Estimating Functions of Canonical Correlation Coefficients*

Robb J. Muirhead and Pui Lam Leung
Department of Statistics
The University of Michigan
Ann Arbor, Michigan 48109-1027

Submitted by George P. H. Styan

ABSTRACT

Let $\rho_1^2,\dots,\rho_p^2$ be the squares of the population canonical correlation coefficients from a normal distribution. This paper is concerned with the estimation of the parameters $\delta_1,\dots,\delta_p$, where $\delta_i=\rho_i^2/(1-\rho_i^2)$, $i=1,\dots,p$, in a decision theoretic way. The approach taken is to estimate a parameter matrix Δ whose eigenvalues are $\delta_1,\dots,\delta_p$, given a random matrix F whose eigenvalues have the same distribution as $r_i^2/(1-r_i^2)$, $i=1,\dots,p$, where $r_1,\dots,r_p$ are the sample canonical correlation coefficients.

1. INTRODUCTION

Problems concerning the estimation of population eigenvalues are of great interest in multivariate analysis. This paper is essentially concerned with estimating certain functions of canonical correlation coefficients. Suppose that the $(p+q)\times(p+q)$ positive definite matrix S has the Wishart distribution with n degrees of freedom and positive definite parameter covariance matrix Σ, written $S\sim W_{p+q}(n,\Sigma)$, and partition S and Σ as

$$S=\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22}\end{bmatrix},\qquad \Sigma=\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22}\end{bmatrix},$$

where S_{11} and Σ_{11} are $p\times p$, and S_{22} and Σ_{22} are $q\times q$, with $p\leqslant q$. The positive square roots $\rho_1,\dots,\rho_p$ $(1\geqslant\rho_1\geqslant\cdots\geqslant\rho_p\geqslant 0)$ of $\rho_1^2,\dots,\rho_p^2$, the eigenvalues of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, are the population canonical correlation coeffi-

*This work was supported by the National Science Foundation.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/85/$3.30

cients. The squares of the sample canonical correlation coefficients are $r_1^2, \ldots, r_p^2$ $(1 > r_1 > \cdots > r_p > 0)$, the eigenvalues of $S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}$. Discussions of canonical correlation analysis may be found in, e.g., [1, Chapter 12] and [3, Chapter 11]. These eigenvalues are also important in the problem of testing $H: \Sigma_{12} = 0$ against $K: \Sigma_{12} \neq 0$, as they form maximal invariants under a natural group of transformations leaving the testing problem invariant. Any invariant test statistic is a function of $r_1^2, \ldots, r_p^2$ and has power function depending only on $\rho_1^2, \ldots, \rho_p^2$.

The work that follows represents an attempt to estimate the parameters $\delta_i = \rho_i^2/(1-\rho_i^2)$, $i = 1, \ldots, p$, in a decision theoretic way. Ideally, such an approach would specify a loss function in terms of these parameters, and risk computations would involve expectations of this loss taken with respect to the joint distribution of $r_1^2, \ldots, r_p^2$. Such an approach, however, does not seem feasible, due primarily to the complexity of the distribution of $r_1^2, \ldots, r_p^2$ [3, Section 11.3.2.]. Instead, we concentrate on estimating a parameter matrix Δ whose eigenvalues are $\delta_1, \ldots, \delta_p$, given a random matrix F whose eigenvalues have the same distribution as the variables $y_i = r_i^2/(1-r_i^2)$, $i = 1, \ldots, p$. It is then natural to hope that the eigenvalues of a "good" estimate $\hat{\Delta}(F)$ of Δ will be reasonable estimates of $\delta_1, \ldots, \delta_p$.

2. ESTIMATES OF Δ

The parameters $\delta_i = \rho_i^2/(1-\rho_i^2)$, $i = 1, \ldots, p$, are the eigenvalues of the parameter matrix

$$\Delta = \Sigma_{11.2}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11.2}^{-1/2}, \tag{1}$$

where $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. We assume for simplicity here that Δ is positive definite. Naive estimates of $\delta_1, \ldots, \delta_p$ are $y_i = r_i^2/(1-r_i^2)$, $i = 1, \ldots, p$, the eigenvalues of $S_{11.2}^{-1}S_{12}S_{22}^{-1}S_{21}$ or, equivalently, the eigenvalues of the random matrix

$$F = B^{1/2}A^{-1}B^{1/2}, \tag{2}$$

where $A = \Sigma_{11.2}^{-1/2}S_{11.2}\Sigma_{11.2}^{-1/2}$, $B = \Sigma_{11.2}^{-1/2}S_{12}S_{22}^{-1}S_{21}\Sigma_{11.2}^{-1/2}$. Put $X = \Sigma_{22}^{-1/2}S_{22}\Sigma_{22}^{-1/2}$; from standard distribution theory (see, e.g., [3, Theorem 3.2.10 and Section 10.3], it follows that $A \sim W_p(n-q, I_p)$ and is independent of X and B, that $X \sim W_q(n, I_q)$, and that, given X, the conditional distribution of B is noncentral Wishart $W_p(q, I_p, \Omega)$, where the noncentrality parameter matrix Ω is

$$\Omega=\Sigma_{11.2}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2}X\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11.2}^{-1/2}. \tag{3}$$

The probability density function (pdf) of F may be shown to be

$$\frac{\Gamma_p(\frac{1}{2}n)}{\Gamma_p(\frac{1}{2}q)\Gamma_p(\frac{1}{2}(n-q))}\frac{(\det F)^{\frac{1}{2}(q-p-1)}}{\det(I+F)^{\frac{1}{2}n}}\det(I+\Delta)^{-n/2}$$

$$\times{}_2F_1\Big(\tfrac{1}{2}n,\tfrac{1}{2}n;\tfrac{1}{2}q;\Delta(I_p+\Delta)^{-1}F(I_p+F)^{-1}\Big), \tag{4}$$

where Δ is given by (1),

$$\Gamma_p(a)=\pi^{\frac{1}{4}p(p-1)}\prod_{i=1}^{p}\Gamma\big(a-\tfrac{1}{2}(i-1)\big),$$

and ${}_2F_1$ is a hypergeometric function of matrix argument, having an infinite series expansion in terms of zonal polynomials (see [2] or [3, Chapter 7]). It is worth emphasizing that although the random matrix F is not observable if Σ is unknown, its eigenvalues are equivalent, in distribution, to $y_1,\dots,y_p$, where $y_i=r_i^2/(1-r_i^2)$, and these are observable. An approach to estimating $\delta_1,\dots,\delta_p$, and the one suggested here, is to estimate Δ by an orthogonally invariant estimate, i.e., an estimate of the form

$$\hat{\Delta}(F)=H\phi(Y)H'$$

where H is a $p\times p$ orthogonal matrix such that $F=HYH'$, with $Y=\operatorname{diag}(y_1,\dots,y_p)$, and $\phi(Y)=\operatorname{diag}(\phi_1(Y),\dots,\phi_p(Y))$. The variables $\phi_i(Y)$, $i=1,\dots,p$, may then be regarded as estimates of $\delta_1,\dots,\delta_p$. The only orthogonally invariant estimates considered in this paper are ones of the form $\alpha F+\beta I$, so that the corresponding estimates of $\delta_1,\dots,\delta_p$ are $\alpha r_i^2/(1-r_i^2)+\beta$, $i=1,\dots,p$.

It has been suggested by an anonymous referee that, rather than using F, it would be better to choose an observable matrix whose eigenvalues are the same as those of F. Such a matrix can be chosen in several ways; however, the distribution theory and associated problems of finding expectations are greatly simplified by the choice of F.

Our starting point, then, is an observation on a random $p\times p$ positive definite matrix F having pdf (4), and we consider the problem of estimating

Δ by $\hat{\Delta}(F)$ using the loss functions

$$L_1(\Delta, \hat{\Delta}) = \operatorname{tr}(\Delta^{-1}\hat{\Delta}) - \ln\det(\Delta^{-1}\hat{\Delta}) - p \tag{5}$$

and

$$L_2(\Delta, \hat{\Delta}) = \operatorname{tr}(\hat{\Delta} - \Delta)^2. \tag{6}$$

The corresponding risk functions, involving expectations of L_1 and L_2 with respect to the distribution of F, will be subscripted similarly.

The pdf (4) of F is not particularly convenient for finding expectations of functions of F. It is easier to use the representation $F = B^{1/2}A^{-1}B^{1/2}$ and the known distributional results for A and B. The expectation of F is, using standard arguments and known results about expectations of Wishart, inverse Wishart, and noncentral Wishart matrices,

$$\begin{aligned} E(F) &= E(B^{1/2}A^{-1}B^{1/2}) \\ &= E\big(E(B^{1/2}A^{-1}B^{1/2} \mid B)\big) \\ &= E\big(B^{1/2}E(A^{-1})B^{1/2}\big) \\ &= \frac{1}{n-q-p-1}E(B) \\ &= \frac{1}{n-q-p-1}E(E(B \mid X)) \\ &= \frac{1}{n-q-p-1}E\big(qI_p + \Sigma_{11.2}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2}X\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11.2}^{-1/2}\big) \\ &= \frac{1}{n-q-p-1}(qI_p + n\Delta) \qquad (n-q-p-1>0). \end{aligned} \tag{7}$$

It follows that an unbiased estimate of Δ is the orthogonally invariant estimate

$$\hat{\Delta}_U = \frac{n-q-p-1}{n}F - \frac{q}{n}I_p. \tag{8}$$

Using the loss function L_1 given by (5), $\hat{\Delta}_U$ is the best estimate in the class of estimates of the form $\alpha\hat{\Delta}_U$, as the following result shows.

Theorem 1. *Using the loss function L_1, the best (smallest risk) estimate of Δ having the form $\alpha\hat{\Delta}_U$ is the unbiased estimate $\hat{\Delta}_U$.*

Proof. The risk of the estimate $\alpha\hat{\Delta}_U$ is

$$R_1(\alpha\hat{\Delta}_U, \Delta) = E\left[\alpha \operatorname{tr}(\Delta^{-1}\hat{\Delta}_U) - \ln\det(\alpha\Delta^{-1}\hat{\Delta}_U) - p\right]$$
$$= \alpha p - p\ln\alpha - E\left[\ln\det(\Delta^{-1}\hat{\Delta}_U)\right] - p.$$

The proof is completed by noting that this is minimized for all Δ when $\alpha = 1$. ■

Using the loss function L_2, however, $\hat{\Delta}_U$ is dominated by $\alpha\hat{\Delta}_U$ for some choices of α. In order to show this, we need the expectation of $\operatorname{tr}(F^2)$. This is given in the following lemma.

Lemma 2. *If $n - q - p - 3 > 0$, then*

$$E\left[\operatorname{tr}(F^2)\right] = \beta_0\left[\beta_1(\operatorname{tr}\Delta)^2 + \beta_2\operatorname{tr}(\Delta^2) + \beta_3\operatorname{tr}\Delta + \beta_4\right], \tag{9}$$

where

$$\beta_0 = \frac{1}{(n-q-p)(n-q-p-1)(n-q-p-3)},$$
$$\beta_1 = n(2n - q - p - 1),$$
$$\beta_2 = n(n^2 - nq - np - q - p + 1),$$
$$\beta_3 = 2n\left[(n-q-p)(p+q+1) + (p-1)(q-1)\right],$$
$$\beta_4 = pq\beta_3/2n. \tag{10}$$

Proof. The proof is rather long and algebraically messy, and we will merely sketch it. Using the standard vec notation and direct products, we can write

$$\begin{aligned}
E\left[\operatorname{tr}(F^2)\right] &= E\left[\operatorname{vec}(A^{-1})'(B\otimes B)\operatorname{vec}(A^{-1})\right] \\
&= E\left[\operatorname{tr}(B\otimes B)\operatorname{vec}(A^{-1})\operatorname{vec}(A^{-1})'\right] \\
&= \operatorname{tr} E(B\otimes B)E\left[\operatorname{vec}(A^{-1})\operatorname{vec}(A^{-1})'\right] \\
&= \operatorname{tr} E(B\otimes B)\left[\beta_0(I_{p^2}+K)+\beta_5\operatorname{vec}(I_p)\operatorname{vec}(I_p)'\right],
\end{aligned}$$

where β_0 is given in (10),

$$\beta_5 = \frac{1+2(n-q-p-1)\beta_0}{(n-q-p-1)^2},$$

and K denotes the $p^2 \times p^2$ matrix

$$K = \sum_{i,j=1}^{p} \left(H_{ij}\otimes H'_{ij}\right),$$

with H_{ij} being the $p \times p$ matrix with i-j element equal to 1 and all other elements zero. It then follows easily that

$$E\left[\operatorname{tr}(F^2)\right] = \beta_0 E\left[(\operatorname{tr} B)^2\right] + \beta_6 E\left[\operatorname{tr}(B^2)\right], \tag{11}$$

where $\beta_6 = \beta_0(n-q-p-1)$. Next, conditioning on X, we have

$$\begin{aligned}
E\left[(\operatorname{tr} B)^2 \middle| X\right] &= E\left[\operatorname{vec}(I_p)'\operatorname{vec}(B)\operatorname{vec}(B)'\operatorname{vec}(I_p)\middle| X\right] \\
&= \operatorname{vec}(I_p)' E\left[\operatorname{vec}(B)\operatorname{vec}(B)' \middle| X\right]\operatorname{vec}(I_p) \\
&= \operatorname{vec}(I_p)'\Big\{(I_{p^2}+K)\left[qI_{p^2}+(I_p\otimes\Omega)+(\Omega\otimes I_p)\right] \\
&\quad + \left[q\operatorname{vec}(I_p)+\operatorname{vec}(\Omega)\right]\left[q\operatorname{vec}(I_p)+\operatorname{vec}(\Omega)\right]'\Big\}\operatorname{vec}(I_p). \\
&= 2\operatorname{vec}(I_p)'\left[qI_{p^2}+(I_p\otimes\Omega)+(\Omega\otimes I_p)\right]\operatorname{vec}(I_p) \\
&\quad + (pq+\operatorname{tr}\Omega)^2 \\
&= 2pq + 4\operatorname{tr}\Omega + (pq+\operatorname{tr}\Omega)^2,
\end{aligned} \tag{12}$$

where Ω is given by (3). A similar argument shows that

$$E\left[\operatorname{tr}(B^2)|X\right] = E\left[\operatorname{tr}\operatorname{vec}(B)\operatorname{vec}(B)'\right]$$
$$= pq(p+q+1)+2(p+q+1)\operatorname{tr}\Omega+\operatorname{tr}(\Omega^2). \qquad (13)$$

Now taking expectations with respect to the $W_q(n, I_q)$ distribution for X, we obtain, using similar calculations,

$$E(\operatorname{tr}\Omega) = n\operatorname{tr}\Delta, \qquad (14)$$

$$E\left[(\operatorname{tr}\Omega)^2\right] = 2n\operatorname{tr}(\Delta^2)+n^2(\operatorname{tr}\Delta)^2, \qquad (15)$$

and

$$E\left[\operatorname{tr}(\Omega^2)\right] = n(\operatorname{tr}\Delta)^2+n(n+1)\operatorname{tr}(\Delta^2). \qquad (16)$$

Hence we have

$$E\left[\operatorname{tr}(F^2)\right] = \beta_0 E\left[E\left[(\operatorname{tr}B)^2|X\right]\right]+\beta_6 E\left\{E\left[\operatorname{tr}(B^2)|X\right]\right\},$$

and using (12)–(16) gives the required result. ■

Using the loss function L_2, the risk of the unbiased estimate $\hat{\Delta}_U$ is, using (7) and (9),

$$R_2(\hat{\Delta}_U, \Delta) = E\left[\operatorname{tr}\left(\frac{n-q-p-1}{n}F-\frac{q}{n}I_p-\Delta\right)^2\right]$$
$$= a(\operatorname{tr}\Delta)^2+b\operatorname{tr}(\Delta^2)+c\operatorname{tr}\Delta+d, \qquad (17)$$

where

$$a = \frac{(2n-q-p-1)(n-q-p-1)}{n(n-q-p)(n-q-p-3)}, \qquad (18)$$

$$b = \frac{(2n-q-p-1)(n-q-p+1)}{n(n-q-p)(n-q-p-3)}, \qquad (19)$$

$$c = \frac{2(n-p-1)[(p+1)(n-q-p)-(p-1)]}{n(n-q-p)(n-q-p-3)}, \qquad (20)$$

and

$$d = pqc/2n, \tag{21}$$

and we are assuming henceforth that $n - q - p - 3 > 0$. The unbiased estimate $\hat{\Delta}_U$ is dominated by estimates of the form $\alpha\hat{\Delta}_U$. The risk of $\alpha\hat{\Delta}_U$ is

$$\begin{aligned} R_2(\alpha\hat{\Delta}_U, \Delta) &= E\left[\operatorname{tr}(\alpha\hat{\Delta}_U - \Delta)^2\right] \\ &= \alpha^2 R_2(\hat{\Delta}_U, \Delta) + (1-\alpha)^2 \operatorname{tr}(\Delta^2). \end{aligned} \tag{22}$$

The following theorem gives conditions under which $\hat{\Delta}_U$ is dominated by $\alpha\hat{\Delta}_U$.

THEOREM 3. *The estimate $\alpha\hat{\Delta}_U$ dominates $\hat{\Delta}_U$ provided that*

$$\max\left(0, \frac{1-b}{b+1}\right) \leqslant \alpha < 1, \tag{23}$$

where b is given by (19). *(Note that for large n, $(1-b)/(b+1) \approx (n-2)/(n+2)$.)*

Proof. From (17) and (22) the difference between the risks of $\hat{\Delta}_U$ and $\alpha\hat{\Delta}_U$ is

$$\begin{aligned} G(\Delta) &\equiv R_2(\hat{\Delta}_U, \Delta) - R_2(\alpha\hat{\Delta}_U, \Delta) \\ &= a(1-\alpha^2)(\operatorname{tr}\Delta)^2 + (1-\alpha)[\alpha(b+1) + b - 1] \operatorname{tr}(\Delta^2) \\ &\quad + c(1-\alpha^2)\operatorname{tr}\Delta + d(1-\alpha^2). \end{aligned}$$

In $G(\Delta)$ the constant term and the coefficients of $\operatorname{tr}\Delta$ and $(\operatorname{tr}\Delta)^2$ are positive; the proof is completed by noting that the coefficient of $\operatorname{tr}(\Delta^2)$ is positive provided $\alpha > (1-\mathrm{b})/(\mathrm{b}+1)$. ■

A referee suggested the following rather more illuminating proof, which shows that (23) is a sufficient condition, independent of Δ, for $G(\Delta) > 0$ to hold. Considered as a concave quadratic form in α, $G(\Delta)$ has two real roots, namely 1 and

$$\alpha_0(\Delta) = -\frac{a(\operatorname{tr}\Delta)^2 + (b-1)\operatorname{tr}(\Delta^2) + c\operatorname{tr}\Delta + d}{a(\operatorname{tr}\Delta)^2 + (b+1)\operatorname{tr}(\Delta^2) + c\operatorname{tr}\Delta + d},$$

where $-1<\alpha_0(\Delta)<1$ with $(1+\alpha_0)/2>0$. Hence if $\max(0,\alpha_0(\Delta))<\alpha<1$ then $G(\Delta)>0$. To obtain a lower limit, which is independent of Δ, on α for $G(\Delta)>0$ to hold, we seek x independent of Δ such that $\alpha_0(\Delta)<x<1$. To this end, consider

$$Q(\Delta)\equiv[x-\alpha_0(\Delta)]\left[a(\operatorname{tr}\Delta)^2+(b+1)\operatorname{tr}(\Delta^2)+c\operatorname{tr}\Delta+d\right].$$

After some minor manipulation $Q(\Delta)$ can be written as

$$Q(\Delta)=\boldsymbol{\delta}'A\boldsymbol{\delta}+\mathbf{b}'\boldsymbol{\delta}+c_0, \tag{24}$$

where $\boldsymbol{\delta}'=(\delta_1,\ldots,\delta_p)$, with $\delta_1,\ldots,\delta_p$ the eigenvalues of Δ; $A=(g-h)I_p+hJ_p$, where J_p is a $p\times p$ matrix with every element equal to 1, $g=(a+b+1)x+a+b-1$, and $h=ax+a$; $\mathbf{b}=(cx+c)\mathbf{1}_p$, with $\mathbf{1}'_p=(1,1,\ldots,1)$; and $c_0=dx+d$. Now, $x>\alpha_0(\Delta)$ is equivalent to $Q(\Delta)>0$, and from (24) this is implied by $A>0$, i.e., to the eigenvalues $\lambda_1,\ldots,\lambda_p$ all being positive. These eigenvalues are $\lambda_1=g+(p-1)h$ and $\lambda_2=\cdots=\lambda_p=g-h$. These are positive when $x>(1-b-ap)/(1+b+ap)$ and $x>(1-b)/(1+b)$. It is readily shown that

$$\frac{1-b-ap}{1+b+ap}<\frac{1-b}{1+b}<1 \quad\text{and}\quad \alpha_0(\Delta)<\frac{1-b}{1+b},$$

which leads to $x>(1-b)/(1+b)$. Hence it is sufficient for $G(\Delta)>0$ to hold, independent of Δ, that α satisfy (23).

When $p=1$, $r^2\equiv r_1^2$ and $\rho^2\equiv\rho_1^2$ are respectively the squares of the sample and multiple correlation coefficients. In this case the unique minimum variance unbiased estimate of $\delta=\rho^2/(1-\rho^2)$ is

$$\hat{\delta}_U=\frac{n-q-2}{n}\,\frac{r^2}{1-r^2}-\frac{q}{n},$$

and an argument similar to that above shows that, using squared-error loss, $\hat{\delta}_U$ is beaten by $\alpha\hat{\delta}_U$ for α satisfying $\max(0,(1-a-b)/(1+a+b))\leqslant\alpha<1$. Moreover, Muirhead [4] has shown that these latter estimates are beaten by nonlinear estimates of the form $\alpha\hat{\delta}_U+\beta(1-r^2)/r^2$.

It is seen from (17) and (22) that the coefficient of $\operatorname{tr}(\Delta^2)$ in $R_2(\alpha\hat{\Delta}_U,\Delta)$ is minimized when $\alpha=1/(b+1)$, and this value of α satisfies (23). The corresponding estimate $\hat{\Delta}_L\equiv(b+1)^{-1}\hat{\Delta}_U$ thus dominates $\hat{\Delta}_U$.

TABLE 1

PRIAL OF $\hat{\Delta}_L$ OVER $\hat{\Delta}_U$ WHEN $p = q = 3$

Σ_{12}	n = 25	50	75	100
diag(.5,.5,.5)	10.57	4.13	4.82	1.64
diag(.3,.2,.1)	16.07	9.52	7.69	0.00
diag(.7,.6,.5)	2.63	3.42	3.94	2.08
diag(.1,.1,.1)	17.65	10.00	0.00	0.00
diag(.9,.9,.9)	3.75	2.95	2.92	0.04
diag(.9,.5,.1)	4.93	0.96	3.71	1.40

A Monte Carlo study was carried out to compare $\hat{\Delta}_L$ with $\hat{\Delta}_U$. For $p = q = 3$ and $n = 25, 50, 75, 100$, a sample of 100 Wishart $W_6(n, \Sigma)$ matrices were generated, where

$$\Sigma = \begin{bmatrix} I_3 & \Sigma_{12} \\ \Sigma_{12}' & I_3 \end{bmatrix},$$

for various choices of a diagonal matrix Σ_{12}. Then 100 F's were formed using (2), and were used to construct both $\hat{\Delta}_U$ and $\hat{\Delta}_L$, and from these average losses (with respect to L_2) were obtained. Table 1 summarizes the results. In this table the value given for each combination of Σ_{12} and n is the percentage reduction in average loss (PRIAL) for $\hat{\Delta}_L$ compared with $\hat{\Delta}_U$, i.e., it is the estimate of

$$\frac{R_2(\hat{\Delta}_U, \Delta) - R_2(\hat{\Delta}_L, \Delta)}{R_2(\hat{\Delta}_U, \Delta)} \times 100$$

obtained by replacing risk with average loss. It appears that $\hat{\Delta}_L$ can represent a reasonable improvement over $\hat{\Delta}_U$, especially when n is small.

Estimates of the parameters $\delta_i = \rho_i^2/(1-\rho_i^2)$, $i = 1, \ldots, p$, obtained from the estimate $\alpha\hat{\Delta}_U$ in Theorem 3, are

$$\hat{\delta}_i = \alpha\left[\frac{n-q-p-1}{n}\frac{r_i^2}{1-r_i^2} - \frac{q}{n}\right],$$

where α satisfies (23). Current work is proceeding on evaluating these estimates and on obtaining other orthogonally invariant estimates of Δ.

The authors would like to express their thanks to the referees for detailed comments and suggestions and to one referee for the illuminating proof of Theorem 3.

REFERENCES

1 T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, 1958.

2 A. T. James, Distributions of matrix variates and latent roots derived from normal samples, *Ann. Math. Statist.* 35:475–501 (1964).

3 R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley, New York, 1982.

4 R. J. Muirhead, Estimating a particular function of the multiple correlation coefficient. *J. Am. Statist. Soc.*, to appear.

Received 24 June 1984; revised 1 March 1985