The Annals of Statistics
1976, Vol. 4, No. 6, 1065–1076

INADMISSIBILITY RESULTS FOR THE BEST INVARIANT ESTIMATOR OF TWO COORDINATES OF A LOCATION VECTOR[1]

BY JAMES O. BERGER

Purdue University

Let $X = (X_1, X_2, X_3)$ be a random vector with density $f(x - \theta)$, where $\theta = (\theta_1, \theta_2, \theta_3)$ is unknown. It is desired to estimate (θ_1, θ_2) using an estimator $(\delta_1(X), \delta_2(X))$, and under a loss function $L(\delta_1 - \theta_1, \delta_2 - \theta_2)$. (Note that θ_3 is a nuisance parameter.) Under certain conditions on f and L, it is shown that the best invariant estimator of (θ_1, θ_2) is inadmissible.

1. Introduction. The problem of inadmissibility of the best invariant estimator of a p-dimensional location vector has received considerable study since Stein (1955) first demonstrated that it could be inadmissible if $p \geqq 3$. For the situation of estimating the full location vector, the answers are now quite complete. Brown (1966) showed for quite general distributions and loss functions that if $p \geqq 3$ then the best invariant estimator is inadmissible, while if $p = 1$ the best invariant estimator is admissible. Brown and Fox (1974) demonstrated that if $p = 2$, then the best invariant estimator is usually admissible. (The above results hold subject only to certain moment conditions and technical assumptions.)

A long outstanding problem in this field has been to determine inadmissibility results when only some coordinates of the location vector are of interest. (The other coordinates are then nuisance parameters.) The two significant problems in this area which cannot be subsumed into the framework of Brown (1966) or Brown and Fox (1974) are (i) estimating one coordinate of a p-dimensional location vector with $p \geqq 3$, and (ii) estimating two coordinates of a p-dimensional location vector with $p \geqq 3$. Problem (i) was recently considered in Berger (1976a) and Berger (1976b). In Berger (1976a) it was shown that if $p \geqq 4$, then the best invariant estimator of one coordinate of a location vector is often inadmissible. In Berger (1976b) it was shown that if $p \leqq 3$, then the best invariant estimator of one coordinate of a location vector is usually admissible. (The conditions needed for the results were fairly restrictive and did leave some questions unanswered.) Portnoy (1975) had earlier obtained similar results for a specially constructed class of distributions and squared error loss.

In this paper, problem (ii) is considered. It is shown that in a wide variety of situations, the best invariant estimator of two coordinates of a three or more

Received September 1975; revised January 1976.

[1] This research was supported by the National Science Foundation under Grant # MPS75-07017.

AMS 1970 *subject classifications*. Primary 62C15; Secondary 62F10, 62H99.

Key words and phrases. Inadmissibility, best invariant estimator, location vector, risk function, loss function.

dimensional location vector is inadmissible. (For simplicity only the three-dimensional case is explicitly considered, since inadmissibility in a three-dimensional subproblem will imply inadmissibility in the full problem.) The results obtained in this paper (and in Berger (1976a) and Berger (1976b)) essentially substantiate conjectures of James and Stein (1960).

Section 2 introduces notation, and states and discusses the assumptions that will be made. Section 3 gives some needed technical lemmas. Section 4 deals with the main inadmissibility theorem. The technique of proof used in the theorem is basically due to Brown (1974). The idea is to approximate the risk functions through the use of Taylor expansions, using a "randomization-of-the-origin" argument to handle the nondominant terms. See Brown (1974) and Berger (1976a) for more general discussions of the relevant techniques used. The proof given in this paper does employ a new idea to significantly reduce the needed technical detail. This idea is discussed in the proof of Theorem 1. The simplified proof is one which should easily generalize to other location parameter problems.

Section 5 of the paper gives two examples of applications of the theory. It should be mentioned that, as in Berger (1976a), the theory will not apply to estimating a normal mean vector. See the beginning of Section 5 for a short discussion of this.

2. Notation and assumptions. Let $X = (X_1, X_2, X_3)$ be a three-dimensional random vector with density $f(x - \theta)$ with respect to Lebesgue measure. Here $\theta = (\theta_1, \theta_2, \theta_3)$ is unknown and it is desired to estimate $\theta^* = (\theta_1, \theta_2)$. Assume that the loss incurred in estimating θ^* by $d^* = (d_1, d_2)$ is $L(d^* - \theta^*)$.

The notation $E[\ ]$ will be used for the expectation of the argument. If the argument is a vector or a matrix, the expectation is to be taken componentwise. Subscripts on E will denote parameter values under which the expectation is to be taken. Superscripts on E will be used to clarify the random variable with respect to which the expectation is being taken. When obvious, subscripts or superscripts will be suppressed.

Under a suitable parameterization of the problem, the best invariant estimator (we assume it exists) will be given by $\delta_0(X) = X^* = (X_1, X_2)$. For an arbitrary estimator $\delta: R^3 \to R^2$ (R^k is k-dimensional Euclidean space), define the risk function

$$R(\delta, \theta) = E_\theta[L(\delta(X) - \theta^*)] .$$

Also define

$$\Delta_\delta(\theta) = R(\delta, \theta) - R(\delta_0, \theta) = E_\theta[L(\delta(X) - \theta^*) - L(X^* - \theta^*)] .$$

It will be shown that, under certain assumptions, there exists an estimator δ for which $\Delta_\delta(\theta) < 0$ for all $\theta \in R^3$, and hence that the best invariant estimator δ_0 is inadmissible.

For notational convenience, partial derivatives of an appropriately differentiable function h will be denoted by

$$h^{(i)}(x) = \frac{\partial}{\partial x_i} h(x) , \qquad h^{(i,j)}(x) = \frac{\partial^2}{\partial x_i \, \partial x_j} h(x) , \quad \text{etc.}$$

Also define

$$m_{ij} = E_0[L^{(i)}(X^*)X_j] , \qquad m_{ijk} = E_0[L^{(i)}(X^*)X_j X_k] ,$$
$$\nabla L(X^*) = (L^{(1)}(X^*), L^{(2)}(X^*))^t , \qquad \text{and} \qquad M = E_0[\nabla L(X^*)X^*] .$$

Note that M is the 2×2 matrix with elements m_{ij} ($i \leqq 2, j \leqq 2$). Let $|x|$ denote the usual Euclidean norm of the vector x, and for $x^* = (x_1, x_2)$ define

$$||x^*|| = x^* M^{-1}(1, 1)^t \qquad \text{and} \qquad |x^*|_\alpha = |x_1|^\alpha + |x_2|^\alpha .$$

(Though $||x^*||$ and $|x^*|_\alpha$ are not necessarily norms, the notation will prove very convenient.) Also for convenience, K will be used as a generic constant throughout the paper.

The following assumptions are needed:

1. $L(x^*) \geqq 0$.
2. All second order partial derivatives of L exist.
3. If $y^* = (y_1, y_2)$ and $|y^*| \leqq D < \infty$, then for some $K < \infty$

$$|L^{(i,j)}(x^* + y^*)| \leqq K(1 + |L^{(i,j)}(x^*)|) .$$

4. For $i = 1, 2$ and $j = 1, 2$, $E_0[L(X^*)] < \infty$, $E_0[|X|^4] < \infty$, $E_0[|X|^7|L^{(i)}(X^*)|] < \infty$, and $E_0[|X|^4|L^{(i,j)}(X^*)|] < \infty$.
5. (a) (w.l.o.g.) $E_0[L^{(1)}(X^*)] = E_0[L^{(2)}(X^*)] = 0$.
   (b) M is nonsingular.
   (c) (w.l.o.g.) $m_{13} = m_{23} = 0$.
   (d) $m_{133} \neq 0$ or $m_{233} \neq 0$ or both.
   (d′) (w.l.o.g.) $m_{133} = m_{233} = -1$.

Discussion of assumptions. Assumptions 2, 3 and 4 are technical assumptions and could probably be somewhat weakened. In their present form, however, they do encompass a broad range of losses and densities and are relatively straightforward to verify. Note that if L is a quadratic loss and if the eighth absolute moment of the density is finite, then Assumptions 1 through 4 are trivially satisfied.

The purpose of Assumption 5 is to ensure that X_3 is related to X^* in a suitable way. Some restriction is clearly necessary, for if X_3 and X^* are independent then $\delta_0(X) = X^*$ is admissible for estimating θ^*. (Independence reduces the problem to two dimensions, where Brown and Fox (1974) show that the best invariant estimator is admissible.)

Assumption 5 (a) is explicitly included only for ease in applications, since it is really an immediate consequence of the assumption that the problem has been parameterized so that $\delta_0(X) = X^*$ is the best invariant estimator. (If X^* is the best invariant estimator, then $E_0 L(X^* + (c_1, c_2))$ is minimized at $c_1 = c_2 = 0$.

Under Assumptions 1 through 4 it is easy to check by differentiating under the integral sign that 5(a) must then hold.)

Assumption 5(b) is quite weak, being satisfied in virtually all situations of practical interest. As an example, if L is the quadratic loss $L(x^*) = x^*Q(x^*)^t$ (where Q is a 2×2 positive definite matrix), then an easy calculation gives $M = 2Q\Sigma$ (where Σ is the covariance matrix of X^*). Providing X^* is nondegenerate, it is clear that M is nonsingular.

Assumption 5(c) can be satisfied without loss of generality. To see this, consider linear transformations of the problem defined by $Y = XP$, where P is a 3×3 nonsingular matrix with elements p_{ij} which satisfy $p_{31} = p_{32} = 0$. Note that the matrix P^* having elements p_{ij} $(i \leqq 2, j \leqq 2)$ is then nonsingular also, and that the transformed loss function is $L^*(y^*) = L(y^*(P^*)^{-1})$. The transformed problem is clearly equivalent to the original problem in terms of admissibility, amounting to nothing more than a change of variables in all risk expressions. Furthermore, Assumptions 1 through 4, 5(a) and 5(b) remain valid in the transformed problem. (To see that 5(b) is still satisfied, note that

$$\nabla L^*(Y^*) = (P^*)^{-1}\nabla L(X)\,, \tag{2.1}$$

and hence that

$$M^* = E_0[\nabla L^*(Y^*)Y^*] = (P^*)^{-1}MP^*\,,$$

which is clearly nonsingular.) To ensure that Assumption 5(c) is satisfied, it suffices to consider the transformed problem defined by

$$Y^* = X^*\,, \qquad Y_3 = X_3 - X^*M^{-1}(m_{13}, m_{23})^t\,.$$

An easy calculation shows that for the transformed problem

$$\begin{pmatrix} m_{13}^* \\ m_{23}^* \end{pmatrix} = E_0[\{\nabla L^*(Y^*)\}Y_3] = E_0\left[\{\nabla L(X^*)\}\left\{X_3 - X^*M^{-1}\begin{pmatrix} m_{13} \\ m_{23} \end{pmatrix}\right\}\right]$$
$$= \begin{pmatrix} m_{13} \\ m_{23} \end{pmatrix} - MM^{-1}\begin{pmatrix} m_{13} \\ m_{23} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and hence that 5(c) is satisfied.

Assumption 5(d) is needed to, in some sense, specify the exact problem which is to be dealt with. (In terms of the more general discussion in Berger (1976a), it is needed to specify the "moment structure." See also Brown (1974).) This assumption essentially states that Y_3 and Y^* are "related" through the moments m_{133} and m_{233}. This should be the most common situation in problems in which X^* and Y_3 are not independent. Assumption 5(d′) can be made, without loss of generality, providing that 5(d) holds. To see this, consider the transformation $Y_3 = X_3$ and $Y^* = X^*P^*$, where P^* is nonsingular. Calculation, using (2.1), gives

$$(m_{133}^*, m_{233}^*)^t = E_0[\nabla L^*(Y^*)Y_3^2] = (P^*)^{-1}E_0[\nabla L(X)X_3^2] = (P^*)^{-1}(m_{133}, m_{233})^t\,.$$

It is easy to check that providing either m_{133} or m_{233} is nonzero, $(P^*)^{-1}$ can be chosen so that $m^*_{133} = m^*_{233} = -1$. Hence Assumption 5(d′) would be satisfied in the transformed problem. (Again it is easy to check that the previous assumptions also still hold.) Section 5 gives examples of the above procedure of transforming a problem to satisfy Assumption 5.

3. Preliminary lemmas. In this section four rather technical lemmas will be given. The first three are fairly straightforward to verify and their proofs are omitted.

LEMMA 1.

(a) *If* $\alpha > 0$, *then* $2^{-\alpha/2}|x^*|^\alpha \leqq |x^*|_\alpha \leqq 2|x^*|^\alpha$.

(b) *If* α, β, c_1, c_2, d_1, *and* d_2 *are all positive constants, then there exist numbers* $0 < K_1 < \infty$ *and* $0 < K_2 < \infty$ *such that*

$$K_1|x^*|^{\alpha-\beta} \leqq \frac{|x_1|^\alpha/c_1 + |x_2|^\alpha/c_2}{|x_1|^\beta/d_1 + |x_2|^\beta/d_2} \leqq K_2|x^*|^{\alpha-\beta}.$$

(c) *If* $\alpha_1 > 0$, $\alpha_2 > 0$, *and* $\alpha = \alpha_1 + \alpha_2$, *then*

$$(1 + |x^*|^\alpha)^{-1} \leqq 3(1 + |x^*|^{\alpha_1})^{-1}(1 + |x^*|^{\alpha_2})^{-1}.$$

LEMMA 2. *If* $0 \leqq \alpha \leqq 3$ *and* $||x^*|| > 0$, *then there exists* $K < \infty$ *such that*

$$\{(1 + x_3^2)/(2||x^*||)\}^\alpha \exp\{-(1 + x_3^2)/(2||x^*||)\} \leqq K(1 + ||x^*||^{.51})(1 + |x_3|^{1.02})^{-1}.$$

LEMMA 3. (a) *For* $i = 1, 2, 3$, *assume that* $0 \leqq \alpha_i \leqq 1.02$, *and that* $t_i: R^3 \times R^3 \to [0, 1]$ *is a measurable function. Then there exists* $K < \infty$ *such that for* $j = 1, 2$ *and* $0 \leqq k \leqq 3$,

$$\begin{aligned}(3.1)\qquad E_\theta[|X - \theta|^k|L^{(j)}(X^* - \theta^*)| \textstyle\prod_{i=1}^3 (1 + |t_i(X, \theta)X_i + \{1 - t_i(X, \theta)\}\theta_i|^{\alpha_i})^{-1}] \\ \leqq K \textstyle\prod_{i=1}^3 (1 + |\theta_i|^{\alpha_i})^{-1}.\end{aligned}$$

(b) *There exists* $K < \infty$ *such that for* $j = 1, 2$ *and* $k = 1, 2$,

$$E_\theta[\{1 + |L^{(j,k)}(X^* - \theta^*)|\} \textstyle\prod_{i=1}^3 (1 + |X_i|^{1.02})^{-1}] \leqq K \prod_{i=1}^3 (1 + |\theta_i|^{1.02})^{-1}.$$

PROOF. The proof is a fairly straightforward Chebyshev type argument, carried out consecutively on the three coordinates. It is for this lemma that the large number of moments of Assumption 4 are needed. □

The following lemma plays a crucial role in the inadmissibility proof. Let $y = (y_1, y_2, y_3)$ and define

$$h_\rho(y) = \textstyle\prod_{i=1}^3 (1 + \rho|y_i|^{1.02})^{-1}.$$

LEMMA 4. *Assume* $g: R^3 \to [0, \infty)$ *is a measurable function satisfying* $\int_{R^3} g(y)\,dy = \infty$. *Let* $K_1 < \infty$ *be an arbitrary fixed constant. There exists* $0 < \rho < 1$ *such that for all* $\theta \in R^3$,

$$K_1 \int_{R^3} h_1(y)h_\rho(y - \theta)\,dy < \int_{R^3} g(y)h_\rho(y - \theta)\,dy.$$

PROOF. Clearly

$$(3.2)\qquad \int_{R^3} h_1(y)h_\rho(y - \theta)\,dy = \textstyle\prod_{i=1}^3 \int_{-\infty}^{\infty} (1 + |y_i|^{1.02})^{-1}(1 + \rho|y_i - \theta_i|^{1.02})^{-1}\,dy_i.$$

Break up the ith integral on the right-hand side of (3.2) into two integrals, I_1 and I_2, over $W_1 = \{y_i \in R^1 : |y_i - \theta_i| > |\theta_i|/2\}$ and W_1^c. If $y_i \in W_1$, it is clear that

$$(1 + \rho|y_i - \theta_i|^{1.02})^{-1} \leqq (1 + \rho[|\theta_i|/2]^{1.02})^{-1} .$$

Hence

$$(3.3) \qquad I_1 = \int_{W_1} (1 + |y_i|^{1.02})^{-1}(1 + \rho|y_i - \theta_i|^{1.02})^{-1} dy_i \leqq K(1 + \rho[|\theta_i|/2]^{1.02})^{-1} .$$

If $y_i \in W_1^c$, then $|y_i| \geqq |\theta_i|/2$. It follows that

$$(3.4) \qquad \begin{aligned} I_2 &= \int_{W_1^c} (1 + |y_i|^{1.02})^{-1}(1 + \rho|y_i - \theta_i|^{1.02})^{-1} dy_i \\ &\leqq K(1 + [|\theta_i|/2]^{1.02})^{-1} \int_{W_1^c} (1 + \rho|y_i - \theta_i|^{1.02})^{-1} dy_i . \end{aligned}$$

A simple change of variables and the condition $\rho < 1$ gives that

$$\int_{W_1^c} (1 + \rho|y_i - \theta_i|^{1.02})^{-1} dy_i \leqq K\rho^{-1/1.02} \leqq K\rho^{-1} .$$

Using this with (3.4) shows that if $\rho(|\theta_i|/2)^{1.02} \geqq 1$, then

$$(3.5) \qquad I_2 \leqq K\rho^{-1}(1 + [|\theta_i|/2]^{1.02})^{-1} \leqq K(\rho[|\theta_i|/2]^{1.02})^{-1} \leqq 2K(1 + \rho[|\theta_i|/2]^{1.02})^{-1} .$$

Note next that

$$\int_{W_1^c} (1 + \rho|y_i - \theta_i|^{1.02})^{-1} dy \leqq \int_{W_1^c} dy_i = |\theta_i| .$$

Using this with (3.4) shows that if $\rho(|\theta_i|/2)^{1.02} < 1$, then

$$(3.6) \qquad \begin{aligned} I_2 &\leqq K|\theta_i|(1 + [|\theta_i|/2]^{1.02})^{-1} \leqq 4K(1 + [|\theta_i|/2]^{(1.02-1)})^{-1} \\ &\leqq 4K(1 + \rho^{1/1.02}[|\theta_i|/2]^{1.02})^{-1} \leqq 4K(1 + \rho[|\theta_i|/2]^{1.02})^{-1} . \end{aligned}$$

Combining (3.2), (3.3), (3.5) and (3.6) gives that for some $K_2 < \infty$,

$$(3.7) \qquad \int_{R^3} h_1(y)h_\rho(y - \theta)\, dy \leqq K_2 \prod_{i=1}^3 (1 + \rho[|\theta_i|/2]^{1.02})^{-1} .$$

Define $V_n = \{y \in R^3 : |y_i| \leqq n, i = 1, 2, 3\}$. Choose n large enough so that

$$(3.8) \qquad \int_{V_n} g(y)\, dy > K_1 K_2 4^{3.06} .$$

Finally, choose $\rho = (2n)^{-1.02}$. Now if $|\theta_i| \leqq n$ and $|y_i| \leqq n$, then

$$(3.9) \qquad (1 + \rho|y_i - \theta_i|^{1.02})^{-1} \geqq (1 + \rho[2n]^{1.02})^{-1} = 2^{-1} .$$

If $|\theta_i| > n$ and $|y_i| \leqq n$, then

$$(3.10) \qquad \begin{aligned} (1 + \rho|y_i - \theta_i|^{1.02})^{-1} &\geqq (1 + \rho[|\theta_i| + n]^{1.02})^{-1} \geqq (1 + \rho[2|\theta_i|]^{1.02})^{-1} \\ &\geqq 4^{-1.02}(1 + \rho[|\theta_i|/2]^{1.02})^{-1} . \end{aligned}$$

Combining (3.9) and (3.10) gives that if $\theta \in R^3$ and $y \in V_n$, then

$$(3.11) \qquad h_\rho(y - \theta) \geqq 4^{-3.06} \prod_{i=1}^3 (1 + \rho[|\theta_i|/2]^{1.02})^{-1} .$$

Using (3.7), (3.8), (3.11) and the fact that $g \geqq 0$ finally gives

$$\begin{aligned} \int_{R^3} g(y)h_\rho(y - \theta)\, dy &\geqq \int_{V_n} g(y)h_\rho(y - \theta)\, dy \\ &\geqq 4^{-3.06} \int_{V_n} g(y)\, dy \prod_{i=1}^3 (1 + \rho[|\theta_i|/2]^{1.02})^{-1} \\ &> K_1 K_2 \prod_{i=1}^3 (1 + \rho[|\theta_i|/2]^{1.02})^{-1} \\ &\geqq K_1 \int_{R^3} h_1(y)h_\rho(y - \theta)\, dy . \end{aligned}$$

□

4. Inadmissibility of the best invariant estimator. Let $\Omega = \{z \in R^3 : ||z^*|| > 0\}$ and let $I_\Omega(z)$ denote the usual indicator function on Ω. Define $\gamma : R^3 \to R^2$ by

$$\gamma(x) = \frac{-(x^* M^{-1})}{(1 + |x^*|_{2.48})} \exp\left\{-\frac{1 + x_3^2}{2||x^*||}\right\} I_\Omega(x) . \tag{4.1}$$

Let γ_i ($i = 1, 2$) be the coordinate functions of γ, and define $\delta(x) = x^* + \gamma(x)$. Finally, let $Y = (Y_1, Y_2, Y_3)$ be a random vector with density $K_\rho h_\rho(y)$ (K_ρ being the normalizing constant), and define the randomized estimator $\delta^\rho : R^3 \times R^3 \to R^2$ by

$$\delta^\rho(X, Y) = \delta(X + Y) - Y^* .$$

THEOREM 1. *If Assumptions* 1 *through* 5 *of Section* 2 *hold, then the best invariant estimator,* $\delta_0(X) = X^*$, *is inadmissible for estimating* θ^*. *Indeed there exists* $0 < \rho < 1$ *such that* $\Delta_{\delta^\rho}(\theta) < 0$ *for all* $\theta \in R^3$.

(Note that if L is convex, then Jensen's inequality shows that the nonrandomized estimator $\delta^*(X) = E_\rho^Y(\delta^\rho(X, Y))$ is also better than δ_0.)

PROOF. Clearly

$$\begin{aligned} \Delta_{\delta^\rho}(\theta) &= E_\theta^X E_\rho^Y\{L(\delta^\rho(X, Y) - \theta^*) - L(X^* - \theta^*)\} \\ &= E_\theta^X E_\rho^Y\{L(\delta(X + Y) - [\theta^* + Y^*]) \\ &\quad - L([X^* + Y^*] - [\theta^* + Y^*])\} \\ &= E_\rho^Y E_{(\theta+Y)}^X\{L(\delta(X) - [\theta^* + Y^*]) - L(\delta_0(X) - [\theta^* + Y^*])\} \\ &= E_\rho^Y \Delta_\delta(\theta + Y) . \end{aligned} \tag{4.2}$$

Define

$$g(\theta) = (.02)(1 + |\theta^*|_{2.48})^{-1} \exp\{-(1 + \theta_3^2)/(2||\theta^*||)\} I_\Omega(\theta) . \tag{4.3}$$

It will be shown that there exists $K_1 < \infty$ such that

$$\Delta_\delta(\theta) \leqq -g(\theta) + r(\theta) , \qquad \text{where} \quad |r(\theta)| \leqq K_1 h_1(\theta) . \tag{4.4}$$

Assume for the moment that this is true. From the definition of $||\theta^*||$ and the fact that M^{-1} is nonsingular, it is clear that Ω is a half plane in R^2. It is thus easy to check (integrating out first over θ_3) that $\int_{R^3} g(\theta)\, d\theta = \infty$. It can be concluded from Lemma 4 that there exists $0 < \rho < 1$ such that

$$E_\rho^Y[K_1 h_1(\theta + Y)] < E_\rho^Y[g(\theta + Y)] .$$

Using this, together with (4.2), (4.3) and (4.4), shows that

$$\begin{aligned} \Delta_{\delta^\rho}(\theta) &= E_\rho^Y \Delta_\delta(\theta + Y) \\ &\leqq E_\rho^Y[-g(\theta + Y) + r(\theta + Y)] \\ &\leqq -E_\rho^Y[g(\theta + Y)] + E_\rho^Y[K_1 h_1(\theta + Y)] < 0 . \end{aligned}$$

The proof will thus be complete as soon as (4.4) is verified. The above argument is basically a "randomization-of-the-origin" argument, first used by Brown. (See Brown (1974) and Berger (1976a) for other examples.) In previous proofs using this type of argument, a region such as Ω was found in which $\Delta_\delta(\theta) < 0$. The estimator was then randomized to "spread out" the improvement in risk to include all of R^3. The arguments involved in proving that $\Delta_\delta(\theta) < 0$ for all $\theta \in \Omega$ are often quite formidable technically. In this paper it is only shown that $\Delta_\delta(\theta) = -g(\theta) + r(\theta)$, where $|r(\theta)| \leqq Kh_1(\theta)$. No attempt is made to show that $-g(\theta) + r(\theta) < 0$ for all $\theta \in \Omega$. The error term, r, is instead handled by a slightly more involved randomization-of-the-origin argument. This different approach results in a significantly easier proof. For example, the present proof is considerably shorter than the analogous proof in Berger (1976a). (Of course, generalized Bayes estimators were also considered in that paper, which precluded the use of a randomization-of-the-origin argument.)

To prove (4.4) first L and then γ will be expanded in Taylor expansions. Since several such arguments have previously been used (Brown (1975), Brown (1974), Berger (1976a), Berger (1976b), and Berger (1976c)), an attempt is made to suppress details not unique to this particular problem.

By definition,

$$\Delta_\delta(\theta) = \int_{R^3} [L(x^* + \gamma(x) - \theta^*) - L(x^* - \theta^*)] f(x - \theta)\, dx\,. \tag{4.5}$$

Expanding $L(x^* + \gamma(x) - \theta^*)$ in a Taylor expansion about $(x^* - \theta^*)$ gives

$$\begin{aligned} L(x^* + \gamma(x) - \theta^*) = L(x^* - \theta^*) &+ \textstyle\sum_{i=1}^2 \gamma_i(x) L^{(i)}(x^* - \theta^*) \\ &+ \tfrac{1}{2} \textstyle\sum_{i=1}^2 \sum_{j=1}^2 \gamma_i(x)\gamma_j(x) L^{(i,j)}(x^* - \theta^* + t(x, \theta^*)\gamma(x))\,, \end{aligned}$$

where $0 \leqq t(x, \theta^*) \leqq 1$. Using this expansion in (4.5) results in

$$\Delta_\delta(\theta) = \Delta^1(\theta) + \Delta^2(\theta) + \Delta^3(\theta)\,, \tag{4.6}$$

where

$$\Delta^i(\theta) = \int_{R^3} \gamma_i(x) L^{(i)}(x^* - \theta^*) f(x - \theta)\, dx\,, \quad i = 1, 2\,,$$

and

$$\Delta^3(\theta) = \tfrac{1}{2} \textstyle\sum_{i=1}^2 \sum_{j=1}^2 \int_{R^3} \gamma_i(x)\gamma_j(x) L^{(i,j)}(x^* - \theta^* + t\gamma) f(x - \theta)\, dx\,.$$

Consider $\Delta^3(\theta)$ first. From the definitions of t and γ, it is clear that $|t\gamma|$ is bounded. Assumption 3 thus gives that

$$|\Delta^3(\theta)| \leqq K \textstyle\sum_{i=1}^2 \sum_{j=1}^2 \int_{R^3} |\gamma(x)|^2 (1 + |L^{(i,j)}(x^* - \theta^*)|) f(x - \theta)\, dx\,. \tag{4.7}$$

From (4.1) and Lemmas 1 and 2 it is easy to see that

$$\begin{aligned} |\gamma(x)|^2 &\leqq K(1 + |x^*|^{2.96})^{-1} \exp\{-x_3^2/(2\|x^*\|)\} \\ &\leqq K(1 + |x^*|^{2.96})^{-1}(1 + \|x^*\|^{.51})(1 + |x_3|^{1.02})^{-1} \\ &\leqq K(1 + |x_1|^{1.02})^{-1}(1 + |x_2|^{1.02})^{-1}(1 + |x_3|^{1.02})^{-1} = Kh_1(x)\,. \end{aligned}$$

Using this bound in (4.7) and applying Lemma 3(b) gives

$$|\Delta^3(\theta)| \leqq Kh_1(\theta)\,. \tag{4.8}$$

Consider $\Delta^1(\theta)$ next. From (4.1) it can be seen that γ_1 has continuous second order partial derivatives. Hence $\gamma_1(x)$ can be expanded in a Taylor series about θ to give

$$(4.9)\qquad \begin{aligned}\gamma_1(x) = \gamma_1(\theta) &+ \textstyle\sum_{i=1}^3 \gamma_1^{(i)}(\theta)(x_i - \theta_i) \\ &+ \tfrac{1}{2} \textstyle\sum_{i=1}^3 \sum_{j=1}^3 \gamma_1^{(i,j)}(\theta)(x_i - \theta_i)(x_j - \theta_j) \\ &+ \textstyle\int_0^1 (1 - t)[\sum_{i=1}^3 \sum_{j=1}^3 \{\gamma_1^{(i,j)}(\theta + t(x - \theta)) - \gamma_1^{(i,j)}(\theta)\} \\ &\times (x_i - \theta_i)(x_j - \theta_j)]\, dt\,.\end{aligned}$$

Using this expansion gives

$$(4.10)\qquad \begin{aligned}\Delta^1(\theta) &= \textstyle\int_{R^3} \gamma_1(x) L^{(1)}(x^* - \theta^*) f(x - \theta)\, dx \\ &= I_1(\theta) + I_2(\theta) + I_3(\theta) + I_4(\theta)\,,\end{aligned}$$

where $I_i(\theta)$ is the integral resulting from the ith expression on the right-hand side of (4.9). By Assumption 5(a)

$$(4.11)\qquad I_1(\theta) = \textstyle\int_{R^3} \gamma_1(\theta) L^{(1)}(x^* - \theta^*) f(x - \theta)\, dx = 0\,.$$

By the definitions of the m_{ij} and m_{ijk},

$$I_2(\theta) + I_3(\theta) = \textstyle\sum_{i=1}^3 m_{1i} \gamma_1^{(i)}(\theta) + \tfrac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 m_{1ij} \gamma_1^{(i,j)}(\theta)\,.$$

Using Lemmas 1 and 2, a simple calculation shows that $|\gamma_1^{(i,j)}(\theta)| \leqq Kh_1(\theta)$ unless $i = j = 3$. Together with Assumptions 4, 5(c) and 5(d′), this shows that

$$(4.12)\qquad I_2(\theta) + I_3(\theta) = m_{11}\gamma_1^{(1)}(\theta) + m_{12}\gamma_1^{(2)}(\theta) - \tfrac{1}{2}\gamma_1^{(3,3)}(\theta) + r_1(\theta)\,,$$

where $|r_1(\theta)| \leqq Kh_1(\theta)$.

It remains to analyze $I_4(\theta)$. Consider first the terms corresponding to $i \neq 3$ or $j \neq 3$. As earlier, Lemmas 1 and 2 will show that $|\gamma_1^{(i,j)}(\theta + t(x - \theta))| \leqq Kh_1(\theta + t(x - \theta))$. Using this, together with Lemma 3(a), and interchanging orders of integration gives

$$(4.13)\qquad \begin{aligned}&\textstyle\int_{R^3} \int_0^1 (1 - t)[|\gamma_1^{(i,j)}(\theta + t(x - \theta))| + |\gamma_1^{(i,j)}(\theta)|] \\ &\qquad \times |x - \theta|^2\, dt\, |L^{(1)}(x^* - \theta^*)| f(x - \theta)\, dx \\ &\qquad\qquad \leqq 2 \textstyle\int_0^1 (1 - t)\, dt\, Kh_1(\theta) = Kh_1(\theta)\,.\end{aligned}$$

To handle the $i = j = 3$ term of $I_4(\theta)$, note that $\gamma_1^{(3,3)}$ has continuous partial derivatives and hence

$$\begin{aligned}&\{\gamma_1^{(3,3)}(\theta + t(x - \theta)) - \gamma_1^{(3,3)}(\theta)\} \\ &\qquad = \textstyle\int_0^1 \sum_{i=1}^3 \{\gamma_1^{(3,3,i)}(\theta + t_2 t(x - \theta)) t(x_i - \theta_i)\}\, dt_2\,.\end{aligned}$$

Lemmas 1, 2 and 3(a) can again be used to give

$$\begin{aligned}&\textstyle\int_{R^3} \int_0^1 \int_0^1 |\gamma_1^{(3,3,i)}(\theta + t_2 t(x - \theta))|(1 - t)t|x - \theta|^3 |L^{(1)}(x^* - \theta^*)| f(x - \theta)\, dt_2\, dt\, dx \\ &\qquad \leqq Kh_1(\theta)\,.\end{aligned}$$

Together with (4.13) this shows that

$$(4.14)\qquad |I_4(\theta)| \leqq Kh_1(\theta)\,.$$

Combining (4.10), (4.11), (4.12) and (4.14) finally gives

$$\Delta^1(\theta) = m_{11}\gamma_1^{(1)}(\theta) + m_{12}\gamma_1^{(2)}(\theta) - \tfrac{1}{2}\gamma_1^{(3,3)}(\theta) + r_2(\theta) \quad \text{where} \quad |r_2(\theta)| \leqq Kh_1(\theta) . \tag{4.15}$$

It can similarly be shown that

$$\Delta^2(\theta) = m_{22}\gamma_2^{(2)}(\theta) + m_{21}\gamma_2^{(1)}(\theta) - \tfrac{1}{2}\gamma_2^{(3,3)}(\theta) + r_3(\theta) , \quad \text{where} \quad |r_3(\theta)| \leqq Kh_1(\theta) . \tag{4.16}$$

A tedious but straightforward calculation shows that

$$m_{11}\gamma_1^{(1)}(\theta) + m_{12}\gamma_1^{(2)}(\theta) + m_{22}\gamma_2^{(2)}(\theta) + m_{21}\gamma_2^{(1)}(\theta) - \tfrac{1}{2}\gamma_1^{(3,3)}(\theta) - \tfrac{1}{2}\gamma_2^{(3,3)}(\theta) \\ \leqq (-.02)(1 + |\theta^*|_{2.48})^{-1} \exp\{-(1 + \theta_3^2)/(2\|\theta^*\|)\} I_\Omega(\theta) .$$

Using this together with (4.6), (4.8), (4.15) and (4.16) finally verifies (4.4) and completes the proof. □

5. Examples. As mentioned in the introduction, the results of this paper will not apply to estimating a multivariate normal mean. This is because a linear transformation of the problem can be made, which will result in the coordinates being independent. Hence when Assumptions 5(a) and 5(c) are satisfied, 5(d) will be violated. Indeed it is clear that the best invariant estimator is admissible for this problem. (The transformed problem reduces, by independence, to estimating a two-dimensional normal mean.) Two examples in which the theory can be applied follow.

EXAMPLE 1. Suppose $Y_i = \ln(X_i - \theta_i)$, $i = 1, 2, 3$, and that $Y = (Y_1, Y_2, Y_3)$ has a multivariate normal density with mean 0 and known positive definite covariance matrix Σ (with elements σ_{ij}). The random vector $X = (X_1, X_2, X_3)$ has a density $f(x - \theta)$ which is a version of a multivariate lognormal density. Let $L(x^*) = x_1^2 + x_2^2$. Assumptions 1 through 4 of Section 2 are clearly satisfied. Note also that

$$E^X_{\theta=0}[X_1^{t_1} X_2^{t_2} X_3^{t_3}] = E^Y[\exp\{t_1 y_1 + t_2 y_2 + t_3 y_3\}] \\ = \exp\{\tfrac{1}{2}(t_1, t_2, t_3)\Sigma(t_1, t_2, t_3)'\} . \tag{5.1}$$

Assumption 5 must now be considered. Note that

$$E_0[L^{(i)}(X^*)] = 2E_0[X_i] = 2\exp\{\sigma_{ii}/2\} \neq 0 ,$$

and hence that Assumption 5(a) is violated. To reparameterize, define $\eta^* = \theta^* + (\exp\{\sigma_{11}/2\}, \exp\{\sigma_{22}/2\})$, $n_3 = \theta_3$, and consider the equivalent problem of estimating η^* based on observing X. Clearly for $i = 1, 2$,

$$E_{\eta=0}[L^{(i)}(X^*)] = 2E_{\theta=0}[X_i - \exp\{\sigma_{ii}/2\}] = 0 ,$$

and Assumption 5(a) is satisfied. Assumption 5(b) is clearly true since L is quadratic and X^* has a nondegenerate distribution. (See the discussion of Assumption 5(b).) Assumption 5(c) can be satisfied by making the transformation

$$Z_3 = X_3 - X^*M^{-1}(m_{13}, m_{23})^t , \qquad Z^* = X^* ,$$

as suggested in Section 2. (Formula (5.1) can be used to calculate the m_{ij} and hence M^{-1}.) It remains only to verify Assumption 5(d), and hence to calculate $m_{i33}^* = 2E_{\eta=0}[Z_3^2 X_i]$, $i = 1, 2$.

For the special case $\sigma_{12} = 0$, $\sigma_{11} = \sigma_{22} = \sigma_{33} = \alpha \neq 0$, and $\sigma_{13} = \sigma_{23} = \beta$, the calculation was carried out using (5.1). The result was (for $i = 1,2$)

$$m_{i33}^* = 4e^{1.5\alpha}(e^\beta - 1)(e^\alpha - 1)^{-1}[e^{(2\alpha+\beta)} - e^{(\alpha+2\beta)} - e^\alpha + 3e^\beta - e^{2\beta} - 1]. \tag{5.2}$$

Providing this quantity is nonzero, Assumption 5(d) will be satisfied.

(For example, if $\alpha = 1$ and $\beta = .5$, then $m_{i33}^* \cong 22.4 \neq 0$.) It can then be concluded that the best invariant estimator $\delta_0(X) = X^*$ is inadmissible for estimating η^*, and hence that $[X^* - (\exp\{\alpha/2\}, \exp\{\alpha/2\})]$ is inadmissible for estimating θ^*.

EXAMPLE 2. Assume that $X = (X_1, X_2, X_3)$ has a multivariate normal distribution with mean 0 and a known correlation matrix with elements ρ_{ij} which satisfy $\rho_{ii} = 1$ ($i = 1, 2, 3$), and $\rho_{i3} \neq 0$ ($i = 1, 2$). The standard deviations, σ_i, are assumed to be unknown, and it is desired to estimate (σ_1, σ_2). Let the loss incurred in estimating (σ_1, σ_2) by (d_1, d_2) be of the form $L(\ln [d_1/\sigma_1], \ln [d_2/\sigma_2])$.

Usual invariant estimators of (σ_1, σ_2) are functions of $(|X_1|, |X_2|)$. Such estimators can often be shown to be inadmissible by transforming in the usual way to a location vector problem. Thus let $Y_i = \ln |X_i|$ and $\theta_i = \ln \sigma_i$. The problem can now be considered to be one of estimating θ^* by an estimator $\delta(Y)$, and under loss $L(\delta(Y) - \theta^*)$.

As a specific example, suppose $L(y^*) = y_1^2 + y_2^2$, $\rho_{12} = 0$, and $\rho_{13} = \rho_{23} = .8$. Assumptions 1 through 4 are easy to check. Validating Assumption 5(a) again requires a reparametrization of the problem. Numerical calculation showed that $\eta^* = \theta^* - (.63, .63)$ is the correct reparameterization, to the nearest hundredth. Assumption 5(b) is obviously satisfied. To satisfy 5(c), it was numerically calculated that the linear transformation $Z_3 = Y_3 - .35Y_1 - .35Y_2$ must be made. A final calculation showed that 5(d) was then satisfied and hence that Theorem 1 could be applied. The conclusion is that the best invariant estimator $[Y^* + (.63, .63)]$ is inadmissible for estimating θ^*, and hence that any invariant estimator based on $(|X_1|, |X_2|)$ is inadmissible for estimating (σ_1, σ_2) (for the given loss and correlation matrix, of course). As in Berger (1976a), this example could be extended to a more general multiplicative model (i.e., a model where the random errors are multiplicative).

Acknowledgments. I would like to express my thanks to the referees for their helpful comments and suggestions.

REFERENCES

[1] BERGER, J. (1976a). Inadmissibility results for generalized Bayes estimators of coordinates of a location vector. *Ann. Statist.* **4** 302–333.

[2] BERGER, J. (1976b). Admissibility results for generalized Bayes estimators of coordinates of a location vector. *Ann. Statist.* **4** 334–356.

[3] BERGER, J. (1976c). Tail minimaxity in location vector problems and its applications. *Ann. Statist.* **4** 33-50.

[4] BROWN, L. D. (1966). On the admissibility of invariant estimators of one or more location parameters. *Ann. Math. Statist.* **37** 1087-1136.

[5] BROWN, L. D. (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems. *Ann. Math. Statist.* **42** 855-903.

[6] BROWN, L. D. (1974). An heuristic method for determining admissibility of estimators—with applications. Technical Report, Rutgers Univ.

[7] BROWN, L. D. (1975). Estimation with incompletely specified loss functions. *J. Amer. Statist. Assoc.* **70** 417-427.

[8] BROWN, L. D. and FOX, M. (1974). Admissibility of procedures in two-dimensional location parameter problems. *Ann. Statist.* **2** 248-266.

[9] FARRELL, R. (1964). Estimators of a location parameter in the absolutely continuous case. *Ann. Math. Statist.* **35** 949-998.

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

[11] PORTNOY, S. (1973). On solutions to a non-linear diffusion equation arising in an admissibility problem. Technical Report, Harvard Univ.

[12] PORTNOY, S. (1975). Admissibility of the best invariant estimator of one coordinate of a location vector. *Ann. Statist.* **3** 448-450.

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