# NOTES

## MINIMUM GENERALIZED VARIANCE FOR A SET OF LINEAR FUNCTIONS[1]

By Robert G. D. Steel

*University of Wisconsin*

**1. Summary.** Let $n$ variates possessing finite first and second moments be partitioned into $k$ sets. A system of equations is developed for which some solution consists of $k$ sets of coefficients which combine the $k$ sets of variates into $k$ variates possessing minimum generalized variance.

**2. Introduction.** Let $x_1, \cdots, x_n$ be observed variates having zero means and finite covariances, $\sigma_{ij}, i, j = 1, \cdots, n$. Let the variates be such that

$$|\sigma_{ij}| \neq 0, \tag{1}$$

where $|\sigma_{ij}|$ is the determinant of the covariance matrix, $\Sigma = (\sigma_{ij})$. Partition the $n$ variates into $k$ vector variates

$$x_{(j)} = (x_{m_j+1}, \cdots, x_{m_j+n_j}), \qquad j = 1, \cdots, k, \tag{2}$$

where $m_j = \sum_{i=1}^{j-1} n_i, j \neq 1, m_1 = 0$, and $n_i \leq n_{i+1}$. Partition $\Sigma$ correspondingly as

$$\Sigma = \begin{pmatrix} \Sigma_{11} \Sigma_{12} & \cdots & \Sigma_{1k} \\ \Sigma'_{12} \Sigma_{22} & \cdots & \Sigma_{2k} \\ \cdots & \cdots & \cdots \\ \Sigma'_{1k} \Sigma'_{2k} & \cdots & \Sigma_{kk} \end{pmatrix} \tag{3}$$

where $\Sigma_{ij}$ is an $n_i$ by $n_j$ matrix with transpose $\Sigma'_{ij}$.

For $k = 2$, Hotelling [1] forms two variates, one linear function of the variates for each vector variate, for which the correlation is maximum. This leads to canonical variates and canonical correlations. Wilks' $\lambda_I$ criterion [2] is the likelihood ratio criterion for testing independence among $k$ sets of normally distributed variates.

We consider the choice of $k$ linear functions, one per vector variate (2), such that the $k$ resulting variates possess minimum generalized variance when each has unit variance.

[1] This paper is based on doctoral dissertation 957 on file at the Iowa State College Library and submitted to the Graduate Faculty of Iowa State College in partial fulfillment of the requirements for the degree Ph.D. in Mathematical Statistics. A part of the research was conducted under the sponsorship of the Office of Naval Research. The principal result in the paper was presented to the American Mathematical Society in Lawrence, Kansas, April 30, 1949.

**3. Some theorems and the procedure.**

THEOREM 1. *A set of linear functions with unit variances exists possessing minimum generalized variance.*

PROOF. From the vector variates (2), form variates

$$a_{(j)} x'_{(j)} \tag{4}$$

($'$ denotes transpose), where $a_{(j)} = (a_{m_j+1}, \cdots, a_{m_j+n_j})$ is a 1 by $n_j$ vector of real numbers such that

$$a_{(j)} \Sigma_{jj} a'_{(j)} = 1, \qquad j = 1, \cdots, k. \tag{5}$$

The variates (4) have zero means and unit variances. Denote the determinant of the covariance matrix by $C$. $C$ is a bounded continuous function of $n$ $a_i$'s and is defined jointly over $k$ closed connected sets, the positive definite quadratic forms (5). Therefore, there exists a minimum for $C$.

Apply real nonsingular transformations to the vector variates (2):

$$x_{(j)} T'_j = \xi_{(j)}, \qquad j = 1, \cdots, k, \tag{6}$$

for $T_j$ an $n_j$ by $n_j$ matrix such that

$$T_j \Sigma_{jj} T'_j = I_j, \qquad j = 1, \cdots, k. \tag{7}$$

$I_j$ is an $n_j$ by $n_j$ identity matrix.

The $n$ by $n$ matrix

$$T = \begin{pmatrix} T_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & T_2 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & T_k \end{pmatrix} \tag{8}$$

will be called an internal linear transformation when applied to the $k$ ordered vector variates simultaneously. $T$ is real and nonsingular.

The covariance matrix of $(\xi_1, \cdots, \xi_n)$ is

$$T\Sigma T' = \begin{pmatrix} I_1 & T_1\Sigma_{12}T'_2 & \cdots & T_1\Sigma_{1k}T'_k \\ T_2\Sigma'_{12}T'_1 & I_2 & \cdots & T_2\Sigma_{2k}T'_k \\ \cdots & \cdots & \cdots & \cdots \\ T_k\Sigma'_{1k}T'_1 & T_k\Sigma'_{2k}T'_2 & \cdots & I_k \end{pmatrix}.$$

Denote it by $\Gamma$ and $T_i\Sigma_{ij}T'_j$ by $\Gamma_{ij}$.

LEMMA 1. *Transforming matrices $T_j$ subject to* (7) *differ only by multiplication by orthogonal matrices.*

Hence matrices such as (8) differ only by multiplication by orthogonal matrices consisting of blocks of orthogonal matrices in the diagonals and zeros elsewhere.

Restrict further internal linear transformations to orthogonal matrices.

From among such transformations, it is required to choose one such that the $k$ resulting variates shall possess minimum generalized variance. Clearly, this restriction on the transformations is no restriction on the value of the minimum generalized variance. Let each of the $k$ desired variates be the first in its respective vector variate. Movement of any variate within a vector variate can be accomplished by means of a permutation matrix, included among the permissible orthogonal matrices.

Consider the effect of an orthogonal internal linear transformation on the covariance matrix of the $k$ variates that are the first ones in the $k$ vector variates (6) subject to (7). To do this, examine the effect of the transforming matrix on the appropriate $k$ by $k$ principal minor of $\Gamma$. The effect may be observed by using compound matrices, in which any principal minor of the matrix compounded occurs as a diagonal element. Denote the $k$th compound of a square matrix $A$ with elements $a_{ij}$ by $A^{(k)}$ with elements $a_{ij}^{(k)}$.

Let the orthogonal internal linear transformation be $P$ with

$$\xi_{(\alpha)} P'_\alpha = \eta_{(\alpha)}, \qquad \alpha = 1, \cdots, k, \tag{9}$$

for $P_\alpha$ and $n_\alpha$ by $n_\alpha$ orthogonal matrix. Denote the elements of $P_\alpha$ by ${}_\alpha p_{ij}$, $i, j = 1, \cdots, n_\alpha$. The covariance matrix of $(\eta_1, \cdots, \eta_n)$ is $P\Gamma P'$. Denote this by $N$. The $k$th compound of $N$ is

$$N^{(k)} = (P\Gamma P')^{(k)} = P^{(k)}\Gamma^{(k)}P^{(k)\prime}. \tag{10}$$

The principal minor of $\Gamma$ which we wish to observe under the transformation (9) is that with diagonal elements the unities in the upper left corners of the $I_j$. This minor appears as a diagonal element of $\Gamma^{(k)}$, its transform being in the same position in $N^{(k)}$.

The transformed principal minor from (10) is given by

$$\begin{aligned} (t_1, \cdots, t_{\binom{n}{k}})\Gamma^{(k)}(t_1, \cdots, t_{\binom{n}{k}})' &= \sum_{\delta,\beta=1}^{\binom{n}{k}} t_\delta \gamma_{\delta\beta}^{(k)} t_\beta \\ &= \nu_{gg}^{(k)}, \end{aligned} \tag{11}$$

where the subscript $g$ is appropriately chosen and the elements $t_i$ are the $k$ by $k$ minors, ordered lexicographically, of the matrix consisting of rows $m_1 + 1, \cdots, m_k + 1$ of $P$. Clearly each nonzero $t_\delta$ consists of a single product of one element from the first row of each $P_\alpha$.

The problem is to determine those ${}_\alpha p_{1i}$'s, $i = 1, \cdots, n_\alpha$ and $\alpha = 1, \cdots, k$, which minimize $\nu_{gg}^{(k)}$. Hence let us obtain partial derivatives of the elements of the $P_\alpha$'s with respect to the parameters of these orthogonal matrices.

**4. The derivative of an orthogonal matrix.** If a nonsingular square matrix $A$ is a function of $x$, then by differentiating $AA^{-1}$ it follows that

$$\frac{dA}{dx} = -A\frac{dA^{-1}}{dx}A, \qquad \frac{dA^{-1}}{dx} = -A^{-1}\frac{dA}{dx}A^{-1}. \tag{12}$$

LEMMA 2. *For a real nonexceptional orthogonal matrix Q,*

$$\frac{\partial q_{\alpha\beta}}{\partial s_{ij}} = -\frac{1}{2}\begin{vmatrix} (\delta_{\alpha i} + q_{\alpha i}) & (\delta_{\alpha j} + q_{\alpha j}) \\ (\delta_{i\beta} + q_{i\beta}) & (\delta_{j\beta} + q_{j\beta}) \end{vmatrix}, \tag{13}$$

*where $q_{\alpha\beta}$ is the $\alpha$, $\beta$th element of Q and $s_{ij}$ is the i, jth element of the skew-symmetric matrix S in Cayley's*[2] *parametrization of Q, viz.,* $Q = (I - S)/(I + S)$.

PROOF. Let the parameters be $s_{ij}$, $i < j$. By (12) and since $(I + S)^{-1} = \frac{1}{2}(I + Q)$,

$$\frac{\partial Q}{\partial s_{ij}} = -\tfrac{1}{2}(I + Q)\,\frac{\partial S}{\partial s_{ij}}\,(I + Q).$$

Now $\partial S/\partial s_{ij}$ is a matrix consisting of a $+1$ in the $i, j$th position, a $-1$ in the $j, i$th position and zeros elsewhere. Hence, elementwise, we have (13) where $\delta$ is Kronecker's delta. The partial derivatives are expressed in terms of the elements of $Q$.

An exceptional orthogonal matrix becomes nonexceptional when multiplied by an appropriate $J$ matrix, a square matrix with $+1$ or $-1$ in each diagonal position and zeros elsewhere. When $JP$ replaces $P$, principal minors of $P\Gamma P'$ will not alter in value though some variates will change sign. Hence, let us consider only real nonexceptional orthogonal transformations.

To maximize $\nu_{gg}^{(k)}$ as in (11), equate to zero:

$$\frac{\partial \nu_{gg}^{(k)}}{\partial\, {}_\alpha s_{ij}} = \sum_{\beta=1}^{\binom{n}{k}} \frac{\partial \nu_{gg}^{(k)}}{\partial t_\beta}\,\frac{\partial t_\beta}{\partial\, {}_\alpha s_{ij}} = 2 \sum_{\beta,\delta=1}^{\binom{n}{k}} t_\delta \nu_{\delta\beta}^{(k)}\,\frac{\partial t_\beta}{\partial\, {}_\alpha s_{ij}}, \tag{14}$$

where ${}_\alpha s_{ij}$ is a parameter of $P_\alpha$, $\alpha = 1, \cdots, k$. The partial derivatives of the $t_\beta$'s are found by use of (13). Equations (14) with those imposing orthogonality restrictions on the $P_\alpha$ are $\Sigma_{\alpha=1}^{k} n_\alpha^2$ simultaneous equations in as many unknowns, the elements of the $P_\alpha$'s.

**5. Two-set case.** For $k = 2$, variates having minimum generalized variance are seen to have maximum correlation. Since Hotelling's canonical form is unique except for the order and signs of the elements in the diagonal of the off-diagonal block and the maximum correlation is always present there, it can be shown that a solution obtained by the present method will agree with that from Hotelling's method.

**6. A three-set case.** For $k = 3$ and $n_1 = n_2 = n_3 = 2$, let $\Gamma$ be such that each of $(\Gamma_{12}, \Gamma_{13})$, $(\Gamma_{12}', \Gamma_{23})$, and $(\Gamma_{13}', \Gamma_{23}')$ has unit rank. Chu [4] has shown that the rows of $(I_1, \Gamma_{12}, \Gamma_{13})$, $(\Gamma_{12}', I_2, \Gamma_{23})$, and $(\Gamma_{13}', \Gamma_{23}', I_2)$ can all be internally

[2] For a recent work, see Weyl [3].

orthogonalized by an orthogonal internal linear transformation. The resulting covariance matrix may be written

$$\begin{pmatrix} 1 & 0 & \rho_{13} & 0 & \rho_{15} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \rho_{13} & 0 & 1 & 0 & \rho_{35} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \rho_{15} & 0 & \rho_{35} & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}, \quad \text{say}, \tag{15}$$

where $X$, $Y$, and $Z$ are 2 by 6 matrices.

Now $XX'$, $YY'$, and $ZZ'$ have distinct characteristic roots. Hence (15) is unique except for internal permutations of rows and columns and changes in the signs of $\rho_{13}$, $\rho_{15}$, and $\rho_{35}$. Also, the orthogonal transforming matrix is unique except that the signs of the elements in any row may be altered simultaneously.

It is now easy to find variates, one per set, with minimum generalized variance. It can be shown that the only permissible orthogonal internal linear transformation consists of an identity matrix. Hence the minimum generalized variance is

$$\begin{vmatrix} 1 & \rho_{13} & \rho_{15} \\ \rho_{13} & 1 & \rho_{35} \\ \rho_{15} & \rho_{35} & 1 \end{vmatrix}.$$

It is unique, as are also the internal linear transformation and the resulting variates.

**7. Acknowledgment.** The author wishes to express his appreciation for the constructive assistance and encouragement given by Professor B. Vinograde of the Mathematics Department of Iowa State College.

## REFERENCES

[1] H. Hotelling, "Relations between two sets of variates," *Biometrika*, Vol. 28 (1936), pp. 321–377.

[2] S. S. Wilks, "On the independence of $k$ sets of normally distributed statistical variables," *Econometrica*, Vol. 3 (1935); pp. 309–326.

[3] H. Weyl, *The Classical Groups, Their Invariants and Representations*, Princeton University Press, 1939, chap. II, B, sec. 10.

[4] J. T. Chu, "Invariants of positive definite matrices under internal transformations," unpublished M. S. thesis, Iowa State College Library, 1948.