the additional constraint to (2) that all eigenvalues of G are greater than or equal to an arbitrary positive constant δ. An algorithm for the latter optimization problem is in progress, but is beyond the scope of the present paper.

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.

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), 697–715.

Shapiro, A. (1985). Extremal problems on the set of nonnegative definite matrices. Linear Algebra and its Applications, 67, 7–18.

Slater, M.L. (1950, November). Lagrange multipliers revisited: A contribution to nonlinear programming. Cowles Commission Discussion Paper, Math. 403.

Stoer, J., & Witzgall, C. (1970). Convexity and optimization in finite dimensions I. Berlin: Springer.

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.

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	1.0000					
2	.4770	1.0000				
3	.6440	.5160	1.0000			
4	.4780	.2330	.5990	1.0000		
5	.6510	.6820	.5810	.7410	1.0000	
6	.8260	.7500	.7420	.8000	.7980	1.0000

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)

			$(G^* - R)$				
n_1	$e(G^*)$	var	1	2	3	4	5
0.1	.002760	2	.0108				
		3	−.0011	−.0015			
		4	.0125	.0173	−.0017		
		5	−.0063	−.0088	.0009	−.0101	
		6	−.0178	−.0248	.0025	−.0286	.0144
2	.002884	3	−.0012	−.0016			
		4	.0131	.0180	−.0019		
		5	−.0067	−.0092	.0009	−.0105	
		6	−.0189	−.0259	.0027	−.0296	.0151
3	.002888	4	.0132	.0181	−.0020		
		5	−.0067	−.0092	.0010	−.0105	
		6	−.0189	−.0259	.0028	−.0296	.0151
4	.003515	5	−.0083	−.0116	.0016	−.0133	
		6	−.0225	−.0312	.0043	−.0359	.0185
5	.004062	6	−.0243	−.0349	.0057	−.0403	.0245

Acknowledgement

The authors are obliged to Alexander Shapiro for his solution in the case of singular R_{11}. However, the responsibility remains entirely to the authors.

[redacted]tment of

[redacted]UCATION

A publication by
the Department of Education
of the University of Tw[redacted]e

[redacted]de
[redacted]ds