THEORETICAL ASPECTS OF COMPONENT ANALYTIC PROCEDURES FOR INTERBATTERY COMPARISONS

BY

HARINDER NANDA and INGRAM OLKIN

TECHNICAL REPORT NO. 175
OCTOBER 1981

PREPARED UNDER THE AUSPICES OF
NATIONAL SCIENCE FOUNDATION GRANT
MCS 78-07736

Ingram Olkin, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
THEORETICAL ASPECTS OF COMPONENT ANALYTIC PROCEDURES FOR INTERBATTERY COMPARISONS

BY

HARINDER NANDA and INGRAM OLKIN

TECHNICAL REPORT NO. 175
OCTOBER 1981

PREPARED UNDER THE AUSPICIES

OF

NATIONAL SCIENCE FOUNDATION GRANT
MCS 78-07736

Ingram Olkin, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
Theoretical Aspects of Component Analytic Procedures for Interbattery Comparisons

by

Harinder Nanda
New Delhi

and

Ingram Olkin\textsuperscript{1}/
Stanford University

Abstract

The solutions of Tucker and Rao for interbattery component analysis are shown to be transforms of one another. These solutions are extended to provide a wider class of solutions from which a specific solution can be chosen to satisfy some criteria.

Key words: canonical correlations, factor analysis, component analysis, interbattery analysis.

\textsuperscript{1}/Supported in part by the National Science Foundation.
1. Introduction. Consider two batteries of $p_1$ and $p_2$ tests with scores $x$ that satisfy the model

$$
(1.1) \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} f + \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \equiv Af + g,
$$

where the column vectors $x_1$ and $x_2$ represent the scores on the $p_1$ and $p_2$ tests in batteries 1 and 2, respectively, $A_1 = (a_{ij}^{(1)})$ and $A_2 = (a_{ij}^{(2)})$ are $p_1 \times m$ and $p_2 \times m$ matrices that represent the loadings in interbattery factors for the variables in batteries 1 and 2, i.e., $a_{ij}^{(k)}$ represents the loading of variable $i$ on interbattery factor $j$ in battery $k$, $i = 1, \ldots, p; \ j = 1, \ldots, m; \ k = 1, 2$. The vector $f$ denotes the score vector on $m$ interbattery factors; $g_1$ and $g_2$ are $p_1$ and $p_2$-dimensional vectors that denote the score on a composite of battery specific factors for the variables.

The variables $f_1, \ldots, f_m$ are assumed to be independently distributed with unit variance. The vectors $f$, $g_1$, $g_2$ are independently distributed with $\text{Var}(g_1) = \psi_1$, $\text{Var}(g_2) = \psi_2$. As a consequence of the assumptions, the covariance matrix $\Sigma$ of $x$ has the structure

$$
(1.2) \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} = \begin{bmatrix} A_1 A_1' + \psi_1 & A_1 A_2' \\ A_2 A_1' & A_2 A_2' + \psi_2 \end{bmatrix} \equiv \Sigma(A, \psi).
$$

The problem is to estimate $A_1$, $A_2$, $\psi_1$, $\psi_2$ based on a sample of $N$ persons each of whom takes $p_1 + p_2$ tests in the two batteries.

If we assume an underlying normal distribution the maximum likelihood estimators can be obtained numerically, and programs are now
available to carry out such an analysis (see e.g., Jöreskog, 1981, for a list of references). Maximum likelihood estimation does not, in general, lead to closed form expressions and does not normally yield exact distributional results. However, it does permit solutions for which asymptotic results can be applied.

Alternatively, the method of least squares can be used, namely

$$\min_{A, \psi} \| \hat{\Sigma} - \Sigma(A, \psi) \|,$$

for some norm $\| \cdot \|$, subject to the constraint that $\Sigma(A, \psi)$ be positive semi-definite.

Historically, other procedures have been proposed. From a sample of $N$ persons we obtain $p_1 \times N$ and $p_2 \times N$ score matrices $X_1$ and $X_2$ on the $p_1$ and $p_2$ tests in batteries 1 and 2, respectively. Using this data an estimate, $\hat{\Sigma}$, of the covariance matrix $\Sigma$ can be obtained. From (1.2) we can determine estimates $\hat{A}_1, \hat{A}_2, \hat{\psi}_1, \hat{\psi}_2$ satisfying

$$\hat{\Sigma}_{11} = \hat{A}_1 \hat{A}_1', \hat{\psi}_1, \hat{\Sigma}_{12} = \hat{A}_1 \hat{A}_2', \hat{\Sigma}_{22} = \hat{A}_2 \hat{A}_2' + \hat{\psi}_2'.$$

A solution of (1.3) can be obtained sequentially, namely, first determine $\hat{A}_1, \hat{A}_2$ to satisfy

$$\hat{\Sigma}_{12} = \hat{A}_1 \hat{A}_2', \hat{\Sigma}_{11} - \hat{A}_1 \hat{A}_1' \geq 0, \hat{\Sigma}_{22} - \hat{A}_2 \hat{A}_2' \geq 0,$$

where $U \geq V$ or $U \succeq V$ means that $U - V$ is positive semi-definite or
positive definite. Given a solution to (1.4), \( \hat{\psi}_1 \) and \( \hat{\psi}_2 \) can be determined uniquely.

The essence of other procedures has been to concentrate on solving (1.4). In this formulation the value of \( m \), the number of interbattery factors is assumed fixed. Because one objective of the interbattery model is scientific parsimony, a minimal value of \( m \) is desired. Since

\[
\text{rank}(\Sigma_{12}) = \text{rank}(A_1 A_2') \leq \min[\text{rank}(A_1), \text{rank}(A_2)] \\
\leq \max[\text{rank}(A_1), \text{rank}(A_2)] \leq m,
\]

and since it is possible for the minimal rank to be achieved, the solution to (1.4) should also satisfy the requirement that \( m = \text{rank}(\hat{\Sigma}_{12}) \).

Tucker (1958) used a standard matrix factorization (the singular-value decomposition) to determine \( A_1 \) and \( A_2 \) from \( \Sigma_{12} \) alone. (For the sake of simplicity, we omit the carat on the estimates and do not distinguish between population parameters and estimates when there is no danger of confusion.) This factorization has the property that it provides the minimum number of common factors, but suffers from the deficiency that the residual matrices \( \Sigma_{11} - A_1 A_1' \) and \( \Sigma_{22} - A_2 A_2' \) need not be positive semi-definite. Consequently, this factorization does not always provide an admissible solution to (1.4).

Rao (1965) offered a solution to (1.4) based on the canonical correlation factorization which also provides a minimum value for \( m \).

Kristof (1967) obtained a solution to (1.4) by first finding an orthogonal factor solution of \( \Sigma \) yielding
\[ \tilde{F} = \begin{bmatrix} \tilde{A}_1 \\ \tilde{A}_2 \end{bmatrix}, \]

and then applying an orthogonal transformation to achieve the format

\[
\begin{bmatrix} \tilde{A}_1 \\ \tilde{A}_2 \end{bmatrix} \rightarrow \begin{bmatrix} A_1 & G_1 & 0 \\ A_2 & 0 & G_2 \end{bmatrix}.
\]

Kristof's solution has the virtue of providing a complete factorization in the sense that \( G_1 \) and \( G_2 \) satisfying

\[ G_1 G_1' = \Sigma_{11} - A_1 A_1', \quad G_2 G_2' = \Sigma_{22} - A_2 A_2' \]

are obtained initially, whereas other procedures require the additional solution of

\[ \psi_1 = G_1 G_1', \quad \psi_2 = G_2 G_2' \]

for \( G_1 \) and \( G_2 \). However, this latter factorization is now relatively standard. More important, the Kristof solution does not necessarily provide the least number of common factors.

Suppose \( A_1 \) and \( A_2 \) satisfy \( A_1 A_2' = \Sigma_{12} \), then for any nonsingular matrix \( M \),

\[ (1.5) \quad \tilde{A}_1 = A_1 M, \quad \tilde{A}_2 = A_2 M^{-1} \]

also satisfies \( \tilde{A}_1 \tilde{A}_2' = \Sigma_{12} \). Conversely, if
\[ A'_1 A'_2 = \tilde{A}'_1 \tilde{A}'_2 = \Sigma_{12}, \]

then there exists a nonsingular matrix \( M \) satisfying (1.5). Consequently, all factorizations of \( \Sigma_{12} \) are connected by (1.5), so that (1.5) provides a mechanism to generate a wider class of solutions from any single one. This freedom permits the choice of a particular matrix \( M \) that has desirable properties. One required property is that the residual matrices be positive semi-definite, i.e., that

\[ \Sigma_{11} - A_1 (M M') A'_1 \geq 0, \quad \Sigma_{22} - A_2 (M M')^{-1} A'_2 \geq 0. \]

That such an \( M \) exists is known from Rao's procedure. However, this result is stated more formally in a theorem.

We note that Gibson (1960, 1961, 1963) considered least squares approximations for \( M \) from each of \( \Sigma_{11} \approx A_1 (M M') A'_1 \) and \( \Sigma_{22} \approx A_2 (M M')^{-1} A'_2 \). A subsequent adjustment was then made so that the solutions were compatible.

The above proposals do not take account of possible connections between the two batteries. For example, if we have two parallel batteries, then any difference between scores in the two batteries is due to errors of measurement in the tests. In this case, the matrices \( A_1 \) and \( A_2 \) of loadings of the variables on interbattery factors should be the same, and hence the estimates of \( A_1 \) and \( A_2 \) should be in close agreement. This suggests that \( A_1 \) and \( A_2 \) be chosen so as to minimize \( \| A_1 - A_2 \| \).

The Tucker and Rao solutions are embedded in a wider class of solutions (Sections 2 and 3) and compared (Section 4). Criteria and an algorithm for choosing a solutions are given in Section 5.

Certain results from matrix theory are needed. These can be found in Bellman (1972), MacDuffee (1946), Marshall and Olkin (1979, Chapters 16, 19),
or Mirsky (1955).

2. Extension of Tucker's Solution. To obtain a factorization of

\[ \Sigma_{12} = A_1 A_2', \]

of rank \( r \), Tucker uses the Eckart-Young singular-value decomposition:

\[ \Sigma_{12} = \Gamma D \gamma \Delta', \]

where \( \Gamma \) and \( \Delta \) are \( p_1 \times r \) and \( p_2 \times r \) matrices, respectively, with orthogonal columns, i.e.,

\[ \Gamma' \Gamma = \Delta' \Delta = I_r, \quad D_{\gamma} = \text{diag}(\gamma_1, \ldots, \gamma_r), \]

\( \gamma_1, \ldots, \gamma_r \) are the positive square roots of the characteristic roots of

\[ \Sigma_{12} \Sigma_{21}'. \]

The determination of \( \Gamma \) and \( \gamma \) can be made from

\[ \Sigma_{12} \Sigma_{21} = \Gamma D_{\gamma}^2 \Gamma', \]

from which

\[ \Delta = \Sigma_{21} \Gamma D_{\gamma}^{-1}. \]

(If \( \gamma_1, \ldots, \gamma_r \) are distinct, then the factorization is unique except possibly for alternations of sign in \( \Gamma \).)

The choice

\[ (2.1) \quad A_1 = \Gamma D_{\gamma}^{1/2}, \quad A_2 = \Delta D_{\gamma}^{1/2}, \]

then satisfies \( \Sigma_{12} = A_1 A_2' \). Since \( r \) is the rank of \( \Sigma_{12} \), this procedure yields \( m = r \), which achieves the smallest number of interbattery factors.

**Remark 1.** Note that \( A_1' A_1 = A_2' A_2 = D_{\gamma} \). Therefore \( A_1 \) and \( A_2 \) are defined so that the amount of variance accounted for by common factors is the same in both batteries. This may not always be a reasonable property. For example, if one of the batteries has more variables measuring a common factor than the other battery, the variance accounted for by that common factor in the first battery will be greater than in the second battery.
A key difficulty in the choice (2.1) is that the residuals \( \Sigma_{11} - A_1 A_1' \) and \( \Sigma_{22} - A_2 A_2' \) need not be positive semi-definite. A sufficient condition that the residuals be positive semi-definite is that the smallest characteristic roots of \( \Sigma_{11} \) and \( \Sigma_{22} \) be less than or equal to unity.

To generate a wider class of solutions from which we choose one that satisfies (1.4), recall that the factorization \( \Sigma_{12} = A_1 A_2' \) is invariant under the transformation

\[
A_1 \rightarrow A_1 M, \quad A_2' \rightarrow M^{-1} A_2'.
\]

Consequently, we wish to choose an \( M \) such that

\[
(2.2) \quad \Sigma_{11} - A_1 M M' A_1' > 0, \quad \Sigma_{22} - A_2 M' M^{-1} A_2' > 0.
\]

The following development shows that such an \( M \) exists.

**Theorem 1.** If \( \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} > 0 \) and \( \Sigma_{12} = A_1 A_2' \), then there exists a nonsingular matrix \( M \) such that (2.2) holds.

Before providing a proof of theorem 1, we require a Lemma.

**Lemma 2.** If \( C \) is a \( p \times n \) matrix of rank \( p \) and \( Z > 0 \), then

\[
(2.3) \quad C' Z C < I
\]

if and only if

\[
(2.4) \quad Z \leq (C C')^{-1}
\]
Proof of Lemma 1. From the fact that $U > V$ implies that $E U E' > E V E'$ for any matrix $E$, (2.3) implies that

$$(C C') Z (C C') \leq C C'.$$

Pre- and post-multiplication by $(C C')^{-1}$ then yields (2.4). Conversely, if (2.4) holds, then $C' Z C \leq (C C')^{-1} C = L$. But $L$ is idempotent, and hence $L \leq I$, which implies (2.3). $||$

Proof of Theorem 1. In (2.2) let $Q = M M'$ so that (2.2) becomes

$$(2.5) \quad \Sigma_{11} \geq A_1^T Q A_1', \quad \Sigma_{22} \geq A_2^{-1} Q^{-1} A_2',$$

or equivalently, because $\Sigma > 0$ and hence $\Sigma_{11} > 0, \Sigma_{22} > 0$,

$$(2.6) \quad I \geq \Sigma^{-\frac{1}{2}}_{11} A_1^T Q A_1' \Sigma^{-\frac{1}{2}}_{11}, \quad I \geq \Sigma^{-\frac{1}{2}}_{22} A_2^{-1} Q^{-1} A_2' \Sigma^{-\frac{1}{2}}_{22},$$

where for $R > 0$, $R^{\frac{1}{2}}$ is the unique positive semi-definite matrix satisfying $(R^{\frac{1}{2}})^2 = R$.

A direct application of Lemma 1 to (2.6) yields

$$(2.7) \quad Q \leq (A_1^T \Sigma_{11}^{-1} A_1)^{-1}, \quad Q^{-1} \leq (A_2^T \Sigma_{22}^{-1} A_2)^{-1}.$$

From the fact that $U > V > 0$ is equivalent to $V^{-1} > U^{-1} > 0$, (2.7) can be written as

$$(2.8) \quad B_2 \equiv A_2^T \Sigma_{22}^{-1} A_2 \leq Q \leq (A_1^T \Sigma_{11}^{-1} A_1)^{-1} \equiv B_1^{-1}.$$
Given any $A_1$ and $A_2$, inequality (2.8) has a solution $Q$ provided $B_2 \leq B_1^{-1}$, or equivalently that $B_1^{1/2} B_2 B_1^{1/2} \leq I$, which is equivalent to the fact that the characteristic roots, $\lambda(B_1^{1/2} B_2 B_1^{1/2})$, are less than or equal to one. But

$$
\lambda(B_1^{1/2} B_2 B_1^{1/2}) = \lambda(B_2 B_1) = \lambda(A_1' \Sigma_{11}^{-1} A_1 + A_2' \Sigma_{22}^{-1} A_2 + \Sigma_{11}^{-1} A_1 A_1')
$$

$$
= \lambda(\Sigma_{22}^{-1} A_2 A_1' \Sigma_{11}^{-1} A_1 A_2') = \lambda(\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}),
$$

which are the squared canonical correlations and are less than or equal to one. ||

To obtain a class of solutions $Q$ satisfying (2.8), simultaneously decompose $B_1$ and $B_2$:

$$
B_1^{-1} = W W', \quad B_2 = W D_\theta W',
$$

where $W$ is a nonsingular $m \times m$ matrix, $D_\theta = \text{diag}(\theta_1, \ldots, \theta_m)$, and $\theta_1, \ldots, \theta_m$ are the characteristic roots of $B_2 B_1$. If we let

$$
Q(\alpha) = W D_\alpha W',
$$

where $D_\alpha = \text{diag}(\alpha_1, \ldots, \alpha_m)$, then condition (2.8) becomes

$$
W D_\theta W' \leq W D_\alpha W' \leq W W'.
$$

Because $W$ is nonsingular, condition (2.11) reduces to
(2.12) \[ \theta_i \leq \alpha_i \leq 1, \quad i = 1, \ldots, m. \]

Every choice of \( D_\alpha \) generates a matrix \( Q(\alpha) = MM' \), for which \( M = M(\alpha) \) can be determined. Although the solution for \( M \) is not unique, all solutions are rotations of one another.

In summary, given any factorization \( \Sigma_{12} = A_1 A_2' \), an extended class of Tucker solutions \( A_1^* = A_1 M, A_2^* = A_2 M' \) can be constructed that satisfies (2.2). Note that some flexibility still remains since \( \alpha_1, \ldots, \alpha_m \) are not fixed.

**Remark 2.** The class of matrices \( Q = MM' \) that define the modified Tucker solutions do not exhaust all the matrices for which the residuals \( \Sigma_{11} - A_1 A_1' \geq 0, \Sigma_{22} - A_2 A_2' \geq 0. \) This is so because every matrix \( Q \) satisfying (2.8) need not be expressible in the form (2.10). To obtain a complete class of solutions we require a characterization of matrices \( \bar{Q} \) satisfying \( D_\theta \leq \bar{Q} \leq I, \) where \( \bar{Q} = W^{-1} Q W'^{-1}. \)

**Remark 3.** Although the class of solutions is now extended, and \( Q = MM' \) can be chosen to satisfy our mathematical conditions, it is not clear how to choose \( M \) to yield psychological interpretation. One approach is to remove the freedom in \( M \) by specifying that certain elements in \( A_1 \) and \( A_2 \) be zero. This would lead to criteria similar to those used in factor analysis.

3. **Extension of Rao's Solution.** To obtain a factorization \( \Sigma_{12} = A_1 A_2' \), Rao (1965) uses the canonical correlation representation of Hotelling (1936):

\[
\Sigma_{11} = L_1 L_1', \quad \Sigma_{22} = L_2 L_2', \quad \Sigma_{12} = L_1 \begin{pmatrix} D_\nu & 0 \\ 0 & 0 \end{pmatrix} L_2'.
\]
where $L_1$ and $L_2$ are $p_1 \times p_1$ and $p_2 \times p_2$ nonsingular matrices, respectively, $D_\rho = \text{diag}(\rho_1, \ldots, \rho_r)$, and $\rho_1, \ldots, \rho_r$ (the canonical correlations) are the positive square roots of the characteristic roots of $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$. Since $r$ is the rank of $\Sigma_{12}$ and $r \leq p_1, p_2, 0 < \rho_i < 1$, $i = 1, \ldots, r$.

For definiteness and with no loss in generality suppose $p_1 \leq p_2$. Then the choice

$$
A_1 = L_1 \begin{bmatrix} L_1^2 \\ D_\rho \\ 0 \end{bmatrix}, \quad A_2 = L_2 \begin{bmatrix} D_\rho^2 \\ 0 \end{bmatrix}
$$

satisfies $\Sigma_{12} = A_1 A_2'$. Furthermore, the residual matrices

$$
\Sigma_{jj} - A_j A_j' = L_j L_j' - L_j \begin{bmatrix} D_\rho \quad 0 \\ 0 \quad 0 \end{bmatrix} L_j',
$$

$$
= L_j \begin{bmatrix} I - D_\rho & 0 \\ 0 & I \end{bmatrix} L_j' \geq 0, \quad j = 1, 2,
$$

by virtue of the fact that $0 < D_\rho < I$.

Following the approach of Section 2 we wish to find $Q = MM'$ to satisfy (2.8). But now by virtue of (3.2), (2.8) becomes

$$
D_\rho \leq Q \leq D_\rho^{-1}
$$

One class of solutions is obtained by letting $M = D_\psi = \text{diag}(\psi_1, \ldots, \psi_r)$, with $\rho_1 \leq \psi_i^2 \leq \rho_i^{-1}$, $i = 1, \ldots, r$. This is equivalent to choosing

$$
A_1 = L_1 \begin{bmatrix} D_\alpha \\ 0 \end{bmatrix}, \quad A_2 = L_2 \begin{bmatrix} D_{\beta} \\ 0 \end{bmatrix},
$$

11
where \( \alpha_1 \beta_1 = \rho_1 \), \( 0 \leq \alpha_1 \leq 1 \), \( 0 \leq \beta_1 \leq 1 \). This class of solutions is called extended Rao solutions.

4. Comparison of Solutions. It is known that \( AA' = BB' \) if and only if \( A = B\Gamma \), where \( \Gamma \) is orthogonal. The following lemma provides a similar result when \( A_1 A_2' = B_1 B_2' \).

**Lemma 3.** Let \( A_1 \) and \( B_1 \) be \( k \times m \) matrices, and let \( A_2 \) and \( B_2 \) be \( \ell \times m \) matrices, each of rank \( m \). Then

\[
(4.1) \quad A_1 A_2' = B_1 B_2'
\]

if and only if

\[
(4.2) \quad A_1 = B_1 H, \quad A_2' = H^{-1} B_2',
\]

where \( H \) is an \( m \times m \) nonsingular matrix.

**Proof.** That (4.2) implies (4.1) is immediate. To prove that (4.1) implies (4.2), note that the rank hypotheses imply that \( A_1 A_1' \) and \( A_2 A_2' \) are nonsingular. Hence,

\[
A_1 = (A_1 A_2') A_2 (A_2' A_2)^{-1} = B_1 (B_2' A_2) (A_2' A_2)^{-1} \equiv B_1 H_1,
\]

\[
A_2 = (A_2 A_1') A_1 (A_1' A_1)^{-1} = B_2 (B_2' A_1) (A_1' A_1)^{-1} \equiv B_2 H_2,
\]

where \( H_1 = B_2' A_2 (A_2' A_2)^{-1} \), \( H_2 = B_1 A_1 (A_1' A_1)^{-1} \). Since
\[ H'_2 H_1 = (A'_1 A_1)^{-1} A'_1 (B'_1 B'_2) A_2 (A'_2 A_2)^{-1} = (A'_1 A_1)^{-1} A'_1 (A_1 A'_2) A_2 (A'_2 A_2)^{-1} = I, \]

we have that \( H'_2 = H_1^{-1} \), which completes the proof. ||

Thus, if we start with either the Tucker or Rao solutions, one is obtained from the other by a nonsingular transformation. Because \( H \) is not orthogonal, each solution extracts a different amount of interbattery variance from the two batteries. By considering the extended Rao or Tucker solutions, we may alter the amount of variance accounted for according to the demands of the investigation.

5. The Choice of Solutions for Some Interbattery Models. Both the extended Tucker and Rao solutions satisfy the requirements (1.4). However, there still remains some flexibility within each class. We now consider some particular choices.

5.1. Parallel Batteries. If we have two parallel batteries in the sense that for each test in one battery there is a classically equivalent test in the other battery, then differences in scores between the two batteries are due to errors of measurement. In this case \( p_1 = p_2 \). Thus, we expect the estimates of \( A_1 \) and \( A_2 \) to be in close agreement, which suggests that we choose a factorization to

\[
\text{Min}_{A_1', A_2'} \text{tr}(A_1 - A_2')(A_1 - A_2)' ,
\]

subject to the conditions (1.4). This means that we can choose a nonsingular matrix \( H \) so as to
\[
\min_{H} \text{tr} \left( A_1 H - A_2 H^{-1} \right) \left( A_1 H - A_2 H^{-1} \right)^{\ast},
\]

(5.1) \[
= \min_{H} \{ \text{tr} \left[ A_1 (H H') A_1^\prime + A_2 (H H')^{-1} A_2^\prime \right] - 2 \text{tr} \Sigma_{12} \}
\]

subject to \( A_1 (H H') A_1^\prime \leq \Sigma_{11}, \) \( A_2 (H H') A_2^\prime \leq \Sigma_{22} \). Because \( A_1 \) and \( A_2 \) are of rank \( m \) and \( H H' > 0 \), the constraints are equivalent to

(5.2) \[
(A_2 \Sigma_{22}^{-1} A_2) \leq H H' \leq (A_1 \Sigma_{11}^{-1} A_1)^{-1}
\]

To simplify the extremal problem further, let

(5.3a) \[
(A_1 \Sigma_{11}^{-1} A_1)^{-1} = W W', \quad (A_2 \Sigma_{22}^{-1} A_2) = W D_\theta W',
\]

(5.3b) \[
H H' = W Q W'
\]

where \( W \) is non-singular, \( D_\theta = \text{diag}(\theta_1, \ldots, \theta_p) \) and \( \theta_1, \ldots, \theta_p \) are the characteristic roots of \( (A_2 \Sigma_{22}^{-1} A_2) (A_1 \Sigma_{11}^{-1} A_1) \), or equivalently, of \( \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \). That is, \( \theta_1, \ldots, \theta_p \) are the squared canonical correlations and are contained in \((0, 1)\). The extremal problem (5.1) is now equivalent to

(5.4) \[
\min_{Q} \text{tr} \left[ Q \left( W' A_1^\prime A_1 W \right) + Q^{-1} (W^{-1} A_2^\prime A_2 W') \right]
\]

subject to the condition

(5.5) \[
W D_\theta W' \leq W Q W' \leq W W',
\]

14
which is equivalent to

\[(5.6) \quad D_0 \leq Q \leq I.\]

A complete solution to this minimization problem is complicated. However, a partial solution can be obtained as follows. Let

\[(5.7) \quad U = W' A_1 A_1' W, \quad V = W^{-1} A_2 A_2' W'^{-1}\]

and \(Q = D_Q = \text{diag} (q_1, \ldots, q_p)\). Then (5.4) becomes

\[
\begin{align*}
\min_{0 \leq q_i \leq 1} & \quad \sum_{i=1}^{p} [q_i u_{ii} + q_i^{-1} v_{ii}]. \\
& \quad \text{for } i = 1, \ldots, m
\end{align*}
\]

Because of the convexity of \(z u + z^{-1} v\), \(0 < z, u, v\), the minimizer is given by

\[(5.8) \quad \hat{q}_i = \begin{cases} 
\max(\theta_i, \sqrt{v_{ii}/u_{ii}}), & \text{if } \sqrt{v_{ii}/u_{ii}} \leq 1, \\
1, & \text{if } \sqrt{v_{ii}/u_{ii}} > 1,
\end{cases} \quad i = 1, \ldots, m.
\]

In summary, the solution of the interbattery model is obtained as follows: (i) Find \(A_1\) and \(A_2\) using the Rao or Tucker procedures. (ii) Determine \(W, D_0\) from (5.3a). (iii) Compute \(U\) and \(V\) from (5.7) and \(\hat{q}_i\) from (5.8). (iv) Choose \(H\) as any factorization of \(HH' = WD_q W'\). (v) Then \(\bar{A}_1 = A_1 H\) and \(\bar{A}_2 = A_2 H'^{-1}\) is a solution.
5.2 One Battery Subsumed in Another. Suppose that one of the batteries, say battery 2, measures additional variables to those measured by battery 1. An example of this occurs when battery 2 is obtained by adding more tests to a parallel from of battery 1. (For an actual example see Nanda, 1967).

Let \( \Sigma_{11} \) and \( \Sigma_{22} \) represent the covariance matrices for the variables in batteries 1 and 2, respectively, and \( \Sigma_{12} \) the covariances of variables between batteries. If we have a factorization \( \Sigma_{12} = A_1 A'_2 \), we no longer expect \( A_1 \) and \( A_2 \) to be close. Thus, we may wish to minimize some composite of the residuals, e.g., a weighted sum of variances

\[
(i) \quad \lambda \, \text{tr} \left( \Sigma_{11} - A_1 A'_1 \right)^2 + (1 - \lambda) \, \text{tr} \left( \Sigma_{22} - A_2 A'_2 \right)^2,
\]

or a geometric mean of generalized variances

\[
(ii) \quad \left| \frac{\lambda}{\Sigma_{11} - A_1 A'_1} \right|^{\lambda} \left| \frac{(1 - \lambda)}{\Sigma_{22} - A_2 A'_2} \right|^{1 - \lambda}
\]

where \( 0 \leq \lambda \leq 1 \), and the residuals are positive semidefinite.

For the criterion (ii), define \( W, D_{\theta, U}, V, D_q \) from (5.3) and (5.7). Then (ii) becomes

\[
\begin{align*}
\text{Min} & \quad \left[ \lambda \log |I - D_q| + (1 - \lambda) \log |I - D_q^{-1} D_{\theta}| \right]_q \\
= & \quad \text{Min} \left[ \lambda \left( \Gamma^m(1 - q_i) + (1 - \lambda) \left( \begin{array}{c} \Gamma^m \Gamma^m \end{array} \right)(1 - \theta_i / q_i) \right) \right]_{0_i \leq q_i \leq 1} \quad , \\
& \quad i = 1, \ldots, m
\end{align*}
\]

which yields
\[ \hat{q}_i = \left[ -\theta_i (1-2\lambda) + \sqrt{\theta_i^2 (1-2\lambda)^2 + 4 \lambda (1-\lambda) \theta_i^2} \right] / 2\lambda, \quad i=1, \ldots, m. \]

A straightforward analysis shows that \( \theta_i \leq \hat{q}_i \leq 1, \quad i=1, \ldots, m. \)

The special case \( \lambda = \frac{1}{2} \) yields the simple result \( \hat{q}_i = \sqrt{\theta_i}, \quad i=1, \ldots, m. \)

In general, \( \lambda \) can be chosen to take account of the difference in the number of tests, e.g., \( \lambda = p_1/(p_1 + p_2) \).
References


