The Annals of Statistics
1986, Vol. 14, No. 2, 405–417

MAXIMUM LIKELIHOOD ESTIMATORS AND LIKELIHOOD RATIO CRITERIA IN MULTIVARIATE COMPONENTS OF VARIANCE[1]

BY BLAIR M. ANDERSON,[2] T. W. ANDERSON, AND INGRAM OLKIN

Stanford University

Maximum likelihood estimators are obtained for multivariate components of variance models under the condition that the effect covariance matrix is positive semidefinite with a maximum rank. The rank of the estimator is random. The estimation procedure leads to a likelihood ratio test that the rank of the effect matrix is not greater than a given number against the alternative that the rank is not greater than a larger specified number. Linear structural relationship models and some factor analytic models can be put into this framework.

1. Introduction. If the effects of factors or classes are random, the analysis of variance model is called the *components of variance model* or model II. When the effects and errors are normally distributed, the multivariate one-way model is described by the covariance matrices of the effects and of the errors and an overall mean vector. In the balanced case with replications a sufficient set of statistics consists of the between class vector sum of squares, the within class vector sum of squares, and the overall sample mean. Linear combinations of the vector sums of squares yield unbiased estimators of the two model covariance matrices, but the estimator of the effect covariance matrix is not necessarily positive semidefinite.

In this paper we find the maximum likelihood estimators under the condition that the covariance matrices are positive semidefinite, in fact, under the condition that the effect covariance matrix is positive semidefinite with a maximum rank. The rank of the estimator is random; it depends on the roots of a certain determinantal equation. The estimators depend on the corresponding vectors associated with a matrix equation. The estimation procedure leads to a likelihood ratio test of the null hypothesis that the rank of the effect matrix is not greater than a given number against the alternative hypothesis that the rank is not greater than a larger specified number. The usual asymptotic theory does not hold; except in special cases -2 times the logarithm of the likelihood ratio criterion is not a χ^2-distribution.

Linear structural relationship models can be put into this framework. The effect vectors can be considered as the random systematic parts. Linear

Received May 1985; revised July 1985.

[1]Supported in part by the National Science Foundation Grants DMS 82-19748 and DMS 84-11411.

[2]Current address is 833 N. Curson Avenue, Los Angeles, CA 90046.

AMS 1980 *subject classifications*. Primary 62H12; secondary 62J10.

Key words and phrases. Multivariate analysis of variance, linear structural models, factor analysis models.

combinations of the components of the systematic parts being constant is equivalent to those linear combinations of the covariance matrix of the systematic parts being zero. We also obtain maximum likelihood estimators of the linear structural relationships.

In Section 2 we consider the case of the error covariance matrix being proportional to the identity; that is, the components of the errors are identically and independently distributed. In this case we do not need replication. In Section 3 we treat the case of the error matrix being proportional to the identity with replications; although this case may not be directly applicable, it is relevant as a transition from Section 2 to Section 4. Finally, in Section 4 we study the case where the error covariance matrix is unrestricted and there are replications.

There is a history of results dealing with components of various models in which particular aspects have been treated. In this paper we attempt to bring together specific results in a coherent manner. Although some of the results are new, we have given special attention to an exposition of the field.

In an unpublished paper Anderson (1946) gave the results for the case of an unrestricted covariance matrix. In another unpublished paper Morris and Olkin (1964) independently obtained these results. Most of these were announced by Anderson (1984), where an extensive list of references is given. Theobald (1975) obtained the estimators of Section 2 (in slightly more generality) by the same method that we are using. More recently Schott and Saw (1984) have derived the maximum likelihood estimators and likelihood ratio criteria in Section 4 by a somewhat different method.

Klotz and Putter (1969) obtained maximum likelihood estimators in a different form when no rank condition is imposed. Amemiya (1985) has also treated this problem. Amemiya and Fuller (1984) have found modified maximum likelihood estimators and likelihood ratio criteria when the rank is specified exactly. Rao (1983) proposed estimation and testing procedures that are related to maximum likelihood.

2. MANOVA without replication. In the simplest case of MANOVA with random factors there is one observation per cell. Let the p-component observable random vector be

$$\mathbf{X}_\alpha = \boldsymbol{\mu} + \mathbf{V}_\alpha + \mathbf{U}_\alpha, \qquad \alpha = 1, \dots, n, \tag{2.1}$$

where $\boldsymbol{\mu}$ is a constant (unknown) vector, $\mathbf{V}_1, \dots, \mathbf{V}_n, \mathbf{U}_1, \dots, \mathbf{U}_n$ are independent unobservable random vectors with means $\mathbf{0}$ and covariance matrices

$$\mathscr{E}\mathbf{V}_\alpha\mathbf{V}_\alpha' = \boldsymbol{\Theta}, \qquad \mathscr{E}\mathbf{U}_\alpha\mathbf{U}_\alpha' = \sigma^2\mathbf{I}. \tag{2.2}$$

A vector $\mathbf{U}_\alpha$ is interpreted as composed of errors that are uncorrelated and have a common variance σ^2. (If $\mathscr{E}\mathbf{U}_\alpha\mathbf{U}_\alpha' = \sigma^2\boldsymbol{\Psi}_0$, where $\boldsymbol{\Psi}_0$ is known, the model can be transformed to replace $\boldsymbol{\Psi}_0$ by $\mathbf{I}$; see Theobald (1975).) The vector $\mathbf{V}_\alpha$ represents the effect of factors and characterizes the cell or class. The covariance matrix $\boldsymbol{\Theta}$ of rank m, $0 \le m \le p$, is not necessarily positive definite. The covariance matrix of the observed $\mathbf{X}_\alpha$ is

$$C(\mathbf{X}_\alpha) = \mathscr{E}(\mathbf{X}_\alpha - \boldsymbol{\mu})(\mathbf{X}_\alpha - \boldsymbol{\mu})' = \sigma^2\mathbf{I} + \boldsymbol{\Theta}. \tag{2.3}$$

The effect vector $\mathbf{V}_\alpha$ is said to satisfy *linear structural relationships* if there is a $q \times p$ matrix $\mathbf{B}$ of rank q such that $\mathbf{B}\mathbf{V}_\alpha = \mathbf{0}$ with probability 1. This implies that $\mathbf{B}\mathscr{E}\mathbf{V}_\alpha\mathbf{V}_\alpha' = \mathbf{0}$, that is,

$$\mathbf{B}\mathbf{\Theta} = \mathbf{0}. \tag{2.4}$$

The matrix $\mathbf{B}$ is not uniquely determined since it can be multiplied on the left by an arbitrary nonsingular matrix of order q. The rank of $\mathbf{B}$ can be taken to satisfy $m + q = p$. We shall obtain the maximum likelihood estimators of $\boldsymbol{\mu}$, σ^2, $\mathbf{\Theta}$, and $\mathbf{B}$ under the assumption that the joint distribution of the random vectors is normal.

Let the observations be $\mathbf{x}_1, \ldots, \mathbf{x}_n$. The sample mean vector $\bar{\mathbf{x}} = (1/n)\sum_{\alpha=1}^n \mathbf{x}_\alpha$ and covariance matrix

$$\mathbf{C} = \frac{1}{n} \sum_{\alpha=1}^{n} (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})' \tag{2.5}$$

are a sufficient set of statistics. The logarithm of the likelihood function L is

$$\begin{aligned} \log L = &-\frac{np}{2}\log 2\pi - \frac{n}{2}\log|\sigma^2\mathbf{I} + \mathbf{\Theta}| \\ &-\frac{n}{2}\operatorname{tr}(\sigma^2\mathbf{I} + \mathbf{\Theta})^{-1}\mathbf{C} - \frac{n}{2}(\bar{\mathbf{x}} - \boldsymbol{\mu})'(\sigma^2\mathbf{I} + \mathbf{\Theta})^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}). \end{aligned} \tag{2.6}$$

For any positive semidefinite $\mathbf{\Theta}$ and positive σ^2, $\log L$ is maximized with respect to $\boldsymbol{\mu}$ at $\hat{\boldsymbol{\mu}} = \bar{\mathbf{x}}$, so that the concentrated likelihood function is equivalent to

$$\log L^* = -\log|\sigma^2\mathbf{I} + \mathbf{\Theta}| - \operatorname{tr}(\sigma^2\mathbf{I} + \mathbf{\Theta})^{-1}\mathbf{C}. \tag{2.7}$$

The canonical form of $\mathbf{C}$ is

$$\mathbf{C} = \mathbf{W}\mathbf{D}_t\mathbf{W}', \tag{2.8}$$

where

$$\mathbf{D}_t = \operatorname{diag}(t_1, \ldots, t_p), \tag{2.9}$$

$$\mathbf{W} = (\mathbf{w}_1, \ldots, \mathbf{w}_p), \tag{2.10}$$

$t_1 > \cdots > t_p$ are the ordered characteristic roots of $\mathbf{C}$ (distinct and positive with probability 1), $\mathbf{w}_1, \ldots, \mathbf{w}_p$ are the corresponding characteristic vectors of $\mathbf{C}$ normalized by $\mathbf{w}_j'\mathbf{w}_j = 1$, and $\operatorname{diag}(t_1, \ldots, t_p)$ represents a diagonal matrix with $t_1, \ldots, t_p$ as the diagonal elements. The matrix $\mathbf{W}$ is orthogonal. The canonical form of $\sigma^2\mathbf{I} + \mathbf{\Theta}$ is

$$\sigma^2\mathbf{I} + \mathbf{\Theta} = \mathbf{\Gamma}\mathbf{D}_\delta\mathbf{\Gamma}', \tag{2.11}$$

where

$$\mathbf{D}_\delta = \operatorname{diag}(\delta_1, \ldots, \delta_p), \tag{2.12}$$

$$\mathbf{\Gamma} = (\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_p), \tag{2.13}$$

$\delta_1 \geq \cdots \geq \delta_p$ are the ordered characteristic roots of $\sigma^2\mathbf{I} + \mathbf{\Theta}$, $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_p$ are corresponding characteristic vectors normalized by $\boldsymbol{\gamma}_i'\boldsymbol{\gamma}_j = \delta_{ij}$, the Kronecker delta; so $\mathbf{\Gamma}$ is orthogonal. If $\mathbf{\Theta}$ is of rank m, $\delta_{m+1} = \cdots = \delta_p = \sigma^2 > 0$.

In terms of the canonical forms the concentrated likelihood is equivalent to

$$\begin{aligned}\log L^* &= -\log|\mathbf{D}_\delta| - \operatorname{tr}\Gamma\mathbf{D}_\delta^{-1}\Gamma'\mathbf{W}\mathbf{D}_t\mathbf{W}' \\ &= -\sum_{i=1}^{p}\log\delta_i - \operatorname{tr}\mathbf{D}_\delta^{-1}(\Gamma'\mathbf{W})\mathbf{D}_t(\Gamma'\mathbf{W})',\end{aligned} \tag{2.14}$$

which is to be maximized with respect to orthogonal Γ and diagonal $\mathbf{D}_\delta$ subject to $\delta_1 \geq \cdots \geq \delta_m > \delta_{m+1} = \cdots = \delta_p > 0$. We use the following theorem of von Neumann (1937).

THEOREM (von Neumann). *For* $\mathbf{Q}$ *orthogonal and* $\mathbf{D}_\delta$ *and* $\mathbf{D}_t$ *diagonal* ($\delta_1 \geq \cdots \geq \delta_p > 0$, $t_1 \geq \cdots \geq t_p > 0$)

$$\min_{\mathbf{Q}} \operatorname{tr}\mathbf{D}_\delta^{-1}\mathbf{Q}\mathbf{D}_t\mathbf{Q}' = \operatorname{tr}\mathbf{D}_\delta^{-1}\mathbf{D}_t, \tag{2.15}$$

and a minimizing value of $\mathbf{Q}$ *is* $\mathbf{Q} = \mathbf{I}$.

REMARK. For any multiplicities of the δ's and t's the set of minimizing $\mathbf{Q}$'s is found from $\operatorname{tr}\mathbf{D}_\delta^{-1}\mathbf{Q}\mathbf{D}_t\mathbf{Q}' = \operatorname{tr}\mathbf{D}_\delta^{-1}\mathbf{D}_t$.

The maximum of (2.14) with respect to orthogonal $\Gamma'\mathbf{W}$ (or orthogonal Γ) is

$$\begin{aligned}-\sum_{i=1}^{p}\log\delta_i - \operatorname{tr}\mathbf{D}_\delta^{-1}\mathbf{D}_t &= -\sum_{i=1}^{p}\left(\log\delta_i + \frac{t_i}{\delta_i}\right) \\ &= -\sum_{i=1}^{m}\left(\log\delta_i + \frac{t_i}{\delta_i}\right) - \left(q\log\sigma^2 + \frac{1}{\sigma^2}\sum_{i=m+1}^{p}t_i\right).\end{aligned} \tag{2.16}$$

The maximum of (2.16) with respect to δ_i is at $\hat{\delta}_i = t_i$, $i = 1,\dots,m$, and with respect to σ^2 is at $\hat{\sigma}^2 = \sum_{i=m+1}^{p} t_i/q$; then $\hat{\delta}_{m+1} = \cdots = \hat{\delta}_p = \hat{\sigma}^2$.

Let

$$\mathbf{D}_t = \begin{bmatrix}\dot{\mathbf{D}}_t & \mathbf{0} \\ \mathbf{0} & \ddot{\mathbf{D}}_t\end{bmatrix}, \qquad \mathbf{W} = (\mathbf{W}_1 \quad \mathbf{W}_2), \tag{2.17}$$

where $\dot{\mathbf{D}}_t$ is $m \times m$ and $\mathbf{W}_1$ has m columns. Then a maximizing $\mathbf{D}_\delta$ and $\mathbf{Q} = \Gamma'\mathbf{W}$ are

$$\hat{\mathbf{D}}_\delta = \begin{bmatrix}\dot{\mathbf{D}}_t & \mathbf{0} \\ \mathbf{0} & \hat{\sigma}^2\mathbf{I}_q\end{bmatrix}, \qquad \mathbf{Q} = \begin{bmatrix}\mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2\end{bmatrix}, \tag{2.18}$$

where $\mathbf{Q}_2$ is any orthogonal matrix of order q. Then

$$\begin{aligned}\hat{\Gamma} = \mathbf{W}\mathbf{Q}' &= (\mathbf{W}_1 \quad \mathbf{W}_2)\begin{bmatrix}\mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2'\end{bmatrix} \\ &= (\mathbf{W}_1 \quad \mathbf{W}_2\mathbf{Q}_2'),\end{aligned} \tag{2.19}$$

and the maximum likelihood estimator of $\sigma^2\mathbf{I} + \boldsymbol{\Theta}$ is

$$\hat{\sigma}^2\mathbf{I}_p + \hat{\boldsymbol{\Theta}} = \hat{\boldsymbol{\Gamma}}\hat{\mathbf{D}}_\delta\hat{\boldsymbol{\Gamma}}'$$

$$= (\mathbf{W}_1 \quad \mathbf{W}_2\mathbf{Q}_2')\begin{bmatrix} \dot{\mathbf{D}}_t & \mathbf{0} \\ \mathbf{0} & \hat{\sigma}^2\mathbf{I}_q \end{bmatrix}\begin{bmatrix} \mathbf{W}_1' \\ \mathbf{Q}_2\mathbf{W}_2' \end{bmatrix} \tag{2.20}$$

$$= \mathbf{W}_1\dot{\mathbf{D}}_t\mathbf{W}_1' + \hat{\sigma}^2\mathbf{W}_2\mathbf{W}_2'.$$

It follows that the maximum likelihood estimator of $\boldsymbol{\Theta}$ is

$$\begin{aligned}\hat{\boldsymbol{\Theta}} &= \mathbf{W}_1\dot{\mathbf{D}}_t\mathbf{W}_1' + \hat{\sigma}^2\mathbf{W}_2\mathbf{W}_2' - \hat{\sigma}^2\mathbf{I}_p \\ &= \mathbf{W}_1(\dot{\mathbf{D}}_t - \hat{\sigma}^2\mathbf{I}_m)\mathbf{W}_1',\end{aligned} \tag{2.21}$$

which is positive semidefinite of rank m.

Since

$$\begin{aligned}\operatorname{tr}\left(\hat{\sigma}^2\mathbf{I}_p + \hat{\boldsymbol{\Theta}}\right)^{-1}\mathbf{C} &= \operatorname{tr}\left(\mathbf{W}\begin{bmatrix} \dot{\mathbf{D}}_t & \mathbf{0} \\ \mathbf{0} & \hat{\sigma}\mathbf{I}_q \end{bmatrix}\mathbf{W}'\right)^{-1}\mathbf{W}\mathbf{D}_t\mathbf{W}' \\ &= \operatorname{tr}\begin{bmatrix} \dot{\mathbf{D}}_t^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\hat{\sigma}^2}\mathbf{I}_q \end{bmatrix}\begin{bmatrix} \dot{\mathbf{D}}_t & \mathbf{0} \\ \mathbf{0} & \ddot{\mathbf{D}}_t \end{bmatrix} \\ &= \operatorname{tr}\begin{bmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & \frac{1}{\hat{\sigma}^2}\ddot{\mathbf{D}}_t \end{bmatrix} \\ &= p,\end{aligned} \tag{2.22}$$

the maximized value of the likelihood function is

$$L(m) = \left[(2\pi)^{np/2}\prod_{i=1}^{m} t_i^{n/2}\left(\sum_{i=m+1}^{p} t_i/q\right)^{nq/2} e^{np/2}\right]^{-1}. \tag{2.23}$$

The likelihood ratio criterion for testing the null hypothesis H_0: $m = m_0$ against the alternative $m_0 < m \leq m_1$, where m_0 and m_1 are specified integers between 0 and p, is

$$\begin{aligned}\frac{L(m_0)}{L(m_1)} &= \frac{\prod_{i=1}^{m_1} t_i^{n/2}\left(\sum_{i=m_1+1}^{p} t_i/q_1\right)^{nq_1/2}}{\prod_{i=1}^{m_0} t_i^{n/2}\left(\sum_{i=m_0+1}^{p} t_i/q_0\right)^{nq_0/2}} \\ &= \prod_{i=m_0+1}^{m_1} t_i^{n/2}\frac{\left(\sum_{i=m_1+1}^{p} t_i/q_1\right)^{nq_1/2}}{\left(\sum_{i=m_0+1}^{p} t_i/q_0\right)^{nq_0/2}},\end{aligned} \tag{2.24}$$

where $q_0 = p - m_0$ and $q_1 = p - m_1$.

A special case is $m_1 = p$, that is, the alternative is that the rank of Θ is greater than m_0. The likelihood ratio criterion is

$$\frac{L(m_0)}{L(p)} = \left[\frac{\left(\prod_{i=m_0+1}^{p} t_i\right)^{1/q_0}}{\sum_{i=m_0+1}^{p} t_i/q_0}\right]^{nq_0/2}. \tag{2.25}$$

This is the $\frac{1}{2}nq_0$ power of the ratio of the geometric mean of the q_0 smallest roots to the arithmetic mean of these roots.

3. MANOVA with independent errors and replications. We now consider the balanced MANOVA with k replications in each cell:

$$\mathbf{X}_{\alpha j} = \boldsymbol{\mu} + \mathbf{V}_\alpha + \mathbf{U}_{\alpha j}, \qquad j = 1, \ldots, k, \ \alpha = 1, \ldots, n. \tag{3.1}$$

The unobservable random vectors $\mathbf{V}_1, \ldots, \mathbf{V}_n, \mathbf{U}_{11}, \ldots, \mathbf{U}_{nk}$ are independent with means $\mathbf{0}$ and covariances

$$\mathscr{E}\mathbf{V}_\alpha \mathbf{V}_\alpha' = \Theta, \qquad \mathscr{E}\mathbf{U}_{\alpha j}\mathbf{U}_{\alpha j}' = \Psi, \qquad j = 1, \ldots, k, \ \alpha = 1, \ldots, n. \tag{3.2}$$

We assume that the rank of Θ is less than or equal to m. The covariance matrix of $\mathbf{X}_\alpha^* = (\mathbf{X}_{\alpha 1}', \ldots, \mathbf{X}_{\alpha k}')'$ is

$$\mathscr{E}(\mathbf{X}_\alpha^* - \mathbf{e} \otimes \boldsymbol{\mu})(\mathbf{X}_\alpha^* - \mathbf{e} \otimes \boldsymbol{\mu})' = \mathbf{I} \otimes \Psi + \mathbf{e}\mathbf{e}' \otimes \Theta, \tag{3.3}$$

where $\mathbf{e} = (1, 1, \ldots, 1)'$, $\mathbf{X}_1^*, \ldots, \mathbf{X}_n^*$ are independently distributed, and $\otimes$ denotes the Kronecker product. The inverse of (3.3) is

$$\left(\mathbf{I} - \frac{1}{k}\mathbf{e}\mathbf{e}'\right) \otimes \Psi^{-1} + \frac{1}{k}\mathbf{e}\mathbf{e}' \otimes (\Psi + k\Theta)^{-1}. \tag{3.4}$$

The determinant of (3.3) is $|\Psi|^{k-1}|\Psi + k\Theta|$. Note that the covariance matrix of $\overline{\mathbf{X}}_\alpha = (1/k)\sum_{j=1}^{k} \mathbf{X}_{\alpha j}$ is $(1/k)(\Psi + k\Theta)$.

If $\mathbf{x}_{11}, \ldots, \mathbf{x}_{1k}, \mathbf{x}_{21}, \ldots, \mathbf{x}_{nk}$ are the observation vectors, the logarithm of the likelihood function is

$$\begin{aligned} \log L = & -\frac{pnk}{2}\log 2\pi - \frac{n(k-1)}{2}\log|\Psi| - \frac{n}{2}\log|\Psi + k\Theta| \\ & -\frac{1}{2}\sum_{\alpha=1}^{n} (\mathbf{x}_\alpha^* - \mathbf{e} \otimes \boldsymbol{\mu})'\left[\left(\mathbf{I} - \frac{1}{k}\mathbf{e}\mathbf{e}'\right) \otimes \Psi^{-1}\right. \\ & \left. + \frac{1}{k}\mathbf{e}\mathbf{e}' \otimes (\Psi + k\Theta)^{-1}\right](\mathbf{x}_\alpha^* - \mathbf{e} \otimes \boldsymbol{\mu}) \\ = & -\frac{pnk}{2}\log 2\pi - \frac{n(k-1)}{2}\log|\Psi| - \frac{n}{2}\log|\Psi + k\Theta| \\ & -\frac{1}{2}\left[\operatorname{tr}\mathbf{G}\Psi^{-1} + \operatorname{tr}\mathbf{H}(\Psi + k\Theta)^{-1} + nk(\overline{\mathbf{x}} - \boldsymbol{\mu})'(\Psi + k\Theta)^{-1}(\overline{\mathbf{x}} - \boldsymbol{\mu})\right], \end{aligned} \tag{3.5}$$

where

$$\mathbf{H} = k \sum_{\alpha=1}^{n} (\bar{\mathbf{x}}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})', \tag{3.6}$$

$$\mathbf{G} = \sum_{\alpha=1}^{n} \sum_{j=1}^{k} (\mathbf{x}_{\alpha j} - \bar{\mathbf{x}}_\alpha)(\mathbf{x}_{\alpha j} - \bar{\mathbf{x}}_\alpha)', \tag{3.7}$$

$\bar{\mathbf{x}}_\alpha = (1/k)\sum_{n=1}^{k} \mathbf{x}_{\alpha j}$, and $\bar{\mathbf{x}} = (1/n)\sum_{\alpha=1}^{n} \bar{\mathbf{x}}_\alpha$. A sufficient set of statistics consists of $\mathbf{H}$, $\mathbf{G}$, and $\bar{\mathbf{x}}$. The maximum of $\log L$ with respect to $\boldsymbol{\mu}$ is at $\hat{\boldsymbol{\mu}} = \bar{\mathbf{x}}$.

First we consider the case of $\boldsymbol{\Psi} = \sigma^2 \mathbf{I}$. Let $(1/n)\mathbf{H} = \mathbf{W}\mathbf{D}_t\mathbf{W}'$ and $\sigma^2\mathbf{I} + k\boldsymbol{\Theta} = \boldsymbol{\Gamma}\mathbf{D}_\delta\boldsymbol{\Gamma}'$, where again $\mathbf{W}$ and $\boldsymbol{\Gamma}$ are orthogonal and $\mathbf{D}_t$ and $\mathbf{D}_\delta$ are diagonal. Then the concentrated likelihood is equivalent to

$$\begin{aligned} &-pn(k-1)\log\sigma^2 - n\sum_{i=1}^{p} \log\delta_i \\ &-\frac{1}{\sigma^2}\operatorname{tr}\mathbf{G} - n\operatorname{tr}\mathbf{D}_\delta^{-1}(\boldsymbol{\Gamma}'\mathbf{W})\mathbf{D}_t(\boldsymbol{\Gamma}'\mathbf{W})'. \end{aligned} \tag{3.8}$$

The maximum of (3.8) with respect to orthogonal $\boldsymbol{\Gamma}'\mathbf{W}$ is

$$\begin{aligned} &-[pn(k-1) + nq]\log\sigma^2 - n\sum_{i=1}^{m} \log\delta_i \\ &-\frac{1}{\sigma^2}\operatorname{tr}\mathbf{G} - n\sum_{i=1}^{m} \frac{t_i}{\delta_i} - \frac{n}{\sigma^2}\sum_{i=m+1}^{p} t_i. \end{aligned} \tag{3.9}$$

We want to maximize (3.9) with respect to $\delta_1 \geq \cdots \geq \delta_m$, and σ^2 subject to $\delta_j \geq \sigma^2$, $j = 1, \ldots, m$. Note that the concentrated likelihood function is a strictly concave function of $1/\sigma^2$ and $1/\delta_j$, $j = 1, \ldots, m$, and hence the maximum is unique. The derivatives of (3.9) with respect to $\delta_1, \ldots, \delta_m$, and σ^2 are

$$-\frac{n}{\delta_j} + \frac{nt_j}{\delta_j^2}, \qquad j = 1, \ldots, m, \tag{3.10}$$

$$-\frac{pn(k-1) + nq}{\sigma^2} + \frac{\operatorname{tr}\mathbf{G} + n\sum_{i=m+1}^{p} t_i}{(\sigma^2)^2}. \tag{3.11}$$

Let $a = pn(k-1) + nq$ and $A = \operatorname{tr}\mathbf{G} + n\sum_{i=m+1}^{p} t_i$. Let $m^* = m$ if $A/a \leq t_m$; otherwise let m^* be such that

$$\begin{aligned} t_{m^*+1} &+ \frac{n}{a}\left[(m - m^* - 1)t_{m^*+1} - t_{m^*+2} - \cdots - t_m\right] \leq \frac{A}{a} \\ &< t_{m^*} + \frac{n}{a}\left[(m - m^*)t_{m^*} - t_{m^*+1} - \cdots - t_m\right]; \end{aligned} \tag{3.12}$$

that is, if

$$[pn(k-1)+nq^*]t_{m^*+1} \le \operatorname{tr}\mathbf{G} + n\sum_{i=m^*+1}^{p} t_i < [pn(k-1)+nq^*]t_{m^*}, \tag{3.13}$$

where $q^* = p - m^*$. Then $\hat{\delta}_i = t_i$, $i = 1,\dots, m^*$, and

$$\hat{\sigma}^{*2} = \frac{\operatorname{tr}\mathbf{G} + n\sum_{i=m^*+1}^{p} t_i}{pn(k-1)+nq^*}. \tag{3.14}$$

Let

$$\dot{\mathbf{D}}_t^* = \operatorname{diag}(t_1,\dots,t_{m^*}), \qquad \ddot{\mathbf{D}}_t^* = \operatorname{diag}(t_{m^*+1},\dots,t_p), \tag{3.15}$$

$$\mathbf{W} = (\mathbf{W}_1^* \quad \mathbf{W}_2^*), \tag{3.16}$$

where $\mathbf{W}_1^*$ has m^* columns. The maximizing $\mathbf{D}_\delta$ and $\mathbf{Q} = \mathbf{\Gamma}'\mathbf{W}$ are

$$\hat{\mathbf{D}}_\delta = \begin{bmatrix} \dot{\mathbf{D}}_t^* & \mathbf{0} \\ \mathbf{0} & \hat{\sigma}^{*2}\mathbf{I}_{q^*} \end{bmatrix}, \qquad \mathbf{Q} = \begin{bmatrix} \mathbf{I}_{m^*} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2^* \end{bmatrix}, \tag{3.17}$$

where $\mathbf{Q}_2^*$ is any orthogonal matrix of order q^*. Then

$$\hat{\mathbf{\Gamma}} = \mathbf{W}\mathbf{Q}' = (\mathbf{W}_1^* \quad \mathbf{W}_2^*\mathbf{Q}_2^{*\prime}), \tag{3.18}$$

and the maximum likelihood estimator of $\sigma^2\mathbf{I} + k\mathbf{\Theta}$ is

$$\hat{\sigma}^{*2}\mathbf{I}_p + k\hat{\mathbf{\Theta}} = \mathbf{W}_1^*\dot{\mathbf{D}}_t^*\mathbf{W}_1^{*\prime} + \hat{\sigma}^{*2}\mathbf{W}_2^*\mathbf{W}_2^{*\prime}. \tag{3.19}$$

The maximum likelihood estimator of $\hat{\mathbf{\Theta}}$ is

$$\begin{aligned} \hat{\mathbf{\Theta}} &= \frac{1}{k}\left(\mathbf{W}_1^*\dot{\mathbf{D}}_t^*\mathbf{W}_1^{*\prime} + \hat{\sigma}^{*2}\mathbf{W}_2^*\mathbf{W}_2^{*\prime} - \hat{\sigma}^{*2}\mathbf{I}_p\right) \\ &= \frac{1}{k}\mathbf{W}_1^*\left(\dot{\mathbf{D}}_t^* - \hat{\sigma}^{*2}\mathbf{I}_{m^*}\right)\mathbf{W}_1^{*\prime}, \end{aligned} \tag{3.20}$$

which is positive semidefinite of rank m^*. The maximized value of the likelihood function is

$$L(m^*) = \left[(2\pi)^{pnk/2}\prod_{i=1}^{m^*} t_i^{n/2}(\hat{\sigma}^{*2})^{[pn(k-1)+nq^*]/2}e^{npk/2}\right]^{-1}. \tag{3.21}$$

The likelihood ratio criterion for testing the null hypothesis H_0: $m = m_0$ against the alternative $m_0 < m \le m_1$ is

$$\frac{L(m_0^*)}{L(m_1^*)} = \prod_{i=m_0^*+1}^{m_1^*} t_i^{n/2}\frac{\left(\hat{\sigma}_1^{*2}\right)^{n[p(k-1)+q_1^*]/2}}{\left(\hat{\sigma}_0^{*2}\right)^{n[p(k-1)+q_0^*]/2}}, \tag{3.22}$$

where $\hat{\sigma}_0^{*2}$ and $\hat{\sigma}_1^{*2}$ are given by (3.14) for $m = m_0^*$, $q = q_0^*$ and $m = m_1^*$, $q = q_1^*$, respectively, for $m_0^* < m_1^*$ and is 1 if $m_0^* = m_1^*$.

4. MANOVA with replications. We now consider the model (3.1) with $\mathscr{E}\mathbf{U}_{\alpha j}\mathbf{U}'_{\alpha j} = \mathbf{\Psi}$ unrestricted. Let the roots of

$$\left|\frac{1}{n}\mathbf{H} - d\frac{1}{n(k-1)}\mathbf{G}\right| = 0 \tag{4.1}$$

be $d_1 > \cdots > d_p > 0$; the roots are distinct and positive with probability 1. Define $\mathbf{Z}$ by

$$\mathbf{H} = n\mathbf{Z}\mathbf{D}_d\mathbf{Z}', \qquad \mathbf{G} = n(k-1)\mathbf{Z}\mathbf{Z}', \tag{4.2}$$

where $\mathbf{D}_d = \mathrm{diag}(d_1, \dots, d_p)$. Then

$$\frac{1}{n}\mathbf{H} - \frac{1}{n(k-1)}\mathbf{G} = \mathbf{Z}\mathbf{D}_d\mathbf{Z}' - \mathbf{Z}\mathbf{Z}' = \mathbf{Z}(\mathbf{D}_d - \mathbf{I}_p)\mathbf{Z}' \tag{4.3}$$

is an estimator of $k\mathbf{\Theta}$.

Let the roots of

$$|(\mathbf{\Psi} + k\mathbf{\Theta}) - \delta\mathbf{\Psi}| = 0 \tag{4.4}$$

be $\delta_1 \geq \cdots \geq \delta_m > \delta_{m+1} = \cdots = \delta_p = 1$. Let diagonal $\mathbf{D}_\delta$ and nonsingular $\mathbf{\Gamma}$ be defined by

$$\mathbf{\Psi} + k\mathbf{\Theta} = \mathbf{\Gamma}\mathbf{D}_\delta\mathbf{\Gamma}', \qquad \mathbf{\Psi} = \mathbf{\Gamma}\mathbf{\Gamma}'. \tag{4.5}$$

Then the log likelihood function concentrated with respect to $\hat{\boldsymbol{\mu}} = \bar{\mathbf{x}}$ is $-\frac{1}{2}npk\log 2\pi$ plus

$$\begin{aligned}
&-nk\log|\mathbf{\Gamma}| - \frac{n}{2}\log|\mathbf{D}_\delta| - \frac{n(k-1)}{2}\mathrm{tr}(\mathbf{\Gamma}\mathbf{\Gamma}')^{-1}\mathbf{Z}\mathbf{Z}' - \frac{n}{2}\mathrm{tr}(\mathbf{\Gamma}\mathbf{D}_\delta\mathbf{\Gamma}')^{-1}\mathbf{Z}\mathbf{D}_d\mathbf{Z}' \\
&\quad = -nk\log|\mathbf{\Gamma}| - \frac{n}{2}\log|\mathbf{D}_\delta| \\
&\qquad - \frac{n(k-1)}{2}\mathrm{tr}\,\mathbf{\Gamma}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{\Gamma}'^{-1} - \frac{n}{2}\mathrm{tr}\,\mathbf{D}_\delta^{-1}\mathbf{\Gamma}^{-1}\mathbf{Z}\mathbf{D}_d\mathbf{Z}'\mathbf{\Gamma}'^{-1} \qquad (4.6)\\
&\quad = -nk\log|\mathbf{\Gamma}| - \frac{n}{2}\log|\mathbf{D}_\delta| - \frac{n(k-1)}{2}\mathrm{tr}\big(\mathbf{\Gamma}^{-1}\mathbf{Z}\mathbf{D}_d^{1/2}\big)\mathbf{D}_d^{-1}\big(\mathbf{\Gamma}^{-1}\mathbf{Z}\mathbf{D}_d^{1/2}\big)' \\
&\qquad - \frac{n}{2}\mathrm{tr}\,\mathbf{D}_\delta^{-1}\big(\mathbf{\Gamma}^{-1}\mathbf{Z}\mathbf{D}_d^{1/2}\big)\big(\mathbf{\Gamma}^{-1}\mathbf{Z}\mathbf{D}_d^{1/2}\big)',
\end{aligned}$$

which is to be maximized with respect to $\mathbf{D}_\delta$ and $\mathbf{\Gamma}$. We use the singular value decomposition:

$$\mathbf{\Gamma}^{-1}\mathbf{Z}\mathbf{D}_d^{1/2} = \mathbf{P}\mathbf{D}_r\mathbf{Q}, \tag{4.7}$$

where $\mathbf{P}$ and $\mathbf{Q}$ are orthogonal, and $\mathbf{D}_r$ is diagonal with $r_1 \geq r_2 \geq \cdots \geq r_p > 0$.

Then

$$\Gamma^{-1} = \mathbf{P}\mathbf{D}_r\mathbf{Q}\mathbf{D}_d^{-1/2}\mathbf{Z}^{-1}, \tag{4.8}$$

$$|\Gamma|^{-1} = |\Gamma^{-1}| = |\mathbf{D}_r| \cdot |\mathbf{Z}|^{-1} \cdot |\mathbf{D}_d|^{-1/2}. \tag{4.9}$$

The concentrated log likelihood function to be maximized with respect to **P**, **Q**, $\mathbf{D}_r$, and $\mathbf{D}_\delta$ is a function of **Z** and $\mathbf{D}_d$ plus

$$\begin{aligned} &nk \log|\mathbf{D}_r| - \frac{n}{2}\log|\mathbf{D}_\delta| - \frac{n(k-1)}{2}\operatorname{tr}\mathbf{P}\mathbf{D}_r\mathbf{Q}\mathbf{D}_d^{-1}(\mathbf{P}\mathbf{D}_r\mathbf{Q})' \\ &\quad - \frac{n}{2}\operatorname{tr}\mathbf{D}_\delta^{-1}\mathbf{P}\mathbf{D}_r\mathbf{Q}(\mathbf{P}\mathbf{D}_r\mathbf{Q})' \\ &= nk\log|\mathbf{D}_r| - \frac{n}{2}\log|\mathbf{D}_\delta| - \frac{n(k-1)}{2}\operatorname{tr}\mathbf{P}\mathbf{D}_r\mathbf{Q}\mathbf{D}_d^{-1}\mathbf{Q}'\mathbf{D}_r\mathbf{P}' \\ &\quad - \frac{n}{2}\operatorname{tr}\mathbf{D}_\delta^{-1}\mathbf{P}\mathbf{D}_r\mathbf{Q}\mathbf{Q}'\mathbf{D}_r\mathbf{P}' \\ &= nk\log|\mathbf{D}_r| - \frac{n}{2}\log|\mathbf{D}_\delta| - \frac{n(k-1)}{2}\operatorname{tr}\mathbf{D}_r^2\mathbf{Q}\mathbf{D}_d^{-1}\mathbf{Q}' \\ &\quad - \frac{n}{2}\operatorname{tr}\mathbf{D}_\delta^{-1}\mathbf{P}\mathbf{D}_r^2\mathbf{P}'. \end{aligned} \tag{4.10}$$

Von Neumann's theorem shows that the maximum of (4.10) with respect to **P** and **Q** is

$$\begin{aligned} &nk\log|\mathbf{D}_r| - \frac{n}{2}\log|\mathbf{D}_\delta| - \frac{n(k-1)}{2}\operatorname{tr}\mathbf{D}_r^2\mathbf{D}_d^{-1} - \frac{n}{2}\operatorname{tr}\mathbf{D}_\delta^{-1}\mathbf{D}_r^2 \\ &= \frac{n}{2}\sum_{i=1}^{p}\left\{k\log r_i^2 - \log\delta_i - (k-1)\frac{r_i^2}{d_i} - \frac{r_i^2}{\delta_i}\right\} \\ &= \frac{n}{2}\sum_{i=1}^{p}\left\{k\log r_i^2 - r_i^2\left[\frac{k-1}{d_i} + \frac{1}{\delta_i}\right] - \log\delta_i\right\}. \end{aligned} \tag{4.11}$$

Since

$$\max_x\left[a\log x - bx\right] = a\log a - a\log b - a \tag{4.12}$$

and occurs at $x = a/b$, the maximum of (4.11) with respect to $r_1, \dots, r_p$ is

$$-\frac{n}{2}\sum_{i=1}^{p}\left\{\log\delta_i + k\log\left(\frac{k-1}{d_i} + \frac{1}{\delta_i}\right) + k - k\log k\right\}, \tag{4.13}$$

and occurs at $r_i^2 = k[(k-1)/d_i + 1/\delta_i]^{-1}$, $i = 1, \dots, p$. The maximum of (4.13) with respect to $\delta_1, \dots, \delta_p$ over the region $\delta_1 \geq \cdots \geq \delta_m > \delta_{m+1} = \cdots = \delta_p = 1$ occurs at

$$\begin{aligned} &\left.\begin{aligned} \hat{\delta}_i &= d_i, \quad \text{if } d_i > 1 \\ \hat{\delta}_i &= 1, \quad \text{if } d_i \leq 1 \end{aligned}\right\} \quad \text{for } i = 1, \dots, m, \\ &\hat{\delta}_i = 1, \qquad\qquad\qquad\quad \text{for } i = m+1, \dots, p. \end{aligned} \tag{4.14}$$

Let p^* be the number of $d_i > 1$, $m^* = \min(m, p)$, and $q^* = p - m^*$. Then $\hat{\delta}_i = d_i$, $i = 1, \ldots, m^*$, and $\hat{\delta}_i = 1$, $i = m^* + 1, \ldots, p$. The maximum of (4.13) is

$$\begin{aligned}(4.15)\qquad &\frac{n(p-m^*)k}{2}\log k - \frac{npk}{2} + \frac{n(k-1)}{2}\sum_{i=1}^{m^*}\log d_i + \frac{nk}{2}\sum_{i=m^*+1}^{p}\log d_i \\ &- \frac{nk}{2}\sum_{i=m^*+1}^{p}\log(k-1+d_i).\end{aligned}$$

When we go back to (4.10) we see that a maximizing $\mathbf{Q}$ is $\hat{\mathbf{Q}} = \mathbf{I}$ (unique except for multiplication of each diagonal element by -1) and a maximizing $\mathbf{P}$ is

$$(4.16)\qquad \mathbf{P} = \begin{bmatrix} \mathbf{I}_{m^*} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 \end{bmatrix},$$

where $\mathbf{P}_2$ is an arbitrary orthogonal matrix of order $q^* = p - m^*$. Let

$$(4.17)\qquad \dot{\mathbf{D}}_d^* = \operatorname{diag}(d_1, \ldots, d_{m^*}), \qquad \ddot{\mathbf{D}}_d^* = \operatorname{diag}(d_{m^*+1}, \ldots, d_p),$$

$$(4.18)\qquad \mathbf{Z} = (\mathbf{Z}_1^* \quad \mathbf{Z}_2^*),$$

where $\mathbf{Z}_1^*$ has m^* columns. Then

$$(4.19)\qquad \hat{\mathbf{D}}_\delta = \begin{bmatrix} \dot{\mathbf{D}}_d^* & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{q^*} \end{bmatrix}, \qquad \mathbf{D}_d = \begin{bmatrix} \dot{\mathbf{D}}_d^* & \mathbf{0} \\ \mathbf{0} & \ddot{\mathbf{D}}_d^* \end{bmatrix}.$$

In these terms

$$(4.20)\qquad \hat{\mathbf{D}}_r^2 = k\left[(k-1)\mathbf{D}_d^{-1} + \hat{\mathbf{D}}_\delta^{-1}\right]^{-1}.$$

From (4.8) we obtain

$$(4.21)\qquad \hat{\mathbf{\Gamma}} = \mathbf{Z}\mathbf{D}_d^{1/2}\hat{\mathbf{Q}}'\hat{\mathbf{D}}_r^{-1}\hat{\mathbf{P}}',$$

from which we obtain

$$\begin{aligned}(4.22)\qquad \hat{\mathbf{\Psi}} &= \mathbf{Z}\mathbf{D}_d^{1/2}\hat{\mathbf{D}}_r^{-2}\mathbf{D}_d^{1/2}\mathbf{Z}' \\ &= \frac{1}{k}(\mathbf{Z}_1^* \quad \mathbf{Z}_2^*)\begin{pmatrix} k\mathbf{I}_{m^*} & \mathbf{0} \\ \mathbf{0} & (k-1)\mathbf{I}_{q^*} + \ddot{\mathbf{D}}_d^* \end{pmatrix}\begin{pmatrix} \mathbf{Z}_1^{*\prime} \\ \mathbf{Z}_2^{*\prime} \end{pmatrix} \\ &= \mathbf{Z}_1^*\mathbf{Z}_1^{*\prime} + \frac{k-1}{k}\mathbf{Z}_2^*\mathbf{Z}_2^{*\prime} + \frac{1}{k}\mathbf{Z}_2^*\ddot{\mathbf{D}}_d^*\mathbf{Z}_2^{*\prime},\end{aligned}$$

$$\begin{aligned}(4.23)\qquad \hat{\mathbf{\Psi}} + k\hat{\mathbf{\Theta}} &= \hat{\mathbf{\Gamma}}\hat{\mathbf{D}}_\delta\hat{\mathbf{\Gamma}}' \\ &= \mathbf{Z}\mathbf{D}_d^{1/2}\hat{\mathbf{D}}_r^{-1}\hat{\mathbf{P}}'\hat{\mathbf{D}}_\delta\mathbf{P}\mathbf{D}_r^{-1}\mathbf{D}_d^{1/2}\mathbf{Z}' \\ &= \frac{1}{k}(\mathbf{Z}_1^* \quad \mathbf{Z}_2^*)\begin{pmatrix} k\dot{\mathbf{D}}_d^* & \mathbf{0} \\ \mathbf{0} & (k-1)\mathbf{I}_{q^*} + \ddot{\mathbf{D}}_d^* \end{pmatrix}\begin{pmatrix} \mathbf{Z}_1^{*} \\ \mathbf{Z}_2^{*\prime} \end{pmatrix} \\ &= \mathbf{Z}_1^*\dot{\mathbf{D}}_d^*\mathbf{Z}_1^{*\prime} + \frac{k-1}{k}\mathbf{Z}_2^*\mathbf{Z}_2^{*\prime} + \frac{1}{k}\mathbf{Z}_2^*\ddot{\mathbf{D}}_d^*\mathbf{Z}_2^{*\prime}.\end{aligned}$$

Subtraction of (4.22) from (4.23) and division by k yields

$$\hat{\mathbf{\Theta}} = \frac{1}{k}\mathbf{Z}_1^*(\dot{\mathbf{D}}_d^* - \mathbf{I}_{m^*})\mathbf{Z}_1^{*\prime}, \tag{4.24}$$

which is positive semidefinite with rank m^*.

The maximized log likelihood function is

$$\begin{aligned} \log L(m^*) &= -\frac{npk}{2}\log 2\pi - nk\log|\mathbf{Z}| - \frac{nk}{2}\log|\mathbf{D}_d| \\ &\quad + \frac{nq^*k}{2}\log k - \frac{npk}{2} + \frac{n(k-1)}{2}\sum_{i=1}^{m^*}\log d_i \\ &\quad + \frac{nk}{2}\sum_{i=m^*+1}^{p}\log d_i - \frac{nk}{2}\sum_{i=m^*+1}^{p}\log(k-1+d_i) \\ &= -\frac{npk}{2}\log 2\pi + \frac{nq^*k}{2}\log k - \frac{npk}{2} - nk\log|\mathbf{Z}| - \frac{n}{2}\sum_{i=1}^{m^*}\log d_i \\ &\quad - \frac{nk}{2}\sum_{i=m^*+1}^{p}\log(k-1+d_i). \end{aligned} \tag{4.25}$$

This expression agrees with the substitution of $\hat{\mathbf{\Psi}}$ and $\hat{\mathbf{\Psi}} + k\hat{\mathbf{\Theta}}$ into (3.5).

Any matrix $\hat{\mathbf{B}}$ satisfying $\hat{\mathbf{B}}\hat{\mathbf{\Theta}} = \mathbf{0}$ is a maximum likelihood estimator of $\mathbf{B}$. In particular, $\mathbf{Y}$, which consists of the last q^* rows of $(\mathbf{Z}')^{-1}$, has the required property. Thus $\hat{\mathbf{B}}$ is any nonsingular multiple of $\mathbf{Y}$.

The likelihood ratio criterion for testing the null hypothesis H_0: $m \le m_0$ against the alternative $m_0 < m \le m_1$ is

$$\begin{aligned} \frac{L(m_0^*)}{L(m_1^*)} &= \frac{k^{nq_0^*k/2}\prod_{i=1}^{m_1^*} d_i^{n/2}\prod_{i=m_1^*+1}^{p}(k-1+d_i)^{nk/2}}{k^{nq_1^*k/2}\prod_{i=1}^{m_0^*} d_i^{n/2}\prod_{i=m_0^*+1}^{p}(k-1+d_i)^{nk/2}} \\ &= \left[\prod_{i=m_0^*+1}^{m_1^*}\frac{k^k d_i}{(k-1+d_i)^k}\right]^{n/2}, \end{aligned} \tag{4.26}$$

if $m_0^* < m_1^*$ and is 1 if $m_0^* = m_1^*$.

REFERENCES

AMEMIYA, Y. (1985). What should be done when an estimated between-group covariance matrix is not nonnegative definite? *Amer. Statist.* **39** 112–117.

AMEMIYA, Y. and FULLER, W. (1984). Estimation for the multivariate errors-in-variables model with estimated error covariance matrix. *Ann. Statist.* **12** 497–509.

ANDERSON, T. W. (1946). Analysis of multivariate variance. Unpublished manuscript.

ANDERSON, T. W. (1984). Estimating linear statistical relationships (The 1982 Wald Memorial Lectures). *Ann. Statist.* **12** 1–45.

KLOTZ, J. and PUTTER, J. (1969). Maximum likelihood estimation of the multivariate covariance components for the balanced one-way layout. *Ann. Math. Statist.* **40** 1100–1105.

MORRIS (ANDERSON), B. and OLKIN, I. (1964). Some estimation and testing problems for factor analysis models. Unpublished manuscript.

VON NEUMANN, J. (1937). Some matrix-inequalities and metrization of matrix space. *Tomsk University Review* **1** 283–300. Reprinted (1962) in *John von Neumann Collected Works* (A. H. Taub, ed.) **4** 205–219. Pergamon, New York.

RAO, C. R. (1983). Likelihood ratio tests for relationships between two covariance matrices. In *Studies in Econometrics, Time Series, and Multivariate Statistics* (S. Karlin, T. Amemiya, and L. A. Goodman, eds.) 529–543. Academic, New York.

SCHOTT, J. R. and SAW, J. G. (1984). A multivariate one-way classification model with random effects. *J. Multivariate Anal.* **15** 1–12.

THEOBALD, C. M. (1975). An inequality with application to multivariate analysis. *Biometrika* **62** 461–466.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305