ON THE EXISTENCE OF PROPER BAYES MINIMAX ESTIMATORS OF THE MEAN OF A MULTIVARIATE NORMAL DISTRIBUTION

WILLIAM E. STRAWDERMAN
STANFORD UNIVERSITY

1. Introduction

Consider the problem of estimating the mean of a multivariate normal distribution with convariance matrix the identity and sum of squared errors loss. In an earlier paper [5] the author showed that if the dimension p is 5 or greater, then proper Bayes minimax estimators do exist. We review this result briefly in Section 2.

The main purpose of the present paper is to show that for p equal to 3 or 4, there do not exist spherically symmetric proper Bayes minimax estimators. The author has been unable, thus far, to disprove the existence of a nonspherical proper Bayes minimax estimator for p equal 3 or 4. Of course, for $p = 1, 2$, the usual estimator $\overline{X}$ is unique minimax but not proper Bayes.

In Section 3 we derive bounds for the possible bias of a minimax estimator. This result should be of some interest independent of its use in proving the main result of the paper. Section 4 is devoted to the proof of the main result.

2. The case of five and higher dimensions

In [5] the author produced a class of estimators for $p \geqq 5$ which are proper Bayes minimax. This was done without loss of generality, in the case of a single observation. For completeness we briefly describe this result.

Let X be a p dimensional random vector distributed according to the multinormal distribution with mean θ and covariance matrix I.

The prior distribution on θ is given as follows: conditional on λ, where $0 < \lambda \leqq 1$, let the distribution of θ be multinormal with mean zero and covariance matrix $[(1 - \lambda)/\lambda]I$. The unconditional density of λ with respect to Lebesgue measure is given by $\lambda^{-a}/(1 - a)$ for any a where $0 \leqq a < 1$.

The Bayes estimator with respect to the above prior distribution on θ is given by

$$(1) \qquad \delta(X) = \left[1 - \left(\frac{p + 2 - 2a}{\|X\|^2} - \frac{2 \exp\{-\frac{1}{2}\|X\|^2\}}{\|X\|^2 \int_0^1 \lambda^{p/2-a} \exp\{-\lambda\|X\|^2\}}\right)\right] X.$$

It then follows from a simple extension of a result due to Baranchik [1], [2] that the estimator $\delta(X)$, which is proper Bayes by definition, is minimax for $p \geqq 6$. In addition for $\frac{1}{2} \leqq a < 1$, $\delta(X)$ is minimax for $p = 5$.

3. The bias of spherically symmetric minimax estimators

In this section we derive upper and lower bounds on the possible bias of spherically symmetric minimax estimators, that is, minimax estimators of the form $\delta(X) = h(\|X\|^2)X$. For such estimators, the ith component of the bias is given by $E_\theta h(\|X\|^2)X_i - \theta_i = -\varphi(\|\theta\|^2)\theta_i$. Using the multivariate information inequality as in Stein [3], p. 202, we have

$$\begin{aligned}(2)\qquad & E_\theta(\|h(\|X\|^2)X - \theta\|^2) \\ & \geqq p + \|\theta\|^2 \varphi^2(\|\theta\|^2) - 2p\varphi(\|\theta\|^2) - 4\|\theta\|^2 \varphi'(\|\theta\|^2).\end{aligned}$$

If $\delta(X) = h(\|X\|^2)X$ is minimax, (2) becomes

$$(3)\qquad 0 \geqq \|\theta\|^2 \varphi^2(\|\theta\|^2) - 2p\varphi(\|\theta\|^2) - 4\|\theta\|^2 \varphi'(\|\theta\|^2).$$

Letting $t = \|\theta\|^2$ and $\psi(t) = t\,\varphi(t)$, we have

$$(4)\qquad 0 \geqq \psi^2(t)/t - 2p\psi(t)/t - 4\psi'(t) - 4\psi(t)/t$$

or

$$(5)\qquad 4\psi'(t) \geqq \psi(t)[\psi(t) - (2p - 4)]/t.$$

We prove the following lemma.

LEMMA 1. *For* $p \geqq 2$, $0 \leqq \psi(t) \leqq 2(p - 2)$, $0 < t < \infty$.

PROOF. Suppose that for some $t_0 > 0$, $\psi(t_0) < 0$. This implies by (5) that $\psi(t) < 0$ for all $0 < t < t_0$, and hence

$$(6)\qquad \psi'(t)[\psi(t)(\psi(t) - 2p - 4)]^{-1} \geqq 1/4t.$$

Integrating from $t(0 < t < t_0)$ to t_0

$$(7)\qquad \log\left[\frac{-\psi(t_0) + (2p - 4)}{-\psi(t_0)}\right]\left[\frac{-\psi(t)}{-\psi(t) + (2p - 4)}\right] \geqq \frac{p - 2}{2}\log\left(\frac{t_0}{t}\right).$$

As $t \to 0$ the right side of (7) approaches $+\infty$ while the left side remains bounded, a contradiction. Hence it cannot happen that $\psi(t_0) < 0$.

Assume next for some $t_0 > 0$ that $\psi(t_0) > 2p - 4$. From (5) it follows that $\psi(t) > 2p - 4$ for all $t > t_0$. Proceeding as above except integrating the inequality (6) from t_0 to t, we obtain

$$(8)\qquad \log\left[\frac{\psi(t) - (2p - 4)}{\psi(t)}\right]\left[\frac{\psi(t_0)}{\psi(t_0) - (2p - r)}\right] \geqq \frac{p - 2}{2}\log\left(\frac{t_0}{t}\right).$$

Again the right side approaches $+\infty$ as t approaches $+\infty$ while the left side remains bounded. This contradiction establishes the lemma.

From the lemma and the definition of $\psi(t)$ we get immediately the following result.

THEOREM 1. *If* $\delta(x) = h(\|X\|^2)X$ *is a minimax estimator, and*

$$E_\theta(h(\|X\|^2)X_i - \theta_i) = -\varphi(\|\theta\|^2)\theta_i. \tag{9}$$

Then $0 \leqq \varphi(\|\theta\|^2) \leqq 2(p-2)/\|\theta\|^2$.

In other words, the bias of a spherically symmetric minimax estimator is always towards the origin but not by a factor larger than $2(p-2)/\|\theta\|2$.

4. The nonexistence of spherically symmetric proper Bayes minimax estimators in three and four dimensions

In this section we apply Theorem 1 to show that spherically symmetric proper Bayes estimators cannot be minimax for $p = 3$ and 4.

We note first that if $\delta(X)$ is a generalized Bayes estimator relative to the prior $dG(\theta)$,

$$\delta_i(X) - X_i = \frac{\partial}{\partial X_i} \log \int_{R^p} \exp\{-\tfrac{1}{2}\|X - \theta\|^2\}\, dG(\theta). \tag{10}$$

The ith components of the $\delta(\cdot)$ is given by

$$\begin{aligned}(11)\quad & A \int_{R^p} \left[\frac{\partial}{\partial x_i} \log \int_{R^p} \exp\{-\tfrac{1}{2}\|x-\theta\|^2\}\, dG(\theta)\right] \exp\{-\tfrac{1}{2}\|x - u\|^2\} \prod_{i=1}^{p} dx_i \\ &= -A \int_{R^p} \left[\log \int_{R^p} \exp\{-\tfrac{1}{2}\|x-\theta\|^2\}\, dG(\theta)\right] \frac{\partial}{\partial x_i} \exp\{-\tfrac{1}{2}\|x - u\|^2 \prod_{i=1}^{p} dx_i \\ &= \frac{\partial}{\partial u_i} A \int_{R^p} \left[\log \int_{R^p} \exp\{-\tfrac{1}{2}\|x-\theta\|^2\}\, dG(\theta)\right] \exp\{-\tfrac{1}{2}\|x - u\|^2\} \prod_{i=1}^{p} dx_i\end{aligned}$$

where $A = (2\pi)^{-p/2}$.

We have used integration by parts with respect to x_i to attain the first equality. To get the second equality we use the fact that the partial derivative of $\exp\{\frac{1}{2}\|x - u\|^2\}$ with respect to x_i is the negative of the partial derivative of the same quantity with respect to u_i. The partial with respect to u_i is then taken outside the integral.

If we now assume $\delta(\cdot)$ is minimax and spherically symmetric (which is equivalent to $dG(\theta)$ being orthogonally invariant [4], p. 42), we have from Theorem 1 that for $u_i > 0$

$$\begin{aligned}(12)\quad & -\varphi(\|u\|^2)u_i \\ &= \frac{\partial}{\partial u_i} A \int_{R^p} \left[\log \int_{R^p} \exp\{-\tfrac{1}{2}\|x-\theta\|^2\}\, dG(\theta)\right] \\ &\qquad\qquad \exp\{-\tfrac{1}{2}\|x - u\|^2\} \prod_{i=1}^{p} dx_i \\ &\geqq \frac{-2(p-2)u_i}{\|u\|^2}.\end{aligned}$$

For any vector $u = (u_1, u_2, \cdots, u_p)$, $\|u\| \geqq \varepsilon > 0$, $u_i \geqq 0$, we define vectors u^i, $i = 0, \cdots, p$ as follows. Let $u^0 = (u_1^0, u_2^0, \cdots, u_p^0)$ be any vector such that $0 \leqq u_i^0 \leqq u_i$ and $\|u^0\| = \eta < \varepsilon$, $\eta > 0$ fixed. Next let

$$u_i^j = \begin{cases} u_i & \text{if } j \geqq i, \\ u_i^0 & \text{if } j < i, \end{cases} \tag{13}$$

and $u^j = (u_1^j, u_2^j, \cdots, u_p^j)$, $j = 1, \cdots, p$.

Note that we are merely changing the jth coordinate of u^{j-1} to that of u by the above construction. Hence $u^p = u$. Integrating the expression (12) from u^{i-1} to u^i with respect to u_i, $i = 1, \cdots, p$, and adding the results, we obtain by collapsing successive terms,

$$\begin{aligned} (14) \quad & A \int_{R^p} \left[\log \int_{R^p} \exp\{-\tfrac{1}{2}\|x - \theta\|^2\}\, dG(\theta)\right] \exp\{-\tfrac{1}{2}\|x - u\|^2\} \prod_{i=1}^p dx_i \\ & -A \int_{R^p} \left[\log \int_{R^p} \exp\{-\tfrac{1}{2}\|x - \theta\|^2\}\, dG(\theta)\right] \exp\{-\tfrac{1}{2}\|x - u^0\|^2\} \prod_{i=1}^p dx_i \\ & \geqq -2(p-2)\log\|u\| - 2(p-2)\log\|u^0\|. \end{aligned}$$

Because of the orthogonal invariance of $dG(\theta)$ and the fact that $\|u^0\| = \eta$, the two terms in the above inequality which depend on u^0 do so only through η. Hence, these two terms are constants. By Jensen's inequality applied to (14) it follows that

$$\begin{aligned} (15) \quad & \log A \left[\int_{R^p} \left(\int_{R^p} \exp\{-\tfrac{1}{2}\|x - \theta\|^2\}\, dG(\theta)\right) \exp\{-\tfrac{1}{2}\|x - u\|^2\} \prod_{i=1}^p dx_i\right] \\ & \geqq -2(p-2)\log\|u\| + c \end{aligned}$$

or

$$\begin{aligned} (16) \quad & A \int_{R^p} \left(\int_{R^p} \exp\{-\tfrac{1}{2}\|x - \theta\|^2\}\, dG(\theta)\right) \exp\{-\tfrac{1}{2}\|x - u\|^2\} \prod_{i=1}^p dx_i \\ & \geqq e^c \|u\|^{-2(p-2)}. \end{aligned}$$

Note that by the orthogonal invariance of G, the above inequality holds for all u such that $\|u\| \geqq \varepsilon$.

The left side of (16) is essentially the (density of) the convolution of the standard normal with the convolution of $dG(\theta)$ and the standard normal. Hence the quantity on the left represents the Radon-Nikodym derivative (with respect to Lebesgue measure on R^p) of a measure which is finite if and only if $\int_{R^p} dG(\theta) < \infty$. However, integrating the right side of (16) over the sphere $\|u\| \geqq \varepsilon$ we see that the result can only be finite if $2(p-2) - (p-1) > 1$, or equivalently if $p > 4$. We therefore have the following result.

THEOREM 2. *Let $\delta(X)$ be a spherically symmetric minimax estimator which is generalized Bayes with respect to the (generalized) prior $dG(\theta)$. If $p \leqq 4$, $dG(\theta)$ cannot be a proper prior distribution.*

5. Remarks

We have shown that there do not exist spherically symmetric proper Bayes minimax estimators in four or lower dimensions. In a previous paper we demonstrated the existence of such estimators for $p \geqq 5$. We are thus far unable to rule out the possibility that nonspherically symmetric proper Bayes minimax estimators exist in three and four dimensions although it seems highly unlikely that such estimators do exist.

Also we have been unable thus far to say anything concrete about the situation where some part of the covariance structure of the problem is unknown.

L. Brown has recently found a proof of the nonexistence of nonspherically symmetric proper Bayes estimators for $p = 3$ and 4.

◇ ◇ ◇ ◇ ◇

The author wishes to thank L. Brown for bringing this problem to his attention and C. Stein for much helpful discussion.

REFERENCES

[1] A. J. BARANCHIK, "Multiple regression and estimation of the mean of a multivariate normal distribution," Stanford University, Technical Report No. 51 (1964).

[2] ———, "A family of minimax estimators of the mean of a multivariate normal distribution," *Ann. Math. Statist.*, Vol. 41 (1970), pp. 642–645.

[3] CHARLES STEIN, "Inadmissibility of the usual estimator for the mean of a multivariate normal distribution," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1955, pp. 197–206.

[4] W. E. STRAWDERMAN, "Generalized Bayes estimators and admissibility of estimators of the mean vector of a multivariate normal distribution with quadratic loss," Ph.D. thesis, Rutgers University, 1969.

[5] ———, "Proper Bayes minimax estimators of the multivariate normal mean," *Ann. Math. Statist.*, Vol. 42 (1971), pp. 385–388.