The Annals of Statistics
1979, Vol. 7, No. 4, 838–846

MINIMAX ESTIMATION OF A NORMAL MEAN VECTOR WHEN THE COVARIANCE MATRIX IS UNKNOWN[1]

BY LEON JAY GLESER

Purdue University

Let X be an observation from a p-variate normal distribution ($p > 3$) with mean vector θ and unknown positive definite covariance matrix Σ. We wish to estimate θ under the quadratic loss $L(\delta; \theta, \Sigma) = [\text{tr}(Q\Sigma)]^{-1}(\delta - \theta)'Q(\delta - \theta)$, where Q is a known positive definite matrix. Estimators of the following form are considered:

$$\delta_{k,h}(X, W) = [I - kh(X'W^{-1}X)\lambda_1(QW/n^*)Q^{-1}W^{-1}]X,$$

where $W : p \times p$ is observed independently of X and has a Wishart distribution with n degrees of freedom and parameter Σ, $\lambda_1(A)$ denotes the minimum characteristic root of A, and $h(t) : [0, \infty) \to [0, \infty)$ is absolutely continuous with respect to Lebesgue measure, is nonincreasing, and satisfies the additional requirements that $th(t)$ is nondecreasing and $\sup_{t \geq 0} th(t) = 1$. With $h(t) = t^{-1}$, the class $\delta_{k,h}$ specializes to that considered by Berger, Bock, Brown, Casella and Gleser (1977). For the more general class considered in the present paper, it is shown that there is an interval $[0, k_{n,p}]$ of values of k (which may be degenerate for small values of $n - p$) for which $\delta_{k,h}$ is minimax and dominates the usual estimator $\delta_0 \equiv X$ in risk.

1. Introduction. Assume that X is a p-dimensional random column vector which is normally distributed with mean vector θ and unknown positive definite covariance matrix Σ. We observe X, and also independently observe the $p \times p$ random matrix W, which has a Wishart distribution with n degrees of freedom and parameter $\Sigma = n^{-1}E(W)$. It is desired to estimate θ by an estimator $\delta(X, W)$ under the quadratic loss

$$L(\delta(X, W); \theta, \Sigma) = \frac{1}{\text{tr}(\Sigma Q)}\left[(\delta(X, W) - \theta)'Q(\delta(X, W) - \theta)\right] \tag{1.1}$$

where Q is a known $p \times p$ positive definite matrix. For this problem it is known that the classical least squares, maximum likelihood estimator $\delta_0(X, W) = X$ is minimax. It has often been conjectured that X is not admissible if $p \geq 3$. Recently, Berger, Bock, Brown, Casella and Gleser (1977) provided the first explicit examples of estimators which dominate X in risk in this context. These estimators have the form

$$\delta_c(X, W) = \left[I_p - \left(\frac{c\lambda_1(QW/n^*)}{X'W^{-1}X}\right)Q^{-1}W^{-1}\right]X \tag{1.2}$$

Received September 1977; revised May 1978.

[1]Research supported by the Air Force Office of Scientific Research, Air Force Systems Command, USAF under Grant AFOSR-77-3291.

AMS 1970 *subject classifications.* Primary 62C99; secondary 62F10, 62H99.

Key words and phrases. Minimax, normal, mean, quadratic loss, unknown covariance matrix, Wishart distribution, risk function.

where $n^* = n - p - 1$, $\lambda_1(A)$ denotes the minimum characteristic root of a matrix A, and $c \geqslant 0$ is a nonnegative constant. A certain constant $c_{n,p}$ depending on the dimension p of X and on the degrees of freedom n of W is analytically defined in Berger, et al., [1977], Equation (2.18), page 768, and it is shown that if $c_{n,p} > 0$, then for any c, $0 < c \leqslant c_{n,p}$, the estimator $\delta_c(X, W)$ dominates X in risk. However, it is extremely difficult to evaluate $c_{n,p}$ for any finite n and p, so Berger, et al., resort to Monte Carlo simulation to evaluate $c_{n,p}$ and find values of n and p for which this constant is positive. Ignoring the very small probability that an extreme error in the Monte Carlo simulations may have occurred, their work shows that X is inadmissible for $p \geqslant 3$ and $n - p$ large enough (usually $n - p \geqslant 5$ is enough).

As noted in Berger, et al, (1977), the class (1.2) contains no admissible procedures, since every member of the class (except $\delta_0(X, W) = X$) has a singularity at $X = 0$. Both for analytic reasons and based on risk simulations, they conjectured that an estimator of the form

$$(1.3) \qquad \delta_c^*(X, W) = \left[I_p - \frac{\min(c, n^* X' W^{-1} X) \lambda_1(QW/n^*)}{X' W^{-1} X} Q^{-1} W^{-1} \right] X$$

would dominate $\delta_c(X, W)$ in risk. However, they were not able to prove that $\delta_c^*(X, W)$ is minimax.

In an earlier paper, the present author [Gleser (1976)] showed that when a lower bound, K, for $\lambda_1(Q\Sigma)$ is known, estimators of the form

$$(1.4) \qquad \delta_h(X, W) = \left[I_p - ch(X' W^{-1} X) Q^{-1} W^{-1} \right] X,$$

where $h(t) : [0, \infty) \rightarrow [0, \infty)$ is absolutely continuous with respect to Lebesgue measure, $th(t)$ is nondecreasing in t, $\sup_{t \geqslant 0} th(t) = 1$, and $0 \leqslant c \leqslant 2(p-2)(n-p) K/(n-1)$ dominate $\delta_0(X, W) = X$ in risk. In that paper it was also shown that if no lower bound to $\lambda_1(Q\Sigma)$ is known, then no estimator of the form (1.4) can dominate $\delta_0(X, W) = X$ in risk.

In the present paper, the methods of Gleser (1976) and Berger, et al, (1977) are combined to prove the following result:

THEOREM 1.1. *Consider the class of estimators*

$$(1.5) \qquad \delta_{k,h}(X, W) = \left[I_p - kh(X' W^{-1} X) \lambda_1(QW/n^*) Q^{-1} W^{-1} \right] X$$

where $h(t) : [0, \infty) \rightarrow [0, \infty)$ *is absolutely continuous with respect to Lebesgue measure, and*

(i) $th(t)$ *is nondecreasing in* t;

(ii) $\sup_{t \geqslant 0} th(t) = 1$;

(iii) $h(t)$ *is nonincreasing in* t.

Let $k_{n,p}$ *be defined by Equation* (2.27). *If* $k_{n,p} > 0$, *then for any* k, $0 < k \leqslant k_{n,p}$, *the estimator* $\delta_{k,h}(X, W)$ *dominates* $X = \delta_0(X, W)$ *in risk, and hence, is minimax.*

It is not difficult to show that the class of rules (1.5) includes both the class of rules (1.2) and the class (1.3).

2. Proof of Theorem 1. Let

(2.1) $\Delta \equiv \Delta(\theta, \Sigma) = [\operatorname{tr}(Q\Sigma)]E^{X,W}[L(\delta_{k,h}(X, W); \theta, \Sigma) - L(X; \theta, \Sigma)]$.

As in Berger, et al, (1977), superscripts on the expected value operator indicate the random variable with respect to which the expectation is to be computed. The quantity $\Delta = \Delta(\theta, \Sigma)$ is the difference in risks between $\delta_{k,h}(X, W)$ and $\delta_0(X, W) = X$, weighted by the positive quantity $\operatorname{tr}(Q\Sigma)$. If $\Delta = \Delta(\theta, \Sigma) \leqslant 0$ for all θ and Σ, then $\delta_{k,h}(X, W)$ dominates $\delta_0(X, W)$ in risk.

The initial steps of our proof closely follow those of Berger, et al, (1977). We expand the quadratic loss of $\delta_{k,h}$, take expected values in the order $E^W E^X$, and use the familiar technique of integration by parts [Berger (1976)] to obtain

$$E^X[h(X'W^{-1}X)(X-\theta)'W^{-1}X] = E^X[h(X'W^{-1}X) \operatorname{tr} \Sigma W^{-1} + 2h^{(1)}(X'W^{-1}X)X'W^{-1}\Sigma W^{-1}X],$$

where $h^{(1)}(t) = dh(t)/dt$ exists almost surely. These steps yield the result:

(2.2)

$$\Delta = E^{X,W}\{k\lambda_1[k\lambda_1 h^2(X'W^{-1}X)X'W^{-1}Q^{-1}W^{-1}X - 2h(X'W^{-1}X) \operatorname{tr} \Sigma W^{-1} - 4h^{(1)}(X'W^{-1}X)X'W^{-1}\Sigma W^{-1}X]\},$$

where $\lambda_1 = \lambda_1(QW/n^*)$. Next we note that

$$X'W^{-1}Q^{-1}W^{-1}X \leqslant \frac{X'W^{-1}X}{n^*\lambda_1(QW/n^*)}, \tag{2.3}$$

and thus, since $th(t) \leqslant 1$,

$$k\lambda_1(QW/n^*)h^2(X'W^{-1}X)X'W^{-1}Q^{-1}W^{-1}X \leqslant kh(X'W^{-1}X)/n^*.$$

Also, since $th(t)$ is nondecreasing,

$$0 \leqslant \frac{d}{dt}(th(t)) = h(t) + th^{(1)}(t),$$

so that

$$h^{(1)}(X'W^{-1}X) \geqslant -\frac{h(X'W^{-1}X)}{X'W^{-1}X}. \tag{2.4}$$

Substituting these results into (2.2), we conclude that

(2.5)

$$\Delta \leqslant E^{X,W}\left\{k\lambda_1(QW/n^*)h(X'W^{-1}X)\left[\frac{k}{n^*} - 2 \operatorname{tr} \Sigma W^{-1} + 4\frac{X'W^{-1}\Sigma W^{-1}X}{X'W^{-1}X}\right]\right\}.$$

For any square root $\Sigma^{-\frac{1}{2}}$ of Σ, let

$$Y = \Sigma^{-\frac{1}{2}}X \qquad U = \Sigma^{-\frac{1}{2}}W\Sigma^{-\frac{1}{2}} \tag{2.6}$$

and

$$\eta = \Sigma^{-\frac{1}{2}}\theta \qquad \phi = \Sigma^{\frac{1}{2}}Q\Sigma^{\frac{1}{2}}.$$

Then Y and U are statistically independent, Y has a p-variate normal distribution with mean vector η and covariance matrix I_p, and U has a Wishart distribution with n degrees of freedom and parameter I_p. In terms of Y and U,

$$(2.7)\quad \Delta \leqslant E^Y E^U\left\{\frac{k}{n^*}\lambda_1(\phi U)h(Y'U^{-1}Y)\left[\frac{k}{n^*} - 2\operatorname{tr} U^{-1} + 4\frac{Y'U^{-2}Y}{Y'U^{-1}Y}\right]\right\}.$$

Finally, let Γ_Y be an orthogonal $p \times p$ matrix satisfying

$$(2.8)\quad Y'\Gamma_Y' = \left((Y'Y)^{\frac{1}{2}}, 0, 0, \cdots, 0\right)$$

and let

$$(2.9)\quad V = \Gamma_Y U \Gamma_Y', \qquad \phi_Y = \Gamma_Y \phi \Gamma_Y'.$$

Then the conditional distribution of V given Y is a Wishart distribution with n degrees of freedom and parameter I_p, independent of Y. Also
(2.10)

$$\operatorname{tr} U^{-1} = \operatorname{tr} V^{-1}, \qquad Y'U^{-1}Y = Y'Y(V^{-1})_{11}, \qquad Y'U^{-2}Y = Y'Y(V^{-2})_{11}.$$

Let

$$(2.11)\quad v_1 = (V^{-1})_{11} \qquad v_2 = (V^{-2})_{11}.$$

It follows that

$$(2.12)\quad \Delta \leqslant E^Y E^V\left\{\frac{k}{n^*}\lambda_1(\phi_Y V)h(Y'Yv_1)\left[\frac{k}{n^*} - 2\operatorname{tr} V^{-1} + 4\frac{v_2}{v_1}\right]\right\}.$$

Let $r(t) = th(t)$. Then from (2.12), we obtain

$$(2.13)\quad \Delta \leqslant \frac{k}{n^*}E^Y\left\{\left(\frac{1}{Y'Y}\right)E^V\left[r(Y'Yv_1)\frac{\lambda_1(\phi_Y V)}{v_1}\left(\frac{k}{n^*} - 2\operatorname{tr} V^{-1} + 4\frac{v_2}{v_1}\right)\right]\right\}.$$

Let

$$\beta = \lambda_1(\phi_Y) = \lambda_1(\Gamma_Y\phi\Gamma_Y') = \lambda_1(\phi) = \lambda_1(Q\Sigma),$$

and let

$$\Sigma^* = \frac{1}{\beta}\phi_Y.$$

Then $\lambda_1(\Sigma^*) = 1$, and
(2.14)

$$\Delta \leqslant -\frac{k\beta}{n^*}E^Y\left\{\left(\frac{1}{Y'Y}\right)E^V\left[r(Y'Yv_1)\frac{\lambda_1(\Sigma^* V)}{v_1}\left(2\operatorname{tr} V^{-1} - 4\frac{v_2}{v_1} - \frac{k}{n^*}\right)\right]\right\}.$$

Hence to show that $\Delta = \Delta(\theta, \Sigma) \leqslant 0$ for all θ, Σ, it suffices to show that for all Σ^* with $\lambda_1(\Sigma^*) = 1$ and all values of $Y'Y$ the following inequality holds:

$$(2.15)\quad \tau = E^V\left\{r(Y'Yv_1)\frac{\lambda_1(\Sigma^* V)}{v_1}\left(2\operatorname{tr} V^{-1} - 4\frac{v_2}{v_1} - \frac{k}{n^*}\right)\right\} \geqslant 0.$$

Except for the term $r(Y'Yv_1)$, this inequality is precisely the inequality in Equation (2.7) of Berger, et al, (1977). (For the class of estimators considered in Berger, et al, (1977), $r(t) \equiv 1$ for all t.) Our problem is to account for the $r(Y'Yv_1)$ term in (2.15). To do so, we make use of a distributional representation previously utilized in Gleser (1976).

LEMMA 2.1. *Let V have a Wishart distribution with degrees of freedom n and parameter I_p, and let V be partitioned as*

$$V = \begin{pmatrix} v_{11} & V_{12} \\ V_{12}' & V_{22} \end{pmatrix}, v_{11} : 1 \times 1,\ V_{22} : (p-1) \times (p-1). \tag{2.16}$$

Let $l' = V_{12}V_{22}^{-\frac{1}{2}}$ for any square root $V_{22}^{\frac{1}{2}}$ of V_{22}. Then $v_1 = (V^{-1})_{11}$, l and V_{22} are mutually statistically independent, v_1^{-1} has a χ^2_{n-p+1} distribution, l has a $(p-1)$-variate standard multivariate normal distribution, and V_{22} has a Wishart distribution with n degrees of freedom and parameter I_{p-1}.

PROOF. This is a well-known result easily proved by making the indicated changes of variables in the density of V. □

Let V^{-1} be partitioned as

$$V^{-1} = \begin{pmatrix} v^{11} & V^{12} \\ (V^{12})' & V^{22} \end{pmatrix} = \begin{pmatrix} v_1 & V^{12} \\ (V^{12})' & V^{22} \end{pmatrix}$$

similar to V in (2.16). Using the well-known relationships between the block elements of V and of V^{-1} it can be shown that

$$v_2 = v_1^2(1 + l'V_{22}^{-1}l),\ \operatorname{tr} V^{-1} = \operatorname{tr} V_{22}^{-1} + v_1(1 + l'V_{22}^{-1}l) \tag{2.17}$$

and that

$$V = \begin{bmatrix} v_1^{-1} + l'l & l'V_{22}^{\frac{1}{2}} \\ (l'V_{22}^{\frac{1}{2}})' & V_{22} \end{bmatrix}. \tag{2.18}$$

Note from (2.18) that, for fixed values of l and V_{22}, V is decreasing in v_1 in the sense of positive definiteness, and that v_1V is increasing in v_1. Thus, for fixed values of l and V_{22},

$$\lambda_1(\Sigma^* V) \quad \text{is decreasing in } v_1,\ v_1\lambda_1(\Sigma^* V) \quad \text{is increasing in } v_1. \tag{2.19}$$

It follows from (2.15), (2.17), (2.18) and Lemma 2.1 that

$$\tau = E^{l, V_{22}}E^{v_1}\left\{ r(Y'Yv_1)\frac{\lambda_1(\Sigma^* V)}{v_1}\left[2 \operatorname{tr} V_{22}^{-1} - 2v_1(1 + l'V_{22}^{-1}l) - \frac{k}{n^*}\right]\right\}. \tag{2.20}$$

Since $r(Y'Yv_1)$ is an increasing function of v_1, we could try to use the following well-known lemma to pull $E^{v_1}[r(Y'Yv_1)]$ out of the expected value in (2.20).

LEMMA 2.2. *Let $g_1(s)$ and $g_2(s)$ map the real line into the real line and let S be any random variable. Then if $g_1(s)$ and $g_2(s)$ are either both nonincreasing in s or both* *nondecreasing in s,*

$$E^S[g_1(S)g_2(S)] \geqslant E^S[g_1(S)]E^S[g_2(S)].$$

PROOF. This lemma follows by direct application of Lemma 1, (i) and (iii), and Lemma 3 of Lehmann (1966). □

Unfortunately $\lambda_1(\Sigma^* V)/v_1$ times the quantity in square brackets in (2.20) is neither a nondecreasing nor a nonincreasing function of v_1. We thus try a more indirect attack, attempting to pull $E^{v_1}[h(Y'Yv_1)]$ out of the expected value in (2.20). (Note that neither $r(Y'Yv_1)$ nor $h(Y'Yv_1)$ depend upon l and V_{22}.)

Let $g_1(v_1) = Y'Yh(Y'Yv_1)$, $g_2(v_1) = 2\lambda_1(\Sigma^* V)\mathrm{tr}\ V_{22}^{-1}$, and $g_3(v_1) = -2\lambda_1(\Sigma^* V)v_1(1 + l'V_{22}^{-1}l)$. Each of these functions is nonincreasing in $v_1 : g_1$ is nonincreasing since $h(t)$ is nonincreasing in t and $Y'Y \geqslant 0$; g_2 is nonincreasing because $\lambda_1(\Sigma^* V)$ is nonincreasing in v_1 (see (2.19)); g_3 is nonincreasing because $v_1\lambda_1(\Sigma^* V)$ is nondecreasing in v_1 (again, see (2.19)). Now

$$\begin{aligned}
&E^{v_1}\left\{r(Y'Yv_1)\frac{\lambda_1(\Sigma^* V)}{v_1}\left[2\ \mathrm{tr}\ V_{22}^{-1} - 2v_1\left(1 + l'V_{22}^{-1}l\right)\right]\right\}\\
&\qquad = E^{v_1}\left\{Y'Yh(Y'Yv_1)\lambda_1(\Sigma^* V)\left[2\ \mathrm{tr}\ V_{22}^{-1} - 2v_1\left(1 + l'V_{22}^{-1}l\right)\right]\right\}\\
&\qquad = E^{v_1}[g_1(v_1)g_2(v_1)] + E^{v_1}[g_1(v_1)g_3(v_1)]\\
&\qquad \geqslant E^{v_1}[g_1(v_1)]E^{v_1}[g_2(v_1)] + E^{v_1}[g_1(v_1)]E^{v_1}[g_3(v_1)]\\
&\qquad = E^{v_1}[g_1(v_1)]\{E^{v_1}[g_2(v_1) + g_3(v_1)]\}\\
&\qquad = E^{v_1}[Y'Yh(Y'Yv_1)]E^{v_1}\left[\lambda_1(\Sigma^* V)\left(2\ \mathrm{tr}\ V^{-1} - 4\frac{v_2}{v_1}\right)\right]
\end{aligned}\tag{2.21}$$

where the inequality follows from two applications of Lemma 2.2. The final equality in (2.21) results by substituting the values of $g_i(v_1)$, $i = 1, 2, 3$, and then applying (2.17).

Similarly, using (2.19), the fact that $r(t)$ is nondecreasing in t while $h(t)$ is nonincreasing in t, and twice applying Lemma 2.2, we find that

$$\begin{aligned}
&E^{v_1}\left\{r(Y'Yv_1)\frac{\lambda_1(\Sigma^* V)}{v_1}\left(-\frac{k}{n^*}\right)\right\}\\
&\qquad \geqslant E^{v_1}\{r(Y'Yv_1)\}E^{v_1}\left\{-\frac{k\lambda_1(\Sigma^* V)}{n^* v_1}\right\}\\
&\qquad = E^{v_1}\{Y'Yh(Y'Yv_1)v_1\}E^{v_1}\left\{\frac{-k\lambda_1(\Sigma^* V)}{n^* v_1}\right\}\\
&\qquad \geqslant E^{v_1}\{Y'Yh(Y'Yv_1)\}E^{v_1}\{v_1\}E^{v_1}\left\{-\frac{k\lambda_1(\Sigma^* V)}{n^* v_1}\right\},\\
&\qquad = E^{v_1}\{Y'Yh(Y'Yv_1)\}E^{v_1}\left\{-\frac{k\lambda_1(\Sigma^* V)}{(n^*)^2 v_1}\right\},
\end{aligned}\tag{2.22}$$

since

$$E^{v_1}[v_1] = E[1/\chi^2_{n-p+1}] = \frac{1}{n-p-1} = \frac{1}{n^*}.$$

Hence,

$$\tau \geqslant E^{v_1}\left\{\frac{Y'Yh(Y'Yv_1)}{n^*}\right\}E^{l,\,V_{22}}E^{v_1}\left\{\frac{\lambda_1(\Sigma^* V)}{v_1}\left[R(V) - \frac{k}{n^*}\right]\right\} \tag{2.23}$$

where

$$R(V) = n^*(2v_1 \operatorname{tr} V^{-1} - 4v_2). \tag{2.24}$$

It now follows from (2.15), (2.20) and (2.23) that $\Delta \leqslant 0$ if for all Σ^* with $\lambda_1(\Sigma^*) = 1$,

$$E^V\left\{\frac{\lambda_1(\Sigma^* V)}{v_1}\left(R(V) - \frac{k}{n^*}\right)\right\} \geqslant 0. \tag{2.25}$$

This requirement closely resembles Equation (2.7) of Berger, et al, (1977), except that, in place of their $\rho(V) = 2 \operatorname{tr} V^{-1} - 4v_2/v_1$, we have $R(V) = n^* v_1 \rho(V)$. Using the arguments above it can be shown that for every Σ^*,

$$E^V\left\{\frac{\lambda_1(\Sigma^* V)}{v_1}\left(\rho(V) - \frac{k}{n^*}\right)\right\} \geqslant E^V\left\{\frac{\lambda_1(\Sigma^* V)}{v_1}\left(R(V) - \frac{k}{n^*}\right)\right\}, \tag{2.26}$$

so that $k_{n,p}$ obtained from (2.25) will be less than or equal to the value of $c_{n,p}$ obtained in Berger, et al, (1977).

Since $R(V)$ shares with $\rho(V)$ the necessary invariance properties under orthogonal rotation of V by a matrix of the form

$$\begin{pmatrix} 1 & 0 \\ 0 & \Delta \end{pmatrix}, \qquad \Delta : (p-1) \times (p-1) \text{ orthogonal},$$

the arguments of Berger, et al, (1977), can be applied to show that (2.25) holds if $0 \leqslant k \leqslant k_{n,p}$, where $k_{n,p}$ is the solution to

$$k = \min\left\{\frac{\tau_0(k) + \tau_1(k)}{\tau_0'(k) + \tau_1'(k)}, \frac{\tau_0(k)}{\tau_0'(k)}\right\} \tag{2.27}$$

and where

$$\tau_0(k) = E^V\{R(V)v_1^{-1}[v_{22}I_{\Omega_k}(V) + \lambda_1(V)I_{\bar{\Omega}_k}(V)]\},$$

$$\tau_1(k) = E^V\{R(V)v_1^{-1}(v_{11} - v_{22})I_{\Omega_k}(V)\},$$

$$\tau_0'(k) = E^V\{v_1^{-1}[v_{22}I_{\Omega_k}(V) + \lambda_1(V)I_{\bar{\Omega}_k}(V) + \lambda_1(V)I_{\bar{\Omega}_k}(V)]\},$$

$$\tau_1'(k) = E^V\{v_1^{-1}(v_{11} - v_{22})I_{\Omega_k}(V)\},$$

v_{22} is the first diagonal element of V_{22}, and

$$I_{\Omega_k}(V) = 1 - I_{\bar{\Omega}_k}(V) = 1 \quad \text{if } R(V) < k/n^*,$$
$$= 0 \quad \text{otherwise.}$$

This completes the proof of Theorem 1. □

With the help of Dr. George Casella, the identical computer program, modified only by replacing $\rho(V)$ by $R(V)$, used in Berger, et al, (1977), to calculate the values of their $c_{n,p}$ was used to calculate values of $k_{n,p}$ by means of simulation. The resulting values of $k_{n,p}$ appear in Table 1. Where $k_{n,p} > 0$, assuming that the Monte Carlo simulation did not produce a gross error in accuracy (an event of very small probability), Theorem 1 shows that X can be dominated in risk by any $\delta_{k,h}(X, W)$ for which $0 < k \leqslant k_{n,p}$.

TABLE 1
Values of $k_{n,p}$.

	n								
p	8	10	12	14	16	18	20	25	30
3	-	--	.07	.41	.58	.75	.87	1.20	1.26
4		--	1.00	1.45	1.79	2.02	2.25	2.78	2.85
5		.21	1.46	2.11	2.78	3.10	3.39	4.15	4.28
6		--	1.38	2.58	3.47	3.98	4.37	5.25	5.58
7			.36	2.87	3.93	4.70	5.27	6.51	6.92
8			----	2.44	4.19	5.12	5.74	7.41	8.14
9				1.22	3.86	5.56	6.28	8.64	9.16
10				----	3.66	5.17	6.80	9.07	10.22
11					1.28	5.18	7.10	9.95	11.14
12					----	4.21	6.52	10.44	12.06
13						.94	6.25	10.94	13.11
14						----	4.58	11.14	13.62
15							----	10.73	14.19
16							----	10.91	14.71
17								10.45	14.23
18								9.30	15.29
19									14.84
20									14.52

3. Remarks. When compared with the corresponding values of $c_{n,p}$ in Berger, et al, (1977), the values of $k_{n,p}$ are admittedly disappointingly small (being, at best, 90% of the values of $c_{n,p}$). However, these values, even when applied in connection with rules of the form (1.2) or (1.3), produce substantial improvements in risk when compared to $\delta_0(X, W) = X$, especially when $\theta'\theta$ is small. In addition, since (1.5) is a far broader class than (1.2) or (1.3), these constants allow for flexibility in the form of the estimator of θ. In the case when Σ is known, certain rules of the form (1.5) with Σ replacing $(n^*)^{-1}W$ are known to be generalized Bayes and admissible. It is doubtful whether the same assertion can be made about the rules (1.5) when Σ is unknown, largely because of the use of $\lambda_1(QW/n^*)$ in the formula for the estimator.

Several very gross inequalities were used to obtain the results of this paper. The use of inequality (2.3) probably does not lose us much, particularly since the inequality is also used in Berger, et al, (1977). The inequality (2.4) is more serious, amounting to adding the quantity

$$4E^{X,W}\left[\lambda_1(QW/n^*)r^{(1)}(X'W^{-1}X)X'W^{-1}\Sigma W^{-1}X\right] \tag{3.1}$$

to the risk of $\delta_{k,h}(X, W)$. Note that $r^{(1)}(t) = dr(t)/dt$ is positive, so that there is a chance that a substantial improvement in risk for $\delta_{k,h}(X, W)$ over $\delta_0(X, W)$ has been ignored in using the inequality (2.4). Finally, the inequalities leading from (2.20) to (2.25) add an additional inaccuracy to the assessment of risk of $\delta_{k,h}(X, W)$, as can be seen from (2.26). However, unless we want to make use of more detailed information about the form of $h(t)$, these inequalities are unavoidable.

The principal accomplishment of this paper lies in demonstrating that a wide and functionally flexible class of estimators can be used to dominate the usual estimator $\delta_0(X, W) = X$ in risk when Σ is unknown (and $p \geqslant 3$). Whether any particular one of these rules, or any other rule, can be recommended for practical application is still an open question.

REFERENCES

[1] BERGER, J. (1976). Minimax estimation of a multivariate normal mean under arbitrary quadratic loss. *J. Multivariate Anal.* **6** 256–264.

[2] BERGER, J., BOCK, M. E., BROWN, L., CASELLA, G. and GLESER, L. (1977). Minimax estimation of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix. *Ann. Statist.* **4** 763–771.

[3] GLESER, LEON JAY (1976). Minimax estimation of a multivariate normal mean with unknown covariance matrix. Technical Report #460, Purdue Univ.

[4] LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137–1153.

DEPARTMENT OF STATISTICS
MATHEMATICAL SCIENCE BUILDING
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907