PSYCHOMETRIKA-VOL. 47, NO. 2.
JUNE, 1982

RANK-REDUCIBILITY OF A SYMMETRIC MATRIX AND SAMPLING THEORY OF MINIMUM TRACE FACTOR ANALYSIS

ALEXANDER SHAPIRO

BEN-GURION UNIVERSITY OF THE NEGEV

One of the intriguing questions of factor analysis is the extent to which one can reduce the rank of a symmetric matrix by only changing its diagonal entries. We show in this paper that the set of matrices, which can be reduced to rank r, has positive (Lebesgue) measure if and only if r is greater or equal to the Ledermann bound. In other words the Ledermann bound is shown to be *almost surely* the greatest lower bound to a reduced rank of the sample covariance matrix. Afterwards an asymptotic sampling theory of so-called minimum trace factor analysis (MTFA) is proposed. The theory is based on continuous and differential properties of functions involved in the MTFA. Convex analysis techniques are utilized to obtain conditions for differentiability of these functions.

Key words: reduced rank, reliability, sample estimates.

1. Introduction

Under the well known factor analysis model the $p \times p$ population covariance matrix Σ is decomposed as follows:

$$\Sigma = (\Sigma - \Psi) + \Psi \tag{1.1}$$

where $\Sigma - \Psi$ is Gramian and Ψ is diagonal. It is clear that without any additional assumptions there exist infinitely many possibilities of choosing Ψ. Two approaches to selection of Ψ will be discussed in this paper. The classical approach solves the problem by searching for Ψ that minimizes the rank of $\Sigma - \Psi$ [e.g., Harman, 1976]. On the other hand in the last decade several authors considered the problem of finding Ψ that minimizes the trace of $\Sigma - \Psi$. We shall refer to this approach as the minimum-trace factor analysis (MTFA). The concept of MTFA has been discussed by Ledermann [1939] who also considered its relation to minimum-rank factor analysis (MRFA). Bentler [1972] applied the MTFA approach to reliability theory. He introduced coefficient ρ that is a lower bound to reliability

$$\rho = 1 - \frac{1'\Psi^* 1}{1'\Sigma\, 1} \tag{1.2}$$

where Ψ^* is the MTFA solution and 1 is the vector of unit weights. However the MTFA (as well as the MRFA) could lead to negative entries of Ψ (Heywood case). Jackson and Agunwamba [1977], Woodhouse and Jackson [1977] and independently Bentler and Woodward [1980] considered the MTFA with further constraint on Ψ to be Gramian. This approach will be referred to as constrained minimum trace factor analysis (CMTFA). The CMTFA provides the greatest lower bound (g.l.b.) to reliability, which is defined as in (1.2) except that Ψ^* is the CMTFA solution.

Requests for reprints should be sent to A. Shapiro, Department of Statistics & O.R., University of South Africa, P.O. Box 392, PRETORIA 0001, South Africa.

0033-3123/82/0600-307/$00.75/0

In practice the population covariance matrix Σ is unknown. Therefore one substitutes for Σ a sample estimate S. Then two questions arise:

(i) To what extent can one reduce the rank of S by only changing its diagonal entries?
(ii) What is the (asymptotic) distribution of sample estimators which the MTFA involves?

The first question has a long history. Ledermann [1937] has suggested that rank can always be reduced to the bound $\phi(p)$, where

$$\phi(p) = \frac{2p + 1 - (8p + 1)^{1/2}}{2}. \tag{1.3}$$

Guttman [1958] gave an example wherein it is impossible to reduce rank below $p - 1$. Shapiro [in press] showed that the minimal reduced rank is $p - 1$ if by changing the sign of a suitable set of variables all the inter-item correlations can be made negative. This is the author's conjecture that loosing the condition for the reduced matrix to be Gramian the Ledermann bound can be achieved *almost surely* (with probability one). At least this is true for $p = 3$.

What we prove in the next section is that with probability one the rank of the sample covariance matrix S cannot be reduced below the Ledermann bound. In this sense $\phi(p)$ can be considered as *almost surely* a lower bound to a reduced rank of S. Then we apply some techniques and definitions of this section to investigate the differential properties of MTFA solutions as functions of a symmetric matrix S. This will provide us with the required base for deriving the asymptotic distribution theory of the MTFA.

The following notations and definitions will be used: R^p denotes the p-dimensional vector space of column vectors. The term by term product (Hadamard product) of matrices V and W will be denoted by V^*W. In particular if $\mathbf{e} = (e_1, \ldots, e_p)'$ and $\mathbf{f} = (f_1, \ldots, f_p)'$ are two column vectors then $\mathbf{e}^*\mathbf{f} = (e_1 f_1, \ldots, e_p f_p)'$ and we denote $\mathbf{e}^2 = \mathbf{e}^*\mathbf{e} = (e_1^2, \ldots, e_p^2)'$.

A property is said to hold *almost everywhere* (on the set of symmetric matrices) if it holds everywhere except for a set of (Lebesgue) measure zero. In statistical language the phrase "*almost everywhere* a nonnegative definite matrix ..." can be said "*almost surely* (with probability one) a sample covariance matrix ...".

For the diagonal matrices $H = \text{diag}(h_i)$, $Y = \text{diag}(y_i)$ we use small letters to denote the corresponding vectors $\mathbf{h} = (h_1, \ldots, h_p)'$, $\mathbf{y} = (y_1, \ldots, y_p)'$ etc. By $N(A')$ we denote the null space of a $p \times r$ matrix A, i.e., $N(A') = \{\mathbf{x} \in R^p : A'\mathbf{x} = 0\}$

2. *Rank-Reducibility of a Symmetric Matrix*

Let S be a $p \times p$ symmetric matrix. The following question will be considered in this section: To what extent can one reduce the rank of S-H by selecting an appropriate diagonal matrix H? It has long been suspected that for randomly chosen S the rank cannot be reduced too much. We shall show that the set of matrices for which the rank can be reduced below the Ledermann bound $\phi(p)$ is thin, or to be more specific has (Lebesgue) measure zero. Another question of interest is the stability of reduced rank. In other words if entries of S are slightly changed can the matrix H be adjusted to preserve the reduced rank? We shall discuss such stability at the end of this section.

Lemma 2.1

Let A be a $p \times r$ matrix and C be a $p \times p$ symmetric matrix. Then the matrix equation

$$AX' + XA' = C \tag{2.1}$$

where X is a $p \times r$ matrix of unknowns, has a solution if and only if

$$\mathbf{e}'C\mathbf{f} = 0 \quad \text{for all } \mathbf{e}, \mathbf{f} \in N(A') \tag{2.2}$$

Proof. The necessity follows from the equality $\mathbf{e}'AX'\mathbf{f} + \mathbf{e}'XA'\mathbf{f} = \mathbf{e}'C\mathbf{f}$ and $A'\mathbf{e} = A'\mathbf{f} = 0$.

To prove the sufficiency we first consider the case when A is of the following form:

$$A' = [I_r \,|\, 0] \tag{2.3}$$

where I_r is the $r \times r$ identity matrix and 0 is the $r \times (p - r)$ zero matrix. Let C be partitioned as follows:

$$C = \left[\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right] \begin{array}{l} r \\ p-r \end{array} \quad \begin{array}{cc} r & p-r \end{array} \tag{2.4}$$

We have by (2.3) that $\mathbf{e} \in N(A')$ if and only if the first r co-ordinates of $\mathbf{e}$ are zero. Then it follows from (2.2) that $C_{22} = 0$. From (2.3) we obtain that

$$AX' + XA' = \left[\begin{array}{c|c} X_1 + X_1' & X_2' \\ \hline X_2 & 0 \end{array}\right] \tag{2.5}$$

where $[X_1' \,|\, X_2'] = X'$. Therefore (2.1) is equivalent to

$$\begin{aligned} X_1 + X_1' &= C_{11} \\ X_2 &= C_{21} \end{aligned} \tag{2.6}$$

It can easily be verified that we solve (2.6) by putting $X_1 = 1/2\, C_{11}$ and $X_2 = C_{21}$.

In the general case we can assume without loss of generality that A has rank r. By some transformation TA, where T is a $p \times p$ nonsingular matrix, the general case can be reduced to the case above. □

We remark that if $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ is some vector basis of $N(A')$, then condition (2.2) can be replaced by

$$\mathbf{e}_i' C \mathbf{e}_j = 0 \qquad \text{for} \qquad i, j = 1, \ldots, k. \tag{2.7}$$

Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ be a set of p-dimensional column vectors. The $p \times m$ matrix with column vectors $\mathbf{e}_i$, $i = 1, \ldots, p$, will be denoted by $[\mathbf{e}_i]$ and the $p \times m(m + 1)/2$ matrix whose column vectors are $\mathbf{e}_i^* \mathbf{e}_j$, $i \leq j = 1, \ldots, m$ will be denoted by $[\mathbf{e}_i^* \mathbf{e}_j]$.

Lemma 2.2

Let L be a subspace of R^p and $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ be some vector basis of L. Then the rank of the matrix $[\mathbf{e}_i^* \mathbf{e}_j]$ is independent of a particular choice of the basis of L and is a property of the subspace L alone.

Proof. Suppose that some linear combination of vectors $\mathbf{e}_i^* \mathbf{e}_j$ is zero. We write this in the following form

$$\sum_{i=1}^{m} t_{ii} \mathbf{e}_i^2 + \sum_{i<j} 2t_{ij} \mathbf{e}_i^* \mathbf{e}_j = 0. \tag{2.8}$$

Equality (2.8) can be written as

$$\mathbf{e}_{k*} T \mathbf{e}_{k*}' = 0, \qquad k = 1, \ldots, p \tag{2.9}$$

where $\mathbf{e}_{k*}$ is the k-th row vector of the $p \times m$ matrix $[\mathbf{e}_i]$ and $T = [t_{ij}]$ is the $m \times m$ symmetric matrix. Now let $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ be another basis of L. Then $[\mathbf{e}_i] = [\mathbf{f}_i]C$, where C is some $m \times m$ nonsingular matrix. Then we obtain from (2.9) that

$$\mathbf{f}_{k*} CTC'\mathbf{f}'_{k*} = 0, \qquad k = 1, \ldots, p \tag{2.10}$$

where $\mathbf{f}_{k*}$ is the k-th row of $[\mathbf{f}_i]$.

Now let $[\mathbf{e}_i^*\mathbf{e}_j]$ be of rank $m(m + 1)/2 - s$. This means that there exist s linearly independent symmetric matrices $T_1, \ldots, T_s$ satisfying (2.9). The corresponding matrices $CT_1C', \ldots, CT_sC'$ in (2.10) are also linearly independent and therefore the rank of $[\mathbf{f}_i^*\mathbf{f}_j]$ is less or equal to $m(m + 1)/2 - s$. Because the two bases were chosen arbitrarily we obtain that $[\mathbf{e}_i^*\mathbf{e}_j]$ and $[\mathbf{f}_i^*\mathbf{f}_j]$ have the same rank. □

Lemma 2.3

Let A be a $p \times r$ matrix of rank r and $\{\mathbf{e}_1, \ldots, \mathbf{e}_{p-r}\}$ be a basis of the null space $N(A')$. Then the matrix equation

$$AX' + XA' + Y = C \tag{2.11}$$

where X is a $p \times r$ matrix of unknowns and Y is a $p \times p$ diagonal matrix of unknowns, has a solution for each $p \times p$ symmetric matrix C if and only if the matrix $[\mathbf{e}_i^*\mathbf{e}_j]$ has rank $(p - r)(p - r + 1)/2$, i.e., vectors $\mathbf{e}_i^*\mathbf{e}_j$, $i \leq j = 1, \ldots, p - r$, are linearly independent.

Proof. We have from lemma 2.1 and (2.7) that for fixed Y matrix equation (2.11) has a solution by X if and only if

$$\mathbf{e}_i'Y\mathbf{e}_j = b_{ij} \tag{2.12}$$

for $i,j = 1, \ldots, p - r$ and $b_{ij} = \mathbf{e}_i'C\mathbf{e}_j$.

Condition (2.12) can be considered now as a system of $(p - r)(p - r + 1)/2$ linear equations with p unknowns and the coefficient matrix $[\mathbf{e}_i^*\mathbf{e}_j]'$. This system has a solution by Y for each choice of b_{ij}, $i \leq j$, if and only if the coefficient matrix has linearly independent rows, i.e., the matrix $[\mathbf{e}_i^*\mathbf{e}_j]$ has the full column rank $(p - r)(p - r + 1)/2$. We complete the proof noting that because vectors $\mathbf{e}_1, \ldots, \mathbf{e}_{p-r}$ are linearly independent each $(p - r) \times (p - r)$ symmetric matrix $B = [b_{ij}]$ can be represented in the form $b_{ij} = \mathbf{e}_i'C\mathbf{e}_j$ for suitable choice of C. □

Definition 2.1

Let S and H be $p \times p$ symmetric and diagonal matrices respectively and let $\{\mathbf{e}_1, \ldots, \mathbf{e}_{p-r}\}$ be a vector basis of the null space $N(S - H)$. We say that $\mathbf{h}$ is a regular point of S (of rank r), if the matrix $[\mathbf{e}_i^*\mathbf{e}_j]$ has the full column rank $(p - r)(p - r + 1)/2$ and singular otherwise. For a given set of indices $I = \{i_1, \ldots, i_k\}$ we say that $\mathbf{h}$ is I-regular if $[\mathbf{e}_i^*\mathbf{e}_j]$ remains of rank $(p - r)(p - r + 1)/2$ after deletion of k rows numbered $i_1, \ldots, i_k$.

Note that by lemma 2.2 the rank of $[\mathbf{e}_i^*\mathbf{e}_j]$ is independent of a particular choice of the basis of $N(S - H)$. It can be shown similar to the proof of lemma 2.2 that the rank of $[\mathbf{e}_i^*\mathbf{e}_j]$ after deletion of k rows is also independent of the basis.

Theorem 2.1

The set of symmetric matrices with singular point has (Lebesgue) measure zero.

Proof. Let $f(\mathbf{x})$ be some infinitely differentiable function from R^{n_1} to R^{n_2}. A point $\mathbf{x} \in R^{n_1}$ is said to be critical if the rank of the Jacobian matrix $Df(\mathbf{x})$ at $\mathbf{x}$ is less than n_2, and a point $\mathbf{z} \in R^{n_2}$ is said to be a critical value if there exists a critical point $\mathbf{x}$ such that $f(\mathbf{x}) = \mathbf{z}$. The famous Sard Theorem [e.g., Sternberg, 1964] states that the set of critical values has measure zero.

An equivalent condition for the Jacobian matrix $Df(\mathbf{x})$ to be of rank n_2 is that the differential $df(\mathbf{x})$ of f at $\mathbf{x}$ as a linear transformation from R^{n_1} to R^{n_2} is *on* (i.e., its range is the whole space R^{n_2}).

Now let R^{n_1} be the vector space of pairs (X, Y), where X is a $p \times r$ matrix and Y is a $p \times p$ diagonal matrix ($n_1 = pr + p$) and R^{n_2} be the vector space of $p \times p$ symmetric matrices ($n_2 = p(p+1)/2$). We define the function $f : R^{n_1} \to R^{n_2}$ as follows:

$$f(X, Y) = X\Phi X' + Y \tag{2.13}$$

where Φ is a fixed $r \times r$ diagonal matrix whose diagonal entries are plus or minus one. Then the differential df of f at (A, Y) is

$$df(dX, dY) = (A\Phi)(dX)' + (dX)(A\Phi)' + dY. \tag{2.14}$$

It follows from lemma 2.3 and the definitions that (A, Y) is a critical point of f if and only if $\mathbf{y}$ is a singular point (of rank r) of the symmetric matrix S, $S = A\Phi A' + Y$. In other words a symmetric matrix S is a critical value of f if and only if S has a singular point $\mathbf{y}$ and $S - Y$ is equal to $A\Phi A'$ for a suitable choice of A. We have by the Sard Theorem that the set of such matrices is of measure zero. Changing the signs of the diagonal entries of Φ and the rank r we can cover all possible cases. Finally, noting that a set which is the union of a finite number of sets of measure zero also has measure zero, we complete the proof. □

It is clear from definition 2.1 that the necessary (but generally not sufficient) condition for $\mathbf{h}$ to be regular of rank r is that $(p-r)(p-r+1)/2 \leq p$ or equivalently $r \geq \phi(p) = [2p + 1 - (8p+1)^{1/2}]/2$. Therefore there are no regular points of rank less than the Ledermann bound $\phi(p)$. Then the next result follows from theorem 2.1.

Theorem 2.2

A reduced rank of the $p \times p$ sample covariance matrix is greater than or equal to $\phi(p)$ *almost surely.*

Definition 2.2

We say that a symmetric matrix S_0 is reducible to rank r at $\mathbf{h}_0$ if the matrix $S_0 - H_0$ has rank r. The reduced rank r is said to be stable (at $\mathbf{h}_0$) if for each symmetric matrix S in some neighborhood of S_0 there exists a vector $\mathbf{h}$ such that $S - H$ is of rank r and $\mathbf{h} \to \mathbf{h}_0$ as $S \to S_0$.

It follows from theorems 2.1 and 2.2 that a reduced rank less than the Ledermann bound cannot be stable. The following theorem shows that the stability is usually expected if $r \geq \phi(p)$.

Theorem 2.3

If $\mathbf{h}_0$ is a regular point of rank r of S_0 then the reduced rank r is stable at $\mathbf{h}_0$.

Proof. Suppose that $\mathbf{h}_0$ is a regular point of S_0 of rank r. Then the matrix $S_0 - H_0$ is of rank r and therefore can be represented in the following form

$$S_0 - H_0 = A\Phi A' \tag{2.15}$$

where A is a $p \times r$ matrix of rank r and Φ is an $r \times r$ nonsingular symmetric (Gramian if $S_0 - H_0$ is Gramian) matrix. We fix Φ and define the function f as in the proof of theorem 2.1:

$$f(X, Y) = X\Phi X' + Y. \tag{2.16}$$

It is clear from (2.15) and (2.16) that $f(A, H_0) = S_0$. As in the proof of theorem 2.1 we have by lemma 2.3 that, because $\mathbf{h}_0$ is a regular point of S_0, the differential df of f at (A, H_0)

(as a linear transformation from R^{n_1} to R^{n_2}) is *on*. Now we use the Implicit-Function Theorem [e.g., Ortega & Rheinboldt, 1970, p. 128], which for our purposes can be formulated in the following form:

> Let g be a continuously differentiable function from R^{n_1} to R^{n_2} and $\mathbf{z}_0 = g(\mathbf{x}_0)$. If the differential of g at $\mathbf{x}_0$ is *on*, then there exists such $\mathbf{x} \in R^{n_1}$ that $g(\mathbf{x}) = \mathbf{z}$ whenever $\mathbf{z}$ in some neighborhood of $\mathbf{z}_0$ and $\mathbf{x} \to \mathbf{x}_0$ as $\mathbf{z} \to \mathbf{z}_0$.

Applying the Implicit Function Theorem in the form above to the function f at (A, H_0) we obtain that any symmetric matrix S in some neighborhood of S_0 can be represented as $S = X\Phi X' + H$, where $\mathbf{h} \to \mathbf{h}_0$ as $S \to S_0$. □

It can be seen that for $r \geq \phi(p)$ there exist matrices, which have a regular point of rank r. If S is such a matrix, then by theorem 2.3 the whole neighborhood of S is reducible to the rank r. Therefore we obtain the following.

Corollary 2.1

The set of such matrices, which can be reduced to rank greater or equal to the Ledermann bound, has positive measure.

Together with theorem 2.2 this shows that $\phi(p)$ is the greatest lower bound (almost surely) to the minimal reduced rank of the sample covariance matrix.

Note that we treated in this section the reduced rank problem from the analytical-algebraic point of view. And neither S nor the reduced matrix $S - H$ have been required to be Gramian.

3. *Differential Properties of MTFA Functions*

This section is concerned with the continuous and differential properties of functions involved in the MTFA. Let S be a $p \times p$ Gramian matrix. We consider the two following functions of S, which correspond to the MTFA and the CMTFA respectively:

$$t(S) = \max_{H} trH, \text{ subject to } S - H \text{ Gramian} \tag{3.1}$$

$$t^*(S) = \max_{H} trH, \text{ subject to } S - H \text{ and } H \text{ Gramian} \tag{3.2}$$

and H is diagonal.

We investigate these functions utilizing convex analysis techniques. The results of this section will provide us with the required tools for the asymptotic theory of the next section.

We begin with the following result that was proven by Della Riccia and Shapiro [1980, Note 1] and independently by Ten Berge, Snijders and Zegers [1981].

Theorem 3.1

The CMTFA problem (3.2) [the MTFA problem (3.1)] has unique solution $H^* = H^*(S)$ for any Gramian matrix S.

Bentler [1972] gave necessary conditions for the solution point of MTFA. Della Riccia and Shapiro [1980, Note 1] and Ten Berge, Snijders and Zegers [1981] independently proved that similar conditions are necessary and sufficient for MTFA and CMTFA. We formulate these conditions in the form proposed by Della Riccia and Shapiro.

Theorem 3.2

A diagonal matrix $H^* = H^*(S)$ is the solution of the CMTFA problem (3.2) if and only if $S - H^*$ and H^* are Gramian and vector $1 = (1, \ldots, 1)'$ can be represented in the following form:

$$1 = \sum_{i=1}^{m} \mathbf{f}_i^2 - \sum_{j \in I(H^*)} \alpha_j \eta_j \tag{3.3}$$

where

$\mathbf{f}_i \in N(S - H^*)$, $i = 1, \ldots, m$; $\quad I(H) = \{i : h_i = 0, 1 \leq i \leq p\}$; $\quad \alpha_j, j \in I(H^*)$,

are nonnegative numbers

and

$$\boldsymbol{\eta}_i = (0, \ldots, 0, \underset{i\text{-th}}{1}, 0, \ldots, 0)'$$

is the i-th co-ordinate vector.

It will be convenient to formulate equality (3.3) in the following form:

Lemma 3.1

Equality (3.3) is equivalent to:

$$1 = \sum_{i,j=1}^{p-r} t_{ij} \mathbf{e}_i^* \mathbf{e}_j - \sum_{j \in I(H^*)} \alpha_j \boldsymbol{\eta}_j \tag{3.4}$$

where $\{\mathbf{e}_1, \ldots, \mathbf{e}_{p-r}\}$ is a vector basis of the null space $N(S - H^*)$ and $T = [t_{ij}]$ is a $(p - r) \times (p - r)$ Gramian matrix.

Proof. Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ and $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ be two sets of vectors, such that $[\mathbf{f}_i] = [\mathbf{e}_i]C$, where C is a $k \times m$ matrix. Then [compare with equalities (2.9) and (2.10) in the proof of lemma 2.2]

$$\sum_{i,j=1}^{k} t_{ij} \mathbf{e}_i^* \mathbf{e}_j = \sum_{i,j=1}^{m} t_{ij}^* \mathbf{f}_i^* \mathbf{f}_j \tag{3.5}$$

for any symmetric matrices $T = [t_{ij}]$ and $T^* = [t_{ij}^*]$ of order $k \times k$ and $m \times m$ respectively such that

$$T = CT^*C'. \tag{3.6}$$

Let now $\{\mathbf{e}_1, \ldots, \mathbf{e}_{p-r}\}$ be a vector basis of $N(S - H^*)$ and $\mathbf{f}_1, \ldots, \mathbf{f}_m$ be vectors of equality (3.3). Then we have that

$$\sum_{i=1}^{m} \mathbf{f}_i^2 = \sum_{i,j=1}^{p-r} t_{ij} \mathbf{e}_i^* \mathbf{e}_j \tag{3.7}$$

where $T = CC'$ and therefore T is Gramian. On the other hand if T is Gramian then $T = CC'$ and defining $[\mathbf{f}_i] = [\mathbf{e}_i]C$ we obtain equality (3.7). □

We note that for MTFA or if H^* is positive definite the sum $\Sigma \alpha_j \boldsymbol{\eta}_j$ in (3.3) and (3.4) has to be omitted.

Theorem 3.3

Let m be the minimal number of vectors $\mathbf{f}_i$ in (3.3). Then $m(m + 1) \leq 2p$.

Proof. Suppose that $m(m + 1) > 2p$. Then the set of $m(m + 1)/2$ vectors $\mathbf{f}_i^* \mathbf{f}_j$, $i \leq j$, is linearly dependent. Therefore there exists an $m \times m$ symmetric matrix $T = [t_{ij}]$ such that

$$\sum_{i,j=1}^{m} t_{ij}\mathbf{f}_i^*\mathbf{f}_j = 0 \tag{3.8}$$

Let C be an $m \times m$ orthogonal matrix such that $C'TC$ is diagonal. We define the set $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ by $[\mathbf{e}_i] = [\mathbf{f}_i]C$ or equivalently $[\mathbf{f}_i] = [\mathbf{e}_i]C'$. Then as in the proof of lemma 3.1 [see equality (3.5)] we obtain

$$\sum_{i,j=1}^{m} t_{ij}\mathbf{f}_i^*\mathbf{f}_j = \sum_{i=1}^{m} q_i\mathbf{e}_i^2 \tag{3.9}$$

where q_i, $i = 1, \ldots, m$, are the diagonal entries of $C'TC$. Thus from (3.8) and (3.9) we have

$$\sum_{i=1}^{m} q_i\mathbf{e}_i^2 = 0. \tag{3.10}$$

Furthermore from orthogonality of C follows

$$\sum_{i=1}^{m} \mathbf{f}_i^2 = \sum_{i=1}^{m} \mathbf{e}_i^2. \tag{3.11}$$

Then from (3.10) and (3.11) we obtain

$$\sum_{i=1}^{m} \mathbf{f}_i^2 = \sum_{i=1}^{m} (1 + \alpha q_i)\mathbf{e}_i^2 \tag{3.12}$$

for each number α. Now we choose α such that all numbers $1 + \alpha q_i$, $i = 1, \ldots, m$, will be nonnegative and at least one of them will be zero. Finally putting $\mathbf{h}_i = (1 + \alpha q_i)^{1/2}\mathbf{e}_i$ we have from (3.12)

$$\sum_{i=1}^{m} \mathbf{f}_i^2 = \sum_{i=1}^{k} \mathbf{h}_i^2.$$

where $k < m$. But this contradicts the minimality of m. □

The result of the following key lemma was suggested by Bentler and Woodward [1980] and proven by Ten Berge, Snijders and Zegers [1981].

Lemma 3.2

$$t(S) = \min_F tr\ F'SF, \text{ subject to } \operatorname{diag}(FF') = I_p \tag{3.14}$$

$$t^*(S) = \min_F tr\ F'SF, \text{ subject to } \operatorname{diag}(FF') \geq I_p \tag{3.15}$$

where F is a $p \times m$ matrix and I_p is the $p \times p$ identity matrix. Furthermore the minimum is attained at $F = F_0$ if and only if $F_0 = [\mathbf{f}_i]$, where vectors $\mathbf{f}_i$, $i = 1, \ldots, m$, satisfy the necessary and sufficient conditions of theorem 3.2.

Lemma 3.2 gives the convenient representation of functions $t(S)$ and $t^*(S)$. It follows from theorem 3.3 that it is enough to take in lemma 3.2 only such m that $m(m + 1)/2 \leq p$.

Note that $tr\ F'SF = tr\ SFF'$, therefore (3.14) and (3.15) are equivalent to:

$$t(S) = \min_Z trSZ, \text{ subject to } Z \text{ Gramian and } \operatorname{diag}(Z) = I_p, \tag{3.16}$$

$$t^*(S) = \min_Z trSZ, \text{ subject to } Z \text{ Gramian and } \operatorname{diag}(Z) \geq I_p. \tag{3.17}$$

The sets of Gramian matrices Z, at which the minimum in (3.16) and (3.17) is attained, will be denoted $Z(S)$ and $Z^*(S)$ respectively. By lemma 3.2 $Z \in Z^*(S)$ if and only if $Z = \sum_{i=1}^{m} \mathbf{f}_i \mathbf{f}_i'$ where vectors $\mathbf{f}_i, i = 1, \ldots, m$, satisfy the necessary and sufficient conditions of theorem 3.2.

Theorem 3.4

The function $t^*(S)$ is concave and if S_0 is a positive definite matrix, then

(a) t^* is continuous at S_0

(b) The directional derivative

$$t^{*\prime}(S_0\,;\,Q) = \lim_{\alpha \to 0^+} \frac{t^*(S_0 + \alpha Q) - t^*(S_0)}{\alpha} \tag{3.18}$$

of t^* at S_0 in a direction Q exists for each symmetric matrix Q and

$$t^{*\prime}(S_0\,;\,Q) = \min_{Z \in Z^*(S_0)} tr\, QZ \tag{3.19}$$

(c) t^* is differentiable at S_0 if and only if the set $Z^*(S_0)$ is singleton, i.e., the set $Z^*(S_0)$ contains only one matrix Z_0. In the last case the differential dt^* of t^* at S_0 is

$$dt^*(dS) = tr\, Z_0\, dS. \tag{3.20}$$

The analogous results follow for the function $t(S)$.

Proof. It is clear that $trSZ$ is a linear, and therefore concave function of S. Then we have from (3.17) that $t^*(S)$ can be considered as the minimum of concave functions. This implies that $t^*(S)$ is concave [e.g., Rockafellar, 1970].

Let S_0 be a positive definite matrix. Then S_0 is an interior point in the set of Gramian matrices. It is known [e.g., Rockafellar, 1970] that a concave function is continuous and has the directional derivatives in all directions at an interior point of its domain of definition. Therefore we obtain that the function t^* is continuous and has the directional derivatives at S_0.

As we mentioned above the function $t^*(S)$ can be considered as the minimum of concave functions $trSZ$, where Z is considered as a parameter belonging to the parameter space (of Gramian matrices with diagonal entries greater than or equal to one). Such min-functions have been a subject of much investigation. It follows from the minimax theory [e.g., Ioffe and Tihomirov, 1979, p. 201] that the directional derivatives of t^* at S_0 are the minimum of the corresponding directional derivatives of $trSZ$ at S_0, where the minimum is taken over all Z from the set $Z^*(S_0)$. It is clear that the directional derivative of $trSZ$ in the direction Q is $trQZ$. Therefore equality (3.19) follows.

It is known that a concave function $f(\mathbf{x})$ is differentiable at $\mathbf{x}_0$ if and only if the directional derivative $f'(\mathbf{x}_0\,;\,q)$ is linear in q [e.g., Rockafellar, 1970]. Together with (3.19) this implies that t^* is differentiable at S_0 if and only if $Z^*(S_0)$ is singleton. Now formula (3.20) follows from (3.19). □

As is known a concave function is differentiable almost everywhere [e.g., Rockafellar, 1970]. Therefore theorem 3.4 implies the following result.

Corollary 3.1

The minimum in (3.17) [in (3.16)] is attained almost everywhere at the unique Gramian matrix Z_0.

Now we are going to discuss the uniqueness of Z_0 in greater detail. Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_{p-r}\}$ be a vector basis of $N(S - H^*)$, where H^* is the solution of the CMTFA problem (3.2), and let $E = [\mathbf{e}_i]$. We have that $Z \in Z^*(S)$ if and only if $Z = FF'$, where $F = [\mathbf{f}_i]$ and vectors $\mathbf{f}_i$,

$i = 1, \ldots, m$, satisfy the necessary and sufficient conditions of theorem 3.2. Because $\{\mathbf{e}_1, \ldots, \mathbf{e}_{p-r}\}$ is a basis we have

$$F = EC \tag{3.21}$$

and then

$$Z = ETE' \tag{3.22}$$

where the matrix $T = CC'$ satisfies equality (3.4) [see equality (3.3) in the proof of Lemma 3.1]. Therefore we obtain that the set $Z^*(S)$ is singleton if and only if the matrix T in (3.4) is unique. It is clear that the linear independence of vectors $\mathbf{e}_i^*\mathbf{e}_j$, $i \le j$, implies the uniqueness of the matrix $T = [t_{ij}]$ in (3.4), if the sum $\Sigma\alpha_i \boldsymbol{\eta}_i$ is omitted. In the general case we need the linear independence of $\mathbf{e}_i^*\mathbf{e}_j$ after deletion of the co-ordinates corresponding to the index set $I(H^*)$. This leads to the following definition.

Definition 3.1

We say that a positive definite matrix S is CMTFA-regular (MTFA-regular) if $\mathbf{h}^*$ is an I-regular (regular) point of S, where H^* is the CMTFA (the MTFA) solution and $I = I(H^*)$.

From the discussion above we have that the set $Z^*(S)$ is singleton if S is CMTFA-regular. Together with theorem 3.4 this implies the following.

Theorem 3.5

If S_0 is CMTFA-regular (MTFA-regular) then the function t^* (the function t) is differentiable at S_0 and

$$dt^*(dS) = trZ_0\, dS = \sum_{i=1}^{m} \mathbf{f}_i'(dS)\mathbf{f}_i \tag{3.23}$$

where vectors $\mathbf{f}_i$, $i = 1, \ldots, m$, satisfy the necessary and sufficient conditions of theorem 3.2.

Note that CMFTA-regularity is only sufficient condition for differentiability of t^* and the matrix T can be unique even when S is not CMTFA-regular. However it seems that such a possibility is unlikely.

From theorem 2.1 we obtain the following.

Corollary 3.2

A Gramian matrix is MTFA-regular almost everywhere.

4. *Sampling Theory*

In this section we consider the applications of previous considerations to the sampling theory of the CMTFA. Let X be a p-dimensional random sample of $N(>p)$ observations, denoted by column vectors X_α, $\alpha = 1, \ldots, N$. The $p \times p$ sample covariance matrix S is defined by $S = n^{-1} \sum_\alpha (X_\alpha - \bar{X})(X_\alpha - \bar{X})'$, where $\bar{X}$ is the sample mean vector $N^{-1} \sum_\alpha X_\alpha$ and $n = N - 1$. Let $\Sigma = [\sigma_{ij}]$ be the population covariance matrix and the matrix U be defined by $U = S - \Sigma$. If the sample is drawn from a normally distributed population then $p(p+1)/2$ distinct elements of $n^{1/2}U$ are known to have an asymptotic multivariate normal distribution with zero mean and with variances and co-variances given by

$$nE(u_{ij}u_{he}) = \sigma_{ih}\sigma_{je} + \sigma_{ie}\sigma_{jh} \tag{4.1}$$

[e.g., Anderson, 1958].

If $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ are arbitrary column vectors, each having p elements, then the covariance of $\mathbf{a}'U\mathbf{b}$ and $\mathbf{c}'U\mathbf{d}$ is given by

$$n \operatorname{cov}(\mathbf{a}'U\mathbf{b}, \mathbf{c}'U\mathbf{d}) = (\mathbf{a}'\Sigma\mathbf{c})(\mathbf{b}'\Sigma\mathbf{d}) + (\mathbf{a}'\Sigma\mathbf{d})(\mathbf{b}'\Sigma\mathbf{c}) \tag{4.2}$$

[Lawley & Maxwell, 1971].

We consider the following parameters of the population co-variance matrix Σ:

$$t = t(\Sigma),\ t^* = t^*(\Sigma),\ \rho = 1 - \frac{t}{1'\Sigma 1} \quad \text{and} \quad \rho^* = 1 - \frac{t^*}{1'\Sigma 1}$$

where $t(\Sigma)$ and $t^*(\Sigma)$ defined by (3.1) and (3.2) respectively. It can be seen that ρ is a lower bound and ρ^* is the greatest lower bound to reliability as it was defined in the introduction. In the absence of Σ one substitutes for it the sample estimate S. Therefore we consider the following sample statistics:

$$\hat{t} = t(S), \qquad \hat{t}^* = t(S), \qquad \hat{\rho} = 1 - \frac{\hat{t}}{1'S1} \quad \text{and} \quad \hat{\rho}^* = 1 - \frac{\hat{t}^*}{1'S1}.$$

We suppose in what follows that the population covariance matrix Σ is positive definite.

Theorem 4.1

The sample statistics $\hat{t}$, $\hat{t}^*$, $\hat{\rho}$ and $\hat{\rho}^*$ are consistent estimates of t, t^*, ρ and ρ^* respectively.

Proof. By theorem 3.4(a) we have that the functions t and t^* are continuous at Σ. Then from the convergence in probability of S to Σ follows the convergence in probability of $\hat{t}$ and $\hat{t}^*$ to t and t^* and therefore of $\hat{\rho}$ and $\hat{\rho}^*$ to ρ and ρ^* respectively. □

Now we are going to discuss the asymptotic distribution of the CMTFA sample estimates. If the function t^* is differentiable at Σ then the differential dt^* of t^* at Σ represents the linear approximation to t^* and we can write [see for example the δ-method of Rao (1973, p. 388)]

$$n^{1/2}(\hat{t}^* - t^*) \stackrel{a}{=} dt^*(n^{1/2}U) \tag{4.3}$$

where "$\stackrel{a}{=}$" is read "is asymptotically equal to" and means that the difference between the left and right sides of (4.3) approaches zero in probability as the sample size approaches infinity. This together with the results of the previous section enables us to obtain the following result.

Theorem 4.2

Let the population covariance matrix Σ be CMTFA-regular. Then $n^{1/2}(\hat{t}^* - t^*)$ and $n^{1/2}(\hat{\rho}^* - \rho^*)$ have asymptotically normal distributions with zero means and variances

$$2 \sum_{i,j=1}^{m} (\mathbf{f}_i'\Psi^*\mathbf{f}_j)^2 \tag{4.4}$$

$$\frac{2\left[\sum_{i,j=1}^{m} (\mathbf{f}_i'\Psi^*\mathbf{f}_j)^2 - \frac{2t^*}{\gamma} \sum_{i=1}^{m} (1'\Psi^*\mathbf{f}_i)^2 + t^{*2}\right]}{\gamma^2} \tag{4.5}$$

respectively, where Ψ^* is the solution of the CMTFA problem (for the matrix Σ), vectors $\mathbf{f}_i$, $i = 1, \ldots, m$, satisfy the necessary and sufficient conditions of theorem 3.2 [i.e., $\mathbf{f}_i$, $i = 1, \ldots, m$, belong to $N(\Sigma - \Psi^*)$ and satisfy equality (3.3)] and $\gamma = 1'\Sigma 1$.

Proof. Because the matrix Σ is CMTFA-regular we have by theorem 3.5 that the function t^* is differentiable at Σ and the differential dt^* of t^* at Σ is given by formula (3.23). Together with (4.3) this implies

$$n^{1/2}(\hat{t}^* - t^*) \overset{a}{=} \sum_{i=1}^{m} \mathbf{f}_i'(n^{1/2}U)\mathbf{f}_i. \tag{4.6}$$

The right side of (4.6) is asymptotically normal with zero mean and variance.

$$\mathrm{var}\left\{\sum_{i=1}^{m} \mathbf{f}_i'(n^{1/2}U)\mathbf{f}_i\right\} = \sum_{i,j=1}^{m} n\ \mathrm{cov}(\mathbf{f}_i'U\mathbf{f}_i, \mathbf{f}_j'U\mathbf{f}_j). \tag{4.7}$$

From (4.2) we obtain

$$n\ \mathrm{cov}(\mathbf{f}_i'U\mathbf{f}_i, \mathbf{f}_j'U\mathbf{f}_j) = 2(\mathbf{f}_i'\Sigma\mathbf{f}_j)^2 \tag{4.8}$$

and because $\mathbf{f}_j$ belongs to $N(\Sigma - \Psi^*)$ we have

$$\Sigma\mathbf{f}_j = (\Sigma - \Psi^*)\mathbf{f}_j + \Psi^*\mathbf{f}_j = \Psi^*\mathbf{f}_j. \tag{4.9}$$

Equalities (4.7), (4.8) and (4.9) imply that the asymptotic variance of the right side of (4.6) is given by formula (4.4). It remains to be noted that because of (4.6) the asymptotic distribution of $n^{1/2}(\hat{t}^* - t^*)$ is the same as that of $\sum_{i=1}^{m} \mathbf{f}_i'(n^{1/2}U)\mathbf{f}_i$ [e.g., Rao, 1973, p. 122]. Analogously

$$n^{1/2}(\hat{\rho}^* - \rho^*) \overset{a}{=} \frac{t^*\mathbf{1}'(n^{1/2}U)\mathbf{1} - \gamma \sum_{i=1}^{m} \mathbf{f}_i'(n^{1/2}U)\mathbf{f}_i}{\gamma^2} \tag{4.10}$$

and therefore $n^{1/2}(\hat{\rho}^* - \rho^*)$ has an asymptotically normal distribution with zero mean and variance

$$\frac{t^{*2}\ \mathrm{var}\{\mathbf{1}'(n^{1/2}U)\mathbf{1}\} - 2\gamma t^*\ \mathrm{cov}\left\{\mathbf{1}'(n^{1/2}U)\mathbf{1}, \sum_{i=1}^{m} \mathbf{f}_i'(n^{1/2}U)\mathbf{f}_i\right\} + \gamma^2\ \mathrm{var}\left\{\sum_{i=1}^{m} \mathbf{f}_i'(n^{1/2}U)\mathbf{f}_i\right\}}{\gamma^4}$$

$$= \frac{2t^{*2}\gamma^2 - 4\gamma t^* \sum_{i=1}^{m} (\mathbf{1}'\Sigma\mathbf{f}_i)^2 + 2\gamma^2 \sum_{i,j=1}^{m} (\mathbf{f}_i'\Sigma\mathbf{f}_j)^2}{\gamma^4} \tag{4.11}$$

which is by (4.9) equivalent to formula (4.5). □

If Σ is MTFA-regular then a similar theorem holds for $n^{1/2}(\hat{t} - t)$ and $n^{1/2}(\hat{\rho} - \rho)$.

We have shown that the sample statistics $n^{1/2}(\hat{t}^* - t^*)$ and $n^{1/2}(\hat{\rho}^* - \rho^*)$ have asymptotically normal distributions provided Σ is CMTFA-regular. A necessary condition for CMTFA-regularity (MTFA-regularity) of Σ is that $r \geq \phi(p)$, where r is the (reduced) rank of $\Sigma - \Psi^*$ and $\phi(p)$ is the Ledermann bound. However the most interesting cases are concerned with small r relative to p. If $r < \phi(p)$ we shall expect the function t^* (the function t) to be nondifferentiable at Σ and therefore the classical differential approach is not applicable in this case. We cannot say very much in the case $r < \phi(p)$. Even to calculate the bias $E(\hat{t}^* - t^*)$ seems to be a very complicated problem. We only note that it follows from the concavity of $t^*(S)$ that the mean of $\hat{t}^* - t^*$ is less than zero (even for CMTFA-regular Σ). Especially the sample estimates become heavily biased ($E\{\hat{t}^*\} < t^*$ and $E\{\hat{\rho}^*\} > \rho^*$) when r is much less than $\phi(p)$.

REFERENCE NOTE

Della Riccia, G. & Shapiro, A. *Minimum rank and minimum trace of covariance matrices.* Department of Mathematics, Ben-Gurion University of the Negev, Beer-Sheva, Israel, 1980.

REFERENCES

Anderson, T. W. *An introduction to multivariate statistical analysis.* New York, Wiley, 1958.

Bentler, P. M. A lower-bound method for the dimension-free measurement of internal consistency. *Social Science Research,* 1972, *1,* 343–357.

Bentler, P. M. & Woodward, J. A. Inequalities among lower bounds to reliability: with applications to test construction and factor analysis. *Psychometrika,* 1980, *45,* 249–267.

Guttman, L. To what extent can communalities reduce rank? *Psychometrika,* 1958, *23,* 297–308.

Harman, H. *Modern factor analysis.* Chicago: University of Chicago Press, 1976.

Ioffe, A. D. & Tihomirov, V. M. *Theory of extremal problems.* North-Holland Publishing Company, 1979.

Jackson, P. H. & Agunwamba, C. C. Lower bounds for the reliability of the total score on a test composed of nonhomogeneous items: I: Algebraic lower bounds. *Psychometrika,* 1977, *42,* 567–578.

Lawley, D. N. & Maxwell, A. E. *Factor analysis as a statistical method.* London Butterworths, second edition, 1971.

Ledermann, W. On the rank of the reduced correlation matrix in multiple-factor analysis. *Psychometrika,* 1937, *2,* 85–93.

Ledermann, W. On a problem concerning matrices with variable diagonal elements. *Proceedings of the Royal Society of Edinburgh,* 1939, *60,* 1–17.

Ortega, J. M. & Rheinboldt, W. C. *Iterative solution of nonlinear equations in several variables.* New York, Academic Press, 1970.

Rao, C. R. *Linear statistical inference and its applications.* John Wiley and Sons, second edition, New York, 1973.

Rockafellar, R. T. *Convex analysis.* Princeton, University Press, 1970.

Shapiro, A. Weighted minimum trace factor analysis. *Psychometrika,* in press.

Sternberg, S. *Lectures on differential geometry.* Prentice Hall, Inc., Englewood Cliffs, N.J., 1964.

Ten Berge, J. M. F., Snijders, T. A. B. & Zegers, F. E. Computational aspects of the greatest lower bound to the reliability and constrained minimum trace factor analysis. *Psychometrika,* 1981, *46,* 201–213.

Woodhouse, B. & Jackson, P. M. Lower bounds for the reliability of the total score on a test composed of nonhomogeneous items: II: A search procedure to locate the greatest lower bound. *Psychometrika,* 1977, *42,* 579–591.

Manuscript received 2/23/81
Final version received 2/3/82