The Annals of Statistics
1980, Vol. 8, No. 5, 1156–1159

INADMISSIBILITY OF THE BEST FULLY EQUIVARIANT ESTIMATOR OF THE GENERALIZED RESIDUAL VARIANCE[1]

BY KAM-WAH TSUI, SAMARADASA WEERAHANDI, AND JIM ZIDEK

Simon Fraser University, University of Sri Jayewardenepura, and University of British Columbia

The estimation of the generalized residual variance is considered when an observable Wishart random matrix is available. It is shown that when the loss function is normalized squared error, the natural fully equivariant estimator is dominated by an alternative. The latter uses either one of two estimators depending on the result of a preliminary test of significance. This alternate estimator has an everywhere smaller mean normalized squared error than the natural estimator.

This paper is concerned with the estimation of the generalized residual variance. To explicate the problem of interest we let $Z = (Y, X)$ denote a $1 \times (q + p)$ multinormal random vector with unknown covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Here Σ_{11} denotes the covariance matrix of the random q-vector, Y, and so on. Then $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ is the conditional covariance of Y given X. The quantity $\Sigma_{11.2}$ is commonly referred to as the residual covariance matrix because it represents the covariation of the residual error vector, $Y - \alpha - X\beta$, where $\beta = \Sigma_{22}^{-1}\Sigma_{21}$ and $\alpha = E(Y) - E(X)\beta$ and hence $\alpha + X\beta$ represents the best linear predictor of Y based on X when the mean of Z is known. The generalized residual variance is defined as $\sigma^2 = |\Sigma_{11.2}|$. It is a numerical measure of the "size" of $\Sigma_{11.2}$ and is closely related to the volume of the ellipsoid of concentration determined by $\Sigma_{11.2}$ (cf. Cramér 1961).

Because σ^2 measures the efficacy of the multivariate regression of Y on X, the problem of estimating it arises in a natural way. This is the problem at the basis of the work presented here. The loss function adopted is the most mathematically convenient one, namely normalized squared error:

$$L(\hat{\sigma}^2; \Sigma) = (\hat{\sigma}^2 - \sigma^2)^2\sigma^{-4}.$$

We show that the natural (best equivariant) estimator is dominated by a preliminary test estimator whether or not Z's mean is known. Following the route traced out by Brewster and Zidek (1974), one could obtain similar results for a large class of potential loss functions.

Received October 1978; revised May 1979.

[1]This work was supported by a grant from the National Research Council of Canada, No. A5241.

AMS 1970 *subject classifications.* Primary 62F10; secondary 62H99.

Key words and phrases. Inadmissible estimator, testimation, regression efficiency.

Our interest in this problem was stimulated by our work on a formally similar problem which is much harder and as yet unsolved, namely that of obtaining an improved estimator of $|\Sigma|$. More explicitly, let $Z_1, \cdots, Z_n$ be n independent observations of $Z \sim N_{q+p}(\mu, \Sigma)$, $\bar{Z} = \Sigma_{i=1}^n Z_i/n$, $V^* = \Sigma_{i=1}^n (Z_i - \bar{Z})'(Z_i - \bar{Z})$ and $V = \Sigma_{i=1}^n (Z_i - \mu)'(Z_i - \mu)$. The minimal sufficient statistic is V or $(\bar{Z}, V^*)$ according as μ is known or unknown. The usual (best equivariant) estimator of $|\Sigma|$ is $(n+2)^{-1}|V|$ or $(n+1)^{-1}|V^*|$ according as μ is known or unknown. As Selliah (1964) shows (and Kiefer's theorem (1957) implies), the usual estimator of $|\Sigma|$ is minimax whether or not the mean is known when loss is squared error. Shorrock and Zidek (1976) do obtain a better estimator than the usual one when the mean is unknown. Their new estimator utilizes the information about $|\Sigma|$ in the sample mean on top of that in the sample covariance in order to achieve a mean normalized squared error which is uniformly smaller than that of the usual estimator. However, it is not known whether or not the usual estimator of $|\Sigma|$ is admissible when the mean is known. This problem, because of its difficulty, is of considerable interest even though the problem considered here would seem to be of greater practical significance.

Inference about σ^2 is to be based on $V \sim W_{q+p}(n, \Sigma)$, i.e., an observable Wishart random variable with n degrees of freedom. It will facilitate our analysis to employ the familiar upper triangular decomposition of Σ and V. Thus $\Sigma = \tau\tau'$ and $V = TT'$, where

$$T = \begin{pmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{pmatrix},$$

etc. Observe that $V =_L \tau\tilde{V}\tau'$ where $\tilde{V} \sim W_{q+p}(n, I)$ and "$=_L$" means "is identically distributed as." Furthermore, $\tilde{V} = \tilde{T}\tilde{T}'$, where (cf. Wijsman (1957), Kshirsagar (1959))

$$\tilde{T} = \begin{bmatrix} \chi_{n-(p+q-1)} & Z_{11} & \cdots & Z_{1,p+q-1} \\ 0 & \chi_{n-(p+q-2)} & \cdots & Z_{2,p+q-2} \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \cdots & \chi_n \end{bmatrix},$$

χ_j^2 has the chi-squared distribution with j degrees of freedom, and $Z_{ij} \sim N(0, 1)$. In summary:

$$\Sigma = \tau\tau', \qquad V = TT' =_L \tau\tilde{V}\tau', \qquad T = \tau\tilde{T}.$$

The best fully equivariant estimator is easily obtained. Let G denote the group of all upper triangular matrices. Under the action of $g \in G$ the problem remains invariant and $T \to gT$, $\tau \to g\tau$, and $\sigma^2 \to |g_{11}|^2\sigma^2$. Thus any equivariant estimator, $\hat{\sigma}^2$, must satisfy the condition $\hat{\sigma}^2(gT) = |g_{11}|^2\hat{\sigma}^2(T)$ which implies, as is easily shown, that $\hat{\sigma}^2(T) = c|T_{11}|^2 = c|V_{11.2}|$ for some constant $c > 0$. Since G operates transitively on the parameter space the expected mean squared error of such an estimator is independent of Σ. Thus there is a best value of c, namely $c = c_0 =$

$\Pi_{i=1}^{q}(n - p - q + i + 2)^{-1}$, and a best fully equivariant estimator, $\hat{\sigma}_0^2$, given by

(1) $$\hat{\sigma}_0^2 = c_0 |V_{11.2}|.$$

The content of this paper amounts to a demonstration that another estimator

(2) $$\hat{\sigma}_1^2 = \min\{c_0|V_{11.2}|, c_1|V_{11}|\}, \qquad c_1 = \Pi_{i=1}^{q}(n + 3 - i)^{-1}$$

has a uniformly smaller mean normalized squared error than $\hat{\sigma}_0^2$. This superior alternative is found by considering a larger class of estimators than those equivariant under G. Recognizing the need to restrict the scope of the search and hence this class of potential alternatives, one is tempted to consider the class of estimators equivariant under a subgroup of G. The choice $H = \{g : g \in G, g_{12} = 0\}$ is the natural one and this leads to estimators of the form

$$\hat{\sigma}^2 = |T_{11}|^2 \psi_1(R) = |V_{11}| \psi_2(R),$$

where $R = T_{11}^{-1} T_{12}$. Even this class proves to be unduly large and technical difficulties force us to consider the subclass of estimators of the form

(3) $$\hat{\sigma}^2 = |V_{11}| \psi(D),$$

where D is a $p \times p$ diagonal matrix whose diagonal elements are the roots d of the equation $|RR' - d(I + RR')| = 0$, or equivalently, the equation $|T_{12}T'_{12} - d(T_{11}T'_{11} + T_{12}T'_{12})| = 0$. In terms of T_{11}, T_{12} and T_{22}, (3) becomes

(4) $$\hat{\sigma}^2 = |T_{11}T'_{11} + T_{12}T'_{12}| \psi(D).$$

Our goal is to find the estimator of the form given in equation (4) which minimizes the risk function, $R(\hat{\sigma}^2, \Sigma) = E[\hat{\sigma}^2\sigma^{-2} - 1]^2$. To this end, observe that $T_{11}T'_{11} \sim W_p(n, \Sigma_{11.2})$ (a central Wishart) and given $\tilde{T}_{22}$, $T_{12}T'_{12}$ is a noncentral Wishart. Now the mathematical results from Shorrock and Zidek (1976) can be applied here by setting up the following mapping of their notation into that of this paper:

$$\begin{array}{ll}
S \to T_{11}T'_{11} = V_{11.2}, & p \to q, \\
\Sigma \to \Sigma_{11.2} = \tau_{11}\tau'_{11}, & n \to n - p, \\
X \to T_{12}, & k \to p, \\
\xi \to \tau_{12}\hat{T}_{22}, & \psi \to c_1, \\
|\Sigma| \to \sigma^2, & c \to c_0, \\
T \to D. &
\end{array}$$

By conditioning on $\tilde{T}_{22}$, and with $\hat{\sigma}_0^2$ and $\hat{\sigma}_1^2$ as given by (1) and (2), respectively, the inequality

$$E\left(\hat{\sigma}_1^2\sigma^{-2} - 1\right)^2 < E\left(\hat{\sigma}_0^2\sigma^{-2} - 1\right)^2$$

follows directly (compare c_1 and equation (2) with (3.20) and (3.22) of Shorrock and Zidek).

REMARK 1. $\hat{\sigma}_1^2$ is a "testimator." It is $c_1|V_{11}|$ if a preliminary test accepts $\Sigma_{12} = 0$, i.e., if $c_1 < c_0|V_{11.2}||V_{11}|^{-1}$. This would be the natural estimator to use were it known that $\Sigma_{12} = 0$. Otherwise $\sigma_1^2 = \hat{\sigma}_0^2$.

REMARK 2. When the mean of the underlying normal law as well as its covariance is unknown, a still better estimator of σ^2 than $\hat{\sigma}_1^2$ may be obtained. Such an estimator would use the information in the sample mean as well as that in the sample covariance. However we have not carried out the details.

REMARK 3. A closely related problem which we have not considered is that of estimating $1 - |\Sigma_{11.2}|/|\Sigma_{11}| = 1 - \prod_{i=1}^{q}(1 - \rho_i^2)$ where the $\{\rho_i\}$ are the canonical correlation coefficients. This is a natural measure of the efficiency of the best linear predictor of Y from X.

Acknowledgments. The authors would like to thank the referees for their insightful comments.

REFERENCES

BREWSTER, J. F. and ZIDEK, J. V. (1974). Improving on equivariant estimators. *Ann. Statist.* **2** 21–38.

CONSTANTINE, A. G. (1963). Some noncentral distribution problems in multivariate analysis. *Ann. Math. Statist.* **34** 1270–1285.

CRAMÉR, H. (1961). *Mathematical Methods of Statistics*. Ninth edition. Princeton Univ. Press.

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

KSHIRSAGAR, A. M. (1959). Bartlett decomposition and Wishart distribution. *Ann. Math. Statist.* **30** 239–241.

SELLIAH, J. B. (1964). Estimation and testing problems in a Wishart distribution. Stanford Univ. Tech. Report no. 10. Dept. of Statist.

SHORROCK, R. W. and ZIDEK, J. V. (1976). An improved estimator of the generalized variance. *Ann. Statist.* **4** 629–638.

WIJSMAN, R. A. (1957). Random orthogonal transformations and their use in some classical distribution problems in multivariate analysis. *Ann. Math. Statist.* **28** 415–423.

KAM-WAH TSUI
DEPARTMENT OF MATHEMATICS
SIMON FRASER UNIVERSITY
BURNABY, B. C.
CANADA V5A 1S6

SAMARADASA WEERAHANDI
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF JAYEWARDENEPURA
NUGEGODA
SRI LANKA

JIM ZIDEK
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
2075 WESBROOK MALL
VANCOUVER, B. C.
CANADA V6T 1W5