The Annals of Statistics
1977, Vol. 5, No. 4, 763-771

MINIMAX ESTIMATION OF A NORMAL MEAN VECTOR FOR ARBITRARY QUADRATIC LOSS AND UNKNOWN COVARIANCE MATRIX

BY J. BERGER,[1] M. E. BOCK,[2] L. D. BROWN,[3] G. CASELLA[4] AND L. GLESER[4]

Purdue University and Rutgers University

Let X be an observation from a p-variate normal distribution ($p \geqq 3$) with mean vector θ and unknown positive definite covariance matrix Σ. It is desired to estimate θ under the quadratic loss $L(\delta, \theta, \Sigma) = (\delta - \theta)^t Q(\delta - \theta)/\mathrm{tr}\,(Q\Sigma)$, where Q is a known positive definite matrix. Estimators of the following form are considered:

$$\delta^c(X, W) = (I - c\alpha Q^{-1}W^{-1}/(X^t W^{-1} X))X,$$

where W is a $p \times p$ random matrix with a Wishart (Σ, n) distribution (independent of X), α is the minimum characteristic root of $(QW)/(n-p-1)$ and c is a positive constant. For appropriate values of c, δ^c is shown to be minimax and better than the usual estimator $\delta^0(X) = X$.

1. Introduction. Assume $X = (X_1, \cdots, X_p)^t$ is a p-dimensional random vector ($p \geqq 3$) which is normally distributed with mean vector $\theta = (\theta_1, \cdots, \theta_p)^t$ and positive definite covariance matrix Σ. It is desired to estimate θ by an estimator $\delta = (\delta_1, \cdots, \delta_p)^t$ under the quadratic loss

$$L(\delta, \theta, \Sigma) = (\delta - \theta)^t Q(\delta - \theta)/\mathrm{tr}\,(Q\Sigma),$$

where Q is a positive definite $(p \times p)$ matrix.

The usual minimax and best invariant estimator for θ is $\delta^0(X) = X$. Since Stein (1955) first showed that δ^0 could be improved upon for $Q = \Sigma = I$ (the identity matrix), a considerable effort by a number of authors (see the references) has gone into finding significant improvements upon δ^0. For the most part these efforts have been directed towards the problems where either Σ was known (or known up to a multiplicative constant) or where $Q = \Sigma^{-1}$ (a rather special situation). For unknown Σ and general Q only a few partial results have been obtained. Berger and Bock (1976a) and (1976b) found minimax estimators (better than δ^0) for problems in which Σ was an unknown diagonal matrix or could be reduced to one. Gleser (1976) found minimax estimators under the assumption that the characteristic roots of $Q\Sigma$ have a known lower bound.

Received August 1976; revised November 1976.

[1] Research supported by National Science Foundation Grant # MCS 76-06627.

[2] Research supported by National Science Foundation Grant # MCS 76-10817.

[3] Research supported by National Science Foundation Grant # MPS 72-05075-A02.

[4] Research supported by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant AFOSR-72-2350C at Purdue University.

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

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

In this paper the fundamental problem of completely unknown Σ will be considered. It will be assumed that an estimate W of Σ is available, where W has a Wishart distribution with parameter Σ and n degrees of freedom, and is independent of X. Let $\text{ch}_{\min}(A)$ denote the minimum characteristic root of A, and define

$$\alpha = [(n - p - 1)\,\text{ch}_{\max}(Q^{-1}W^{-1})]^{-1} = \text{ch}_{\min}(QW)/(n - p - 1)\,.$$

The estimators considered in this paper will be of the form

$$\delta^c(X, W) = \left(I - \frac{c\alpha Q^{-1}W^{-1}}{X^t W^{-1} X}\right) X\,, \tag{1.1}$$

where c is a positive constant. For known Σ, estimators of this form (with $(n - p - 1)W^{-1}$ replaced by Σ^{-1}) were shown to be minimax in Bock (1974) and Berger (1976b), providing $0 \leqq c \leqq 2(p - 2)$. In this paper δ^c is shown to be minimax for

$$0 \leqq c \leqq c_{n,p}\,,$$

where the $c_{n,p}$ are solutions to equation (2.17), and are numerically calculated in Table 1 for certain values of n and p.

TABLE 1
Values of $c_{n,p}$

p	n = 8	10	12	14	16	18	20	25	30
3	.14	.41	.72	.88	1.03	1.10	1.23	1.51	1.53
4	.65	1.37	1.88	2.27	2.42	2.60	2.81	3.07	3.12
5		1.83	2.85	3.37	3.80	4.02	4.26	4.78	4.87
6		1.71	3.32	4.27	4.81	5.33	5.66	6.36	6.50
7			3.42	4.99	5.78	6.42	6.96	7.92	8.14
8			2.50	5.15	6.57	7.64	8.19	9.24	9.84
9				4.50	7.02	8.40	9.22	10.60	11.28
10				2.61	6.79	8.90	10.25	11.98	12.84
11					5.78	9.15	10.84	13.14	14.24
12					2.73	8.42	11.10	14.20	15.65
13						7.11	11.09	15.48	17.15
14						2.43	9.70	15.74	18.44
15							7.93	16.61	19.51
16							2.26	16.67	20.62
17								16.67	21.56
18								16.34	22.38
19									22.83
20									23.47

2. Minimaxity of δ^c. The notation $E(Z)$ will be used for the expectation of Z. Subscripts on E will refer to parameter values, while superscripts on E will refer to the random variables with respect to which the expectation is to be taken. When obvious, subscripts and superscripts will be omitted.

For an estimator, δ, define the risk function

$$R(\delta, \theta, \Sigma) = E_{\theta,\Sigma}^{X,W}[L(\delta(X, W), \theta, \Sigma)] .$$

For notational convenience define $n^* = (n - p - 1)$ and

$$\Delta_c = \Delta_c(\theta, \Sigma) = \operatorname{tr}(Q\Sigma)[R(\delta^c, \theta, \Sigma) - R(\delta^0, \theta, \Sigma)] .$$

The estimator δ^c is clearly minimax (and as good as or better than δ^0) providing $\Delta_c(\theta, \Sigma) \leqq 0$ for all θ and Σ.

Expanding the quadratic loss L for δ^c verifies that

$$\Delta_c = -2E\left[\frac{c\alpha(X-\theta)^t W^{-1}X}{X^t W^{-1}X}\right] + E\left[\frac{c^2\alpha^2 X^t W^{-1}Q^{-1}W^{-1}X}{(X^t W^{-1}X)^2}\right]. \tag{2.1}$$

As in Berger (1976b) an integration by parts with respect to the X_i gives

$$E\left[\frac{(X-\theta)^t W^{-1}X}{X^t W^{-1}X}\right] = E\left[\frac{\operatorname{tr}(\Sigma W^{-1})}{X^t W^{-1}X} - \frac{2X^t W^{-1}\Sigma W^{-1}X}{(X^t W^{-1}X)^2}\right].$$

(The usefulness of such an integration by parts was earlier noticed by Stein (1973).) Thus (2.1) becomes

$$\Delta_c = -E\left[\frac{c\alpha}{(X^t W^{-1}X)}\left\{2\operatorname{tr}(\Sigma W^{-1}) - \frac{4X^t W^{-1}\Sigma W^{-1}X}{X^t W^{-1}X} - \frac{c\alpha X^t W^{-1}Q^{-1}W^{-1}X}{X^t W^{-1}X}\right\}\right]. \tag{2.2}$$

Note that

$$\frac{\alpha X^t W^{-1}Q^{-1}W^{-1}X}{X^t W^{-1}X} \leqq \frac{\alpha}{\operatorname{ch}_{\min}(QW)} = \frac{1}{n^*}.$$

Using this in (2.2) gives

$$\Delta_c \leqq -E\left[\frac{\alpha c}{(X^t W^{-1}X)}\left\{2\operatorname{tr}(\Sigma W^{-1}) - \frac{4X^t W^{-1}\Sigma W^{-1}X}{X^t W^{-1}X} - \frac{c}{n^*}\right\}\right]. \tag{2.3}$$

In this expression, perform the change of variables

$$Y = \Sigma^{-\frac{1}{2}}X, \qquad V = \Sigma^{-\frac{1}{2}}W\Sigma^{-\frac{1}{2}} .$$

Note that V is now Wishart with parameter I and n degrees of freedom, and that $\alpha = \operatorname{ch}_{\min}(\Sigma^{\frac{1}{2}}Q\Sigma^{\frac{1}{2}}V)/n^*$. Clearly (2.3) becomes

$$\Delta_c \leqq -E\left[\frac{\alpha c}{(Y^t V^{-1}Y)}\left\{2\operatorname{tr}(V^{-1}) - \frac{4Y^t V^{-2}Y}{Y^t V^{-1}Y} - \frac{c}{n^*}\right\}\right]. \tag{2.4}$$

For convenience, define

$$\beta = \operatorname{ch}_{\min}(Q\Sigma), \qquad Z = Y/|Y|, \qquad \text{and} \qquad \Sigma^* = \Sigma^{\frac{1}{2}}Q\Sigma^{\frac{1}{2}}/\beta .$$

Note that $\operatorname{ch}_{\min}(\Sigma^*) = 1$. Line (2.4) can then be rewritten

$$\Delta_c \leqq \frac{-\beta c}{n^*} E^Y\left[\frac{1}{|Y|^2} E^V\left\{\frac{\operatorname{ch}_{\min}(\Sigma^* V)}{(Z^t V^{-1}Z)}\left(2\operatorname{tr}(V^{-1}) - \frac{4Z^t V^{-2}Z}{Z^t V^{-1}Z} - \frac{c}{n^*}\right)\right\}\right]. \tag{2.5}$$

To show that $\Delta_c \leqq 0$, it suffices to show for all $Z \in U_p$ (the unit p-sphere) and all Σ^* with $\mathrm{ch}_{\min}(\Sigma^*) = 1$, that the following inequality holds:

$$E^V \left\{ \frac{\mathrm{ch}_{\min}(\Sigma^* V)}{(Z^t V^{-1} Z)} \left[2\,\mathrm{tr}\,(V^{-1}) - \frac{4Z^t V^{-2} Z}{Z^t V^{-1} Z} - \frac{c}{n^*} \right] \right\} \geqq 0 . \tag{2.6}$$

(Note that the distribution of V does not depend on Z or on Σ^*.)

Let Γ be a $p \times p$ orthogonal matrix such that $\Gamma Z = (1, 0, \cdots, 0)^t$. Define $V^* = \Gamma V \Gamma^t$ and $\Sigma_Z = \Gamma \Sigma^* \Gamma^t$. Clearly V^* is also Wishart (I) and $\mathrm{ch}_{\min}(\Sigma_Z) = 1$. For convenience, let v_1 denote the (1, 1) element of $(V^*)^{-1}$, v_2 denote the (1, 1) element of $(V^*)^{-2}$, and let

$$\rho(V^*) = [2\,\mathrm{tr}\,\{(V^*)^{-1}\} - 4v_2/v_1] .$$

It is straightforward to verify that under the above change of variables for V, (2.6) becomes

$$E^{V^*} \left\{ \frac{\mathrm{ch}_{\min}(\Sigma_Z V^*)}{v_1} \left[\rho(V^*) - \frac{c}{n^*} \right] \right\} \geqq 0 . \tag{2.7}$$

(Note that for $p \leqq 2$, $\rho(V^*)$ will be negative with considerable probability. Hence no solutions, c, to (2.7) could be obtained for $p \leqq 2$. This is as would be expected in analogy to the known variance case.)

Since $\mathrm{ch}_{\min}(\Sigma_Z) = 1$, it is clear that

$$\mathrm{ch}_{\min}(\Sigma_Z V^*) \geqq \mathrm{ch}_{\min}(V^*) . \tag{2.8}$$

Also if $a \in U_p$ (i.e., $|a| = 1$) then

$$\mathrm{ch}_{\min}(\Sigma_Z V^*) \leqq a^t \Sigma_Z^{\frac{1}{2}} V^* \Sigma_Z^{\frac{1}{2}} a .$$

Choosing a to be a^1, the characteristic vector of the root 1 of $\Sigma_Z^{\frac{1}{2}}$, it follows that

$$\mathrm{ch}_{\min}(\Sigma_Z V^*) \leqq (a^1)^t V^* a^1 . \tag{2.9}$$

For convenience define

$$\Omega_c = \{V^* : \rho(V^*) < c/n^*\} ,$$

let $\bar{\Omega}_c$ denote the complement of Ω_c, and let $I_A(V^*)$ denote the usual indicator function on A. Using (2.8) and (2.9), it then follows that (2.7) will hold (and δ^c will be minimax) if

$$\begin{aligned} E^{V^*} \Bigg\{ & \frac{(a^1)^t V^* a^1}{v_1} \left[\rho(V^*) - \frac{c}{n^*} \right] I_{\Omega_c}(V^*) \\ & + \frac{\mathrm{ch}_{\min}(V^*)}{v_1} \left[\rho(V^*) - \frac{c}{n^*} \right] I_{\bar{\Omega}_c}(V^*) \Bigg\} \geqq 0 \end{aligned} \tag{2.10}$$

for all $a^1 \in U_p$.

To simplify this expression further, let

$$T = \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & S \\ 0 & \end{pmatrix} ,$$

where S is a $(p-1)\times(p-1)$ orthogonal matrix such that

$$Ta^1 = (b, (1-b^2)^{\frac{1}{2}}, 0, \cdots, 0)^t \qquad -1 \leqq b \leqq 1 .$$

In (2.10), performing the change of variables $V = TV^*T^t$ (again Wishart (I)) then gives as the condition for minimaxity

$$(2.11) \qquad E^V\left\{\frac{(Ta^1)^tV(Ta^1)}{v_1}\left[\rho(V)-\frac{c}{n^*}\right]I_{\Omega_c}(V) + \frac{\mathrm{ch}_{\min}(V)}{v_1}\left[\rho(V)-\frac{c}{n^*}\right]I_{\bar{\Omega}_c}(V)\right\} \geqq 0$$

for all $a^1 \in U_p$. (Note that $v_1 = (V^{*-1})_{11} = (T^tV^{-1}T)_{11} = (V^{-1})_{11}$ and likewise $v_2 = (V^{-2})_{11}$.) The inequality (2.11) can be rewritten

$$(2.12) \qquad c \leqq \frac{n^*E^V\{\rho(V)v_1^{-1}[(Ta^1)^tV(Ta^1)I_{\Omega_c}(V) + \mathrm{ch}_{\min}(V)I_{\bar{\Omega}_c}(V)]\}}{E^V\{v_1^{-1}[(Ta^1)^tV(Ta^1)]I_{\Omega_c}(V) + \mathrm{ch}_{\min}(V)I_{\bar{\Omega}_c}(V)]\}} .$$

Note that

$$(Ta^1)^tV(Ta^1) = b^2(V_{11}-V_{22}) + b(1-b^2)^{\frac{1}{2}}(V_{12}+V_{21}) + V_{22} .$$

Hence defining

$$\begin{aligned}
\tau_0(c) &= E^V\{\rho(V)v_1^{-1}[V_{22}I_{\Omega_c}(V) + \mathrm{ch}_{\min}(V)I_{\bar{\Omega}_c}(V)]\} ,\\
\tau_1(c) &= E^V\{\rho(V)v_1^{-1}(V_{11}-V_{22})I_{\Omega_c}(V)\} ,\\
\tau_2(c) &= E^V\{\rho(V)v_1^{-1}(V_{12}+V_{21})I_{\Omega_c}(V)\} ,\\
\tau_0'(c) &= E^V\{v_1^{-1}[V_{22}I_{\Omega_c}(V) + \mathrm{ch}_{\min}(V)I_{\bar{\Omega}_c}(V)]\} ,\\
\tau_1'(c) &= E^V\{v_1^{-1}(V_{11}-V_{22})I_{\Omega_c}(V)\} , \qquad \text{and}\\
\tau_2'(c) &= E^V\{v_1^{-1}(V_{12}+V_{21})I_{\Omega_c}(V)\} ,
\end{aligned}$$

it is clear that (2.12), the condition for minimaxity, can be rewritten

$$(2.13) \qquad c \leqq \frac{n^*[\tau_0(c) + \tau_1(c)b^2 + \tau_2(c)b(1-b^2)^{\frac{1}{2}}]}{\tau_0'(c) + \tau_1'(c)b^2 + \tau_2'(c)b(1-b^2)^{\frac{1}{2}}}$$

for all $-1 \leqq b \leqq 1$. Note, however, that $\tau_2(c) = \tau_2'(c) = 0$. This follows from consideration of

$$V^* = AVA ,$$

where A is diagonal with all ones except the (2, 2) element which is -1. It is easy to check that $\rho(V^*) = \rho(V)$, $[(V^*)^{-1}]_{11} = [V^{-1}]_{11}$, $(V^*)_{12} = -V_{12}$, and $(V^*)_{21} = -V_{21}$. Noting that $f(V) = f(V^*)$ (where f is the Wishart density), it is thus clear that $\tau_2 = \tau_2' = 0$. Finally, defining $\tilde{b} = (b, (1-b^2)^{\frac{1}{2}})$,

$$A(c) = \begin{pmatrix} \tau_0(c)+\tau_1(c) & 0 \\ 0 & \tau_0(c)\end{pmatrix}, \qquad \text{and} \qquad B(c) = \begin{pmatrix} \tau_0'(c)+\tau_1'(c) & 0 \\ 0 & \tau_0'(c)\end{pmatrix} .$$

line (2.13) becomes

$$(2.14) \qquad c \leqq \frac{n^*\tilde{b}^tA(c)\tilde{b}}{\tilde{b}^tB(c)\tilde{b}}, \qquad \text{for all} \quad -1 \leqq b \leqq 1 .$$

Now for fixed $\tilde{b}$, the nonnegative solutions to (2.14) lie in an interval $0 \leqq c \leqq c_{\tilde{b}}$. This can most easily be seen by looking at (2.11) (an expression equivalent to (2.14)) and noting that the left-hand side is decreasing in c. Thus defining

$$c_{n,p} = \inf_{-1 \leqq b \leqq 1} c_{\tilde{b}} ,$$

it follows that if

$$0 \leqq c \leqq c_{n,p} \tag{2.15}$$

then (2.14) will be satisfied for all $-1 \leqq b \leqq 1$, and hence δ^c will be minimax.

To get a more explicit equation for $c_{n,p}$, note that $B(c)$ is positive definite. Hence from (2.14) it follows that

$$c \leqq n^* \operatorname{ch}_{\min} [B(c)^{-1}A(c)] . \tag{2.16}$$

Thus (2.15) $\Rightarrow$ (2.14) for all $-1 \leqq b \leqq 1 \Rightarrow$ (2.16). It is also clear that the reverse implications hold, so that

$$\{c : 0 \leqq c \leqq c_{n,p}\} = \{c : c \leqq n^* \operatorname{ch}_{\min} [B(c)^{-1}A(c)]\} .$$

It is also easy to check that

$$\begin{aligned} c_{n,p} &= n^* \operatorname{ch}_{\min} [B(c_{n,p})^{-1}A(c_{n,p})] , \\ c &< n^* \operatorname{ch}_{\min} [B(c)^{-1}A(c)] \qquad \text{if} \quad 0 \leqq c < c_{n,p} , \\ c &> n^* \operatorname{ch}_{\min} [B(c)^{-1}A(c)] \qquad \text{if} \quad c > c_{n,p} . \end{aligned} \tag{2.17}$$

Hence $c_{n,p}$ is the unique solution to

$$c = n^* \operatorname{ch}_{\min} [B(c)^{-1}A(c)] = \min \left\{ \frac{\tau_0(c) + \tau_1(c)}{\tau_0'(c) + \tau_1'(c)} , \frac{\tau_0(c)}{\tau_0'(c)} \right\} . \tag{2.18}$$

As there appeared to be little hope of analytically obtaining solutions to (2.18), the computer was used to numerically compute the solutions. For a given n and p, the values of the $\tau_i(c)$ and $\tau_i'(c)$ (and hence $A(c)$ and $B(c)$) were calculated by Monte Carlo methods using 4000 generations of V (for $n = 8$) to 1000 generations of V (for $n = 30$). (Unfortunately, a larger number of generations could not be used due to the considerable expense of generating V and performing the calculations involving V^{-1}.) The resulting estimated solution $c_{n,p}$, to (2.18) was found by using the relationshps in (2.17) to obtain a sequence of c^i converging to the solution. The standard deviations of these simulated solutions ranged from about .02 (for $p = 3$) to about .1 (for $n - p = 4$).

3. Comments.

1. The values $c_{n,p}$ are not the largest values of c for which δ^c is minimax. Approximations were made in the proof in lines (2.8) and (2.9) (and to a lesser degree in the passage from (2.2) to (2.3)) which resulted in a smaller than necessary upper bound. If one could somehow determine the "least favorable" matrix Σ_Z in (2.7), the approximations could be improved.

2. The estimators δ^c have a singularity as $X \to 0$. There are numerous ways of eliminating the singularity, one of the simplest being used in the following estimator:

$$\delta^{*c}(X, W) = \left(I - \frac{\min(n^* X^t W^{-1} X, c)\alpha Q^{-1} W^{-1}}{X^t W^{-1} X}\right) X .$$

Through analogy with the known Σ situation, it seems quite likely that δ^{*c} is itself minimax (for $0 \leqq c \leqq c_{n,p}$) and considerably better than δ^c.

3. If the linear restriction $R\theta = r^0$ is thought to hold, where R is an $(m \times p)$ matrix of rank m and r^0 is an $(m \times 1)$ vector, then the estimators δ^c and δ^{*c} can be modified so that their regions of significant risk improvement coincide with the linear restriction. Indeed, defining $Y = RX - r^0$, $W^* = RWR^t$ and $\alpha^* = \mathrm{ch}_{\min}[(RQ^{-1}R^t)^{-1}W^*]/(n - m - 1)$, Theorem 2 of Berger and Bock (1976 b) can be used to show that

$$\delta_R^c = X - c\alpha^* Q^{-1} R^t (W^*)^{-1} Y / [Y^t (W^*)^{-1} Y]$$

is minimax if $0 \leqq c \leqq c_{n,m}$. The appropriate modification of δ^{*c} is the above estimator with c replaced by $\min\{(n - m - 1)Y^t(W^*)^{-1}Y, c\}$.

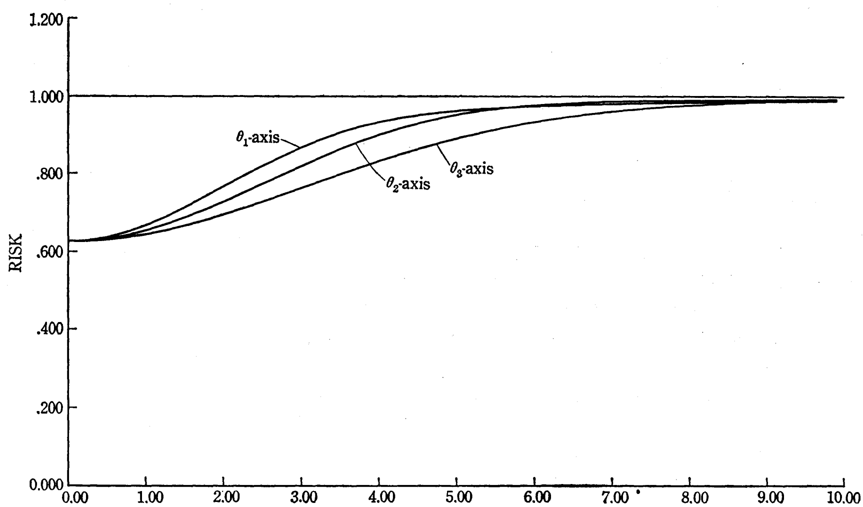

FIG. 1

4. If $(Q\Sigma)$ has a characteristic root considerably smaller than the other characteristic roots, then $\mathrm{ch}_{\min}(Q\Sigma)$ will be small compared to $\mathrm{tr}(Q\Sigma)$. From the definition of $\Delta_c(\theta, \Sigma)$ and line (2.2), it is apparent that the improvement obtained in using δ^c will be quite small. The estimator, δ^c, will therefore perform best when $(Q\Sigma)$ has no exceptionally small roots. (If it is suspected that a coordinate X_i might give rise to an exceptionally small root of $(Q\Sigma)$, it would probably pay to eliminate that coordinate in the construction of δ^c, providing of course that there are at least three coordinates left.)

5. As an example of the type of improvement that can be obtained using the suggested estimator, the risk function of δ^{*c} was numerically calculated for the situation $p = 4$, $n = 16$, $Q = I$, Σ diagonal with elements (2, 3, 4, 4), and $c = c_{n,p} = 2.42$. The risk function is given in Figure 1 along the coordinate axes. (The constant line $y = 1.000$ is, of course, the risk of the usual estimator δ^0.)

REFERENCES

[1] ALAM, KHURSHEED (1975). Minimax and admissible minimax estimators of the mean of a multivariate normal distribution for unknown covariance matrix. *J. Multivariate Anal.* **5** 83–95.

[2] BARANCHIK, A. J. (1970). A family of minimax estimators of the mean of a multivariate normal distribution. *Ann. Statist.* **41** 642–645.

[3] BERGER, J. (1976a). Admissible minimax estimation of a multivariate normal mean with arbitrary quadratic loss. *Ann. Statist.* **4** 223–226.

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

[5] BERGER, J. and BOCK, M. E. (1976a). Combining independent normal mean estimation problems with unknown variances. *Ann. Statist.* **4** 642–648.

[6] BERGER, J. and BOCK, M. E. (1976b). Improved minimax estimators of normal mean vectors for certain types of covariance matrices. To appear in *Statistical Decision Theory and Related Topics II*. S. S. Gupta and D. S. Moore (eds.). Academic Press, New York.

[7] BERGER, J. and BOCK, M. E. (1976c). Eliminating singularities of Stein-type estimators of location vectors. *J. Roy. Statist. Soc. Ser. B* **38** 166–170.

[8] BHATTACHARYA, P. K. (1966). Estimating the mean of a multivariate normal population with general quadratic loss function. *Ann. Math. Statist.* **37** 1819–1927.

[9] BOCK, M. E. (1974). Certain minimax estimators of the mean of a multivariate normal distribution. Ph. D. thesis, Univ. of Illinois.

[10] BOCK, M. E. (1975). Minimax estimators of the mean of a multivariate normal distribution. *Ann. Statist.* **3** 209–218.

[11] EFRON, B. and MORRIS, C. (1973). Stein's estimation rule and its competitors—an empirical Bayes approach. *J. Amer. Statist. Assoc.* **68** 117–130.

[12] EFRON, B. and MORRIS, C. (1976). Families of minimax estimators of the mean of a multivariate normal distribution. *Ann. Statist.* **4** 11–21.

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

[14] HUDSON, M. (1974). Empirical Bayes estimation. Technical Report #58, Stanford Univ.

[15] JAMES, W. and STEIN, C. (1960). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 361–379, Univ. of California Press.

[16] LIN, PI-ERH and TSAI, HUI-LIANG (1973). Generalized Bayes minimax estimators of the multivariate normal mean with unknown covariance matrix. *Ann. Statist.* **1** 142–145.

[17] RAO, C. RADHAKRISHNA (1976). Simultaneous estimation of parameters—a compound decision problem. *Statistical Decision Theory and Related Topics II*. S. S. Gupta and D. S. Moore (eds.). Academic Press, New York.

[18] SHINOZAKI, N. (1974). A note on estimating the mean vector of a multivariate normal distribution with general quadratic loss function. *Keio Engrg. Rep.* **27** 105–112.

[19] STEIN, C. (1955). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **1** 197–206, Univ. of California Press.

[20] STEIN, C. (1973). Estimation of the mean of a multivariate distribution. *Proc. Prague Symp. Asymptotic Statist.* 345–381.

[21] STRAWDERMAN, W. E. (1973). Proper Bayes minimax estimators of the multivariate normal mean vector for the case of common unknown variances. *Ann. Statist.* **1** 1189–1194.

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

DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903