Ann. Inst. Statist. Math.
Vol. 45, No. 3, 551–565 (1993)

SHRINKAGE ESTIMATORS OF THE LOCATION PARAMETER FOR CERTAIN SPHERICALLY SYMMETRIC DISTRIBUTIONS

ANN COHEN BRANDWEIN[1], STEFAN RALESCU[2] AND WILLIAM E. STRAWDERMAN[3]*

[1] *Department of Statistics, Baruch College of the City University of New York, Box 513, 17 Lexington Av., New York, NY 10010, U.S.A.*
[2] *Department of Mathematics, Queens College of the City University of New York, 65-30 Kissena Boulevard, Flushing, NY 11367, U.S.A.*
[3] *Department of Statistics, Hill Center, Busch Campus, Rutgers University, New Brunswick, NJ 08903, U.S.A.*

(Received September 4, 1991; revised September 16, 1992)

Abstract. We consider estimation of a location vector for particular subclasses of spherically symmetric distributions in the presence of a known or unknown scale parameter. Specifically, for these spherically symmetric distributions we obtain slightly more general conditions and larger classes of estimators than Brandwein and Strawderman (1991, *Ann. Statist.*, **19**, 1639–1650) under which estimators of the form $X + ag(X)$ dominate X for quadratic loss, concave functions of quadratic loss and general quadratic loss.

Key words and phrases: Spherical symmetry, quadratic loss, concave loss, location parameter, unknown scale.

1. Introduction

We consider estimation of a p-dimensional location parameter of a spherically symmetric (s.s.) distribution. Specifically, let $X \sim f(\|x - \theta\|^2)$ and consider estimation of θ with loss function $L(\theta, \delta) = \|\delta - \theta\|^2$. Stein (1981) derived an expression for the risk of estimators of the form $X + ag(X)$ using integration by parts when X has a normal distribution. He also gives conditions under which $X + ag(X)$ dominates X, the best equivariant estimator. Chou and Strawderman (1990) derive a similar result for a mixture of normal distributions. Brandwein and Strawderman (1991) using the divergence theorem and (essentially) superharmonicity of $\|g\|^2$ give conditions on a and $g(\cdot)$ for dominance when $f(\cdot)$ is an arbitrary spherically symmetric distribution.

Ralescu *et al.* (1992) using integration by parts in case of a general spherically symmetric distribution gave necessary and sufficient conditions on the constant a for $X + ag(X)$ to improve on X for the case where $\text{div}g(X) \le 0$ for all X.

* Research supported by NSF grant DMS-88-22622.

Berger (1975), Bock (1985), Akai (1986), Cellier *et al.* (1989), and others have used the integration by parts technique, first introduced by Stein (1973), in the spherically symmetric case.

In this paper, we use the techniques in Ralescu *et al.* (1992) to show that the bounds on the constant a given in Brandwein and Strawderman (1991) may be improved in certain cases.

In particular, if $q(t) = \int_t^\infty f(u)du/f(t)$ is either monotone increasing or monotone decreasing, the bound on a may be increased over that in Brandwein and Strawderman (1991). The result reduces to results of Bock (1985) if $X + ag(X)$ is taken to be a James-Stein type estimator.

Additionally, if $f(\cdot)$ is a spherically symmetric unimodal (s.s.u.) distribution (i.e., $f(t)$ is nonincreasing in t), then the upper bound on the constant a may also be enlarged. Here, our result for the James-Stein estimator reduces to that in Brandwein and Strawderman (1978).

Another subclass of spherically symmetric distributions for which alternate bounds on the constant a are found is the class of distributions controlled by a s.s.u. distribution. Specifically, we say $f(\cdot)$ is controlled by $(f_1(\cdot), m)$ if $f_1(t) \le f(t) \le m f_1(t)$ for all $t > 0$ where $f_1(t)$ is nonincreasing in t.

Sections 2 and 3 are devoted to the above developments.

In Section 4, we obtain analogous results for loss functions which are nondecreasing concave functions of $\|\delta - \theta\|^2$ and for general quadratic loss, $L(\delta, \theta) = (\delta - \theta)' D(\delta - \theta)$ where D is positive definite. In Section 5, results for estimating θ in the presence of an unknown scale parameter are given.

The following gives the flavor of results in the paper: Let $X \sim f(\|x - \theta\|^2)$, $p \ge 3$ and the loss be $L(\delta, \theta) = \|\delta - \theta\|^2$. Let

$$\delta_{a,g}(X) = X + ag(X) \tag{1.1}$$

be an estimator of θ such that

(a) $\text{div} g(X) \le h(X)$ where $h(\cdot)$ is such that $E_\theta h(W)$ is nondecreasing in R and $E_\theta[R^2 h(W)]$ is nonincreasing in R if $W \sim \text{Uniform}\{\|W - \theta\|^2 = R^2\}$.

(b) $\|g\|^2 + 2h \le 0$.

Then, provided $E_0\|X\|^2$ and $E_0(1/\|X\|^2) < \infty$, $\delta_{a,g}(X)$ beats X if any of the following hold.

A. $0 < a \le 1/pE_0(1/\|X\|^2)$ for any $f(\cdot)$.

B. $q(t) = \int_t^\infty f(u)du/f(t)$ is nonincreasing and $0 < a \le E_0(\|X\|^2)/p$.

C. $q(t)$ is nondecreasing and $0 < a \le \frac{1}{(p-2)E_0(1/\|X\|^2)}$.

D. $f(\cdot)$ is s.s.u. and $0 < a \le \frac{p}{(p^2-4)E_0(1/\|X\|^2)}$.

E. The pair $(f_1(t), m)$ controls $f(\|x - \theta\|^2)$ and $0 < a \le \frac{p}{(p^2-4)m^2 E_0(1/\|X\|^2)}$.

2. Generalized stein estimators for certain spherically symmetric distributions

Consider throughout this paper a $p \times 1$, $p \ge 3$, random vector $X = [X_1, X_2, \ldots, X_p]'$ having a density $f_\theta(x) = f(\|x - \theta\|^2)$, $\theta = (\theta_1, \theta_2, \ldots, \theta_p)'$, which is continuous a.e. with respect to Lebesgue measure. We will denote this by $X \sim \text{s.s.d.}(\theta)$ from this point on. With respect to quadratic loss, $\|\delta - \theta\|^2$, we investigate conditions on a and g under which estimators of the form $\delta_{a,g}(X)$ (1.1) will dominate X (see Stein (1981)).

Brandwein and Strawderman (1991) found conditions under which $\delta_{a,g}(X)$ dominates X for any spherically symmetric distribution. In this section, we look at particular *subclasses of spherically symmetric distributions* and under slightly more general conditions show that $\delta_a(X)$ improves on X and is thus minimax for a larger range of values of a.

Using integration by parts, Ralescu *et al.* (1992), in Lemma 2.1, establish necessary and sufficient conditions for the generalized Stein estimator to dominate X. We will use this lemma to establish our minimax results. We restate this lemma for completeness. We assume throughout this section that $E_0(1/\|X\|^2)$ and $E_0(\|X\|^2)$ exist and are finite.

LEMMA 2.1. *Let $X \sim s.s.d.(\theta)$, then for $p \geq 3$, provided the divergence of g is nonpositive, the risk of $\delta_{a,g}(X)$ dominates (is less than or equal to, for all θ) the risk of X with respect to quadratic loss if and only if $0 < a \leq \inf_\theta \xi_f(\theta)$ where*

$$\xi_f(\theta) = \frac{\int -(\mathrm{div} g(x))(\int_{\|x-\theta\|^2}^{\infty} f(t)dt)dx}{\int \|g(x)\|^2 f(\|x-\theta\|^2)dx}. \tag{2.1}$$

Let

$$q(t) = \int_t^\infty f(u)du/f(t) \quad \text{for} \quad f(t) > 0. \tag{2.2}$$

In Theorem 2.1, *we will* first consider s.s. distributions for which $q(t)$ is nonincreasing in t *and in Theorem* 2.2, we consider distributions for which $q(t)$ is nondecreasing. *In both cases we will get a larger bound on a than the Brandwein-Strawderman general s.s. bound, $1/pE_0(1/\|X\|^2)$. These improvements are similar in spirit to the results of Bock* (1985) *for the James-Stein estimator.*

THEOREM 2.1. *If $X \sim s.s.d.(\theta)$, then with respect to quadratic loss, $\delta_{a,g}(X)$ dominates X provided:*

(i) *$q(t)$ is nonincreasing in t,*

(ii) *$-\mathrm{div} g \geq -h$ where h is such that $E_\theta[h(W)]$ is nondecreasing in R, when $W \sim \mathcal{U}\{\|W-\theta\|^2 = R^2\}$,*

(iii) *$\|g\|^2 + 2h \leq 0$, and*

(iv) *$0 < a \leq E_{\theta=0}(\|X\|^2)/p$.*

PROOF. If we condition on $\|X-\theta\| = R$, then the numerator of $\xi_f(\theta)$, defined by (2.1), equals $E_\theta[-\mathrm{div} g(X) q(\|X-\theta\|^2)]$.

$$\begin{aligned} E_\theta[-\mathrm{div} g(X) q(\|X-\theta\|^2)] &= E_{f_R}[q(R^2)E_\theta[-\mathrm{div} g(X) \mid \|X-\theta\| = R]] \\ &\geq E_{f_R}[q(R^2)E_\theta[-h(X) \mid \|X-\theta\| = R]], \end{aligned} \tag{2.3}$$

by condition (ii).

Condition (i) implies

$$E_{f_R}[q(R^2)E_\theta[-h(X) \mid \|X-\theta\| = R]] \geq E_{f_R}[q(R^2)]E_\theta[-h(X)]. \tag{2.4}$$

Now,

$$E_{f_R}[q(R^2)] = \int q(r^2) f_R(r)dr = \int cq(r^2)r^{p-1}f(r^2)dr \tag{2.5}$$

where $f_R(r)$ is the density of R and $c = 1/\int_0^\infty r^{p-1}f(r^2)dr$.

Using integration by parts, it is straightforward to show $E_{f_R}[q(R^2)] = (2/p)\cdot E_0[\|X\|^2]$. Thus, (2.3) and (2.4) and condition (iii) imply

$$\xi_f(\theta) \geq \frac{(2/p)E_0[\|X\|^2]E_\theta[-h(X)]}{E_\theta[\|g(X)\|^2]} \geq E_0[\|X\|^2]/p. \qquad \square$$

Remark 2.1. It is well known that if k is superharmonic ($\nabla^2 k(x) = \sum(\partial/\partial x_i^2)k(x) \leq 0$) and $W \sim \mathcal{U}\{\|W-\theta\|^2 = R^2\}$, then $E_\theta\{k(W)\}$ is nonincreasing in R. Thus, if $k = -h$ is superharmonic, then condition (ii) of Theorem 2.1 is satisfied.

THEOREM 2.2. *Suppose $X \sim s.s.d.(\theta)$ and assume that*

(i) *$q(t)$ is nondecreasing in t,*

(ii) *$-\mathrm{div} g \geq -h$ where h is such that $E_\theta[R^2 h(W)]$ is nonincreasing in R, when $W \sim \mathcal{U}\{\|W-\theta\|^2 = R^2\}$,*

(iii) *$\|g\|^2 + 2h \leq 0$, and*

(iv) *$0 < a \leq 1/(p-2)E_0(1/\|X\|^2)$.*

Then $\delta_{a,g}(X)$ has (uniformly) smaller risk than X.

PROOF. The numerator of $\xi_f(\theta)$, defined by (2.1), equals

$$E_\theta[-\mathrm{div} g(X) q(\|X-\theta\|^2)] = E_{f_R}[q(R^2)E_\theta[-\mathrm{div} g(X) \mid \|X-\theta\| = R]]. \tag{2.6}$$

If $G(R) = E_\theta[-\mathrm{div} g(X) \mid \|X-\theta\| = R]$ and $f_R(r)$ is the density of R as defined by (2.5), then (2.6) becomes

$$E_\theta[-\mathrm{div} g(X) q(\|X-\theta\|^2)] = \int_0^\infty (cr^{p-3}f(r^2))[r^2 G(r)]q(r^2)dr. \tag{2.7}$$

Let $\rho_R(r) = cr^{p-3}f(r^2)/E_0(\|X\|^{-2})$ be another density of R and $G^*(R) = R^2 G(R)$. Then (2.7) is equivalent to $E_0(\|X\|^{-2})E_{\rho_R}[q(R^2)G^*(R)]$. Thus, (2.6), (2.7), and assumption (ii) imply

$$\text{numerator of } \xi_f(\theta) \geq E_0(\|X\|^{-2})E_{\rho_R}[q(R^2)G^{**}(R)], \tag{2.8}$$

where $G^{**}(R) = E[R^2(-h(X)) \mid \|X-\theta\| = R]$ and by assumption (ii), $G^{**}(R)$ is nondecreasing in R. Since by assumption (i), $q(R^2)$ is nondecreasing in R as well, we have

$$\text{numerator } \xi_f(\theta) \geq E_0(\|X\|^{-2})E_{\rho_R}[q(R^2)]E_{\rho_R}[G^{**}(R)]. \tag{2.9}$$

Now, by (2.2)

$$\begin{aligned} E_{\rho_R}(q(R^2)) &= \frac{\int_0^\infty q(r^2)cr^{p-3}f(r^2)dr}{E_0(\|X\|^{-2})} \\ &= \frac{1}{E_0(\|X\|^{-2})}\int_0^\infty \left(\int_{r^2}^\infty f(u)du\right) r^{p-3}dr. \end{aligned} \tag{2.10}$$

By integration by parts and since $\int_0^\infty cr^{p-1}f(r^2)dr = 1$, $E_{\rho_R}(q(R^2)) = \frac{2}{(p-2)E_0(\|X\|^{-2})}$. Finally, by assumption (iii),

$$\begin{aligned} E_{\rho_R}(G^{**}(R)) &\geq (1/2)E_{\rho_R}[R^2\|g(X)\|^2 \mid \|X-\theta\| = R] \\ &= \frac{(1/2)E_\theta[\|g(X)\|^2]}{E_0[\|X\|^{-2}]}. \end{aligned} \tag{2.11}$$

Clearly then, (2.9), (2.10) and (2.11) together imply that $\xi_f(\theta) \geq \frac{1}{(p-2)E_0(1/\|X\|^2)}$. □

Remark 2.2. Clearly, in the context of Stein estimation, spherically symmetric functions play a central role. In the case h is such a s.s. function (i.e. $h(W) = h_1(\|W\|^2)$), it is of interest to characterize the h's for which (ii) of Theorem 2.2 is satisfied. The following result (presented without proof) gives sufficient conditions for the validity of (ii) when $k = -h_1$.

PROPOSITION 2.1. *Let $k(t)$ be a non-negative function of t such that $k(t)$ is nonincreasing in t and $tk(t)$ is nondecreasing in t for $t > 0$. If $W \sim \mathcal{U}\{\|W-\theta\|^2 = R^2\}$ then, if $p \geq 4$, $E_\theta\{R^2k(W)\}$ is nondecreasing in R.*

Example 2.1. Consider the James-Stein estimator, $\delta_a(X) = X + a(-b/\|X\|^2)X$, so $g(X) = (-b/\|X\|^2)X$.

(1) $-\text{div}g = (p-2)b/\|X\|^2 = k(\|X\|^2)$. Clearly $k(t) = b(p-2)/t$ satisfies the assumptions of Proposition 2.1, so let $-h = k$.

(2) $\|g\|^2 + 2h = b^2/\|X\|^2 - 2(p-2)b/\|X\|^2 \leq 0$, if $0 \leq b \leq 2(p-2)$. So, if $X \sim$ s.s.d.(θ) with $q(t)$ nonincreasing, by Theorem 2.1 and Remark 2.1, $\delta_a(X)$ dominates X for $0 < ab \leq (2(p-2)/p)E_0(\|X\|^2)$ provided $p \geq 4$. If $X \sim$ s.s.d.(θ) with $q(t)$ nondecreasing, by Theorem 2.2 and Remark 2.2, $\delta_{a,g}(X)$ dominates X for $0 < ab \leq 2/E_0(1/\|X\|^2)$. These bounds coincide with those of Bock (1985).

Note. When $\delta_{a,g}(X)$ (1.1) is the James-Stein estimator, $\xi_f(0) = (2/b) \cdot (1/E_0(\|X\|^{-2}))$ by integration by parts.

Since $0 < b \leq 2(p-2)$, $\xi_f(0) \geq \frac{1}{(p-2)E_0(\|X\|^{-2})}$. So, for the class $\mathcal{F}_1 = \{f : q_f(t)$ is nondecreasing on $(t : f(t) > 0)\}$, we obtain the best bound in Theorem 2.2, when $\delta_a(X)$ is the James-Stein estimator.

Note. For any θ, we can restrict our attention to $\theta = [\|\theta\|, 0, 0, \ldots, 0]$ and for $\delta_{a,g}(X)$, the James-Stein estimator, $\lim_{\theta\to\infty} \xi_f(\theta) = (2(p-2)/bp)E_0(\|X\|^2) \geq E_0(\|X\|^2)/p$. So, for the class $\mathcal{F}_2 = \{f : q_f(t)$ is nonincreasing on $(t : f(t) > 0)\}$, we obtain the best bound in Theorem 2.1 when $\delta_a(X) = X + a(-b/\|X\|^2)X$.

Even for h, nonspherically symmetric, there are simple conditions under which the assumptions of Theorems 2.1 and 2.2 will be satisfied.

Remark 2.3. If k is homogeneous of degree -2 (i.e., $k(W) = (1/s^2)k(W/s)$), then for $W \sim \mathcal{U}\{\|W-\theta\|^2 = R^2\}$, $E_\theta[R^2k(W)]$ is nondecreasing in R and letting $k = -h$, condition (ii) of Theorem 2.2 is satisfied.

Note. If $g : R^p \to R^p$ is homogeneous of degree -1, then all derivatives $\partial g/\partial x_i$ and thus div$g$ will be homogeneous of degree -2.

Example 2.2. "Limited Translation Rule" for special spherically symmetric distributions.

Suppose we consider the "limited translation" rule based on order statistics given by Stein (1981) for $X \sim$ s.s.(θ) and $q(t)$ nonincreasing or nondecreasing. For k a positive integer, let $\delta_{a,g}(X)$ (1.1) be defined by

$$g_i(X) = \begin{cases} \dfrac{-bX_i}{\sum(X_j^2 \wedge Z_{(k)}^2)}, & \text{if } |X_i| \leq Z_{(k)} \\ \dfrac{-b}{\sum(X_j^2 \wedge Z_{(k)}^2)} Z_{(k)}\text{sgn}X_i, & \text{if } |X_i| > Z_{(k)} \end{cases}$$

and $Z_i = |X_i|$, $Z_{(1)} < Z_{(2)} < \cdots < Z_{(p)}$ are the order statistics and $c \wedge d = \min(c,d)$. For this estimator

(1) g is homogeneous of degree -1, and

(2) $-\text{div}g = (k-2)b/\sum(X_i^2 \wedge Z_{(k)}^2)$ according to Stein (1981). So, if $-\text{div}g = -h$, clearly $-h$ is superharmonic (if $b > 0$ and $p \geq 4$), and

(3) $\|g\|^2 + 2h = (b^2 - 2(k-2))(1/\sum(X_j^2 \wedge Z_{(k)}^2)) \leq 0$ if $0 < b \leq 2(k-2)$. Therefore, for $q(t)$ nonincreasing in t, by Theorem 2.1 and Remark 2.1, $\delta_{a,g}(X)$ is minimax for $0 < ab \leq (2(k-2)/p)E_0(\|X\|^2)$, provided $p \geq 4$, and for $q(t)$ nondecreasing in t, by Theorem 2.2 and Remark 2.2, $\delta_{a,g}(X)$ is minimax for $0 < ab \leq \frac{2(k-2)}{(p-2)E_0(\|X\|^{-2})}$ provided $p \geq 4$.

Consider now the problem of estimating θ, the mean of a special class of spherically symmetric distributions, namely spherically symmetric unimodal distributions (i.e., $X \sim$ s.s.u.(θ)). A $p \times 1$ random vector X is said to have a s.s.u. distribution about θ if the density $f(\|x-\theta\|^2)$ with respect to Lebesgue measure is a nonincreasing function. If $X \sim$ s.s.u.(θ), then $X \mid R \sim \mathcal{U}\{\|X-\theta\|^2 \leq R^2\}$.

If a p-dimensional random vector X has a uniform distribution on the ball $\|X-\theta\|^2 \le R^2$, with known radius R, then $f(\|x-\theta\|^2) = (p/cR^p)I(\|x-\theta\|^2 \le R^2)$ where c is the proper constant that makes $f(\|x-\theta\|^2)$ a density. Thus, for $X \sim \mathcal{U}\{\|X-\theta\|^2 \le R^2\}$,

$$(2.12) \qquad q(t) = \int_t^\infty I(u \le R^2)du, \qquad \text{for} \quad f(t) > 0 = \int_t^{R^2} du = R^2 - t$$

which is a nonincreasing function of t.

From this fact and the results of the proof of Theorems 2.1, by conditioning on the radius of the ball, in the following theorem we will find improved minimax estimators of the form $\delta_{a,g}(X)$ (1.1) when $X \sim$ s.s.u.(θ). The bound on a will be slightly better than the general spherically symmetric bound found by Brandwein and Strawderman (1991).

THEOREM 2.3. *If the* $p \times 1$ *random vector* $X = [X_1, X_2, \ldots, X_p]'$ *has a s.s.u.*(θ) *distribution and* $\delta_{a,g}(X)$ *is defined by* (1.1), *then with respect to quadratic loss,* $\delta_{a,g}(X)$ *dominates* X *provided:*

(i) $-\text{div}g \ge -h$ *where* h *is a function such that if* $W \sim \mathcal{U}\{\|W-\theta\|^2 = R^2\}$, $E_\theta[h(W)]$ *is nondecreasing in* R *and* $E_\theta[R^2 h(W)]$ *is nonincreasing in* R,

(ii) $\|g\|^2 + 2h \le 0$, *and*

(iii) $0 < a \le \frac{p}{(p^2-4)E_0(1/\|X\|^2)}$.

PROOF. If $p \times 1$ random vector $Y \sim \mathcal{U}\{\|Y-\theta\|^2 \le R^2\}$, then since by (2.12), $q(t)$ is nonincreasing, and conditions (ii) and (iii) of Theorem 2.1 are satisfied, similarly as in the proof of Theorem 2.1

$$(2.13) \qquad \text{numerator } \xi_f(\theta) \ge \frac{2}{p} E_0[\|Y\|^2] E_\theta[-h(Y)] = \frac{2}{p+2} R^2 E_\theta[-h(Y)].$$

For $X \sim$ s.s.u.(θ), by (2.13) and since $Y = X \mid R$,

$$(2.14) \qquad \begin{aligned} \xi_f(\theta) &= \frac{E[E_\theta[-\text{div}g(X)q(\|X-\theta\|^2) \mid R]]}{E_\theta[\|g(X)\|^2]} \\ &\ge \frac{2}{p+2} \frac{E[R^2 E_\theta[-h(X) \mid R]]}{E_\theta[\|g(X)\|^2]} \\ &= \frac{2}{p+2} \frac{E[R^2 B_{\theta,R}(-h)]}{E_\theta[\|g(X)\|^2]} \end{aligned}$$

where

$$(2.15) \quad \begin{aligned} B_{\theta,R} = E_\theta[-h(X) \mid R] &= \frac{p}{cR^p} \int_{\{\|x-\theta\|^2 \le R^2\}} -h(x)dx \\ &= \frac{p}{R^p} \int_0^R \rho^{p-1} E_\theta[-h(X) \mid \|X-\theta\| = \rho] d\rho. \end{aligned}$$

Since by assumption (i), $E_\theta[-h(X) \mid \|x-\theta\| = \rho]$ is nonincreasing in ρ,

$$B_{\theta,R}(-h) \geq E_\theta[-h(X) \mid \|X-\theta\| = R]. \tag{2.16}$$

Thus, (2.14) and (2.16) imply

$$\xi_f(\theta) \geq \frac{2}{p+2} \frac{E[R^2 E_\theta[-h(X) \mid \|X-\theta\| = R]]}{E_\theta[\|g(X)\|^2]}. \tag{2.17}$$

Now, by assumption (i) $R^2 E_\theta[-h(X) \mid \|X-\theta\| = R]$ is nondecreasing in R, and by assumption (ii)

$$\begin{aligned} &E[R^2 E_\theta[-h(X) \mid \|X-\theta\| = R]] \\ &= \frac{1}{E(1/R^2)}[E(1/R^2)]E[R^2 E_\theta[-h(X) \mid \|X-\theta\| = R]] \\ &\geq \frac{1}{E(1/R^2)} E_\theta[-h(X)] \geq \frac{1}{2E(1/R^2)} E_\theta[\|g(X)\|^2]. \end{aligned}$$

Hence, (2.17) becomes $\xi_f(\theta) \geq 2/(p+2)E(1/R^2)$. But, $E(1/R^2) = ((p-2)/p) \cdot E_0(1/\|X\|^2)$ when $X \sim$ s.s.u.(θ). Thus, $\xi_f(\theta) \geq (p/(p^2-4))(1/E_0(1/\|X\|^2))$. □

Note. From Example 2.1, it is easy to show that when $\delta_{a,g}(X)$ is the James-Stein estimator and $p \geq 4$, from Theorem 2.3, we will obtain the Brandwein-Strawderman (1978) bound $(2p/(p+2))(1/E_0(1/\|X\|^2))$ for s.s.u. distributions.

Remark 2.4. For $k(W) = k_1(\|W\|^2) = -h(W)$ we can apply the following proposition along with Remark 2.2 to verify Condition (i) of Theorem 2.3.

PROPOSITION 2.2. *Let $W \sim \mathcal{U}\{\|W-\theta\|^2 = R^2\}$ and define $F(R) = E_\theta k_1 \cdot (\|W\|^2)$. If*

$$t^{pi/2} \frac{d^i}{dt^i} k_1(t) \text{ is nonincreasing in } t \text{ for } i = 0, 1, \tag{2.18}$$

then $F(R)$ is nonincreasing in R.

Clearly, there exist functions k_1 satisfying (2.18) and such that $k_1(\|W\|^2)$ is not superharmonic. For such k_1, if $tk_1(t)$ is nondecreasing, condition (i) of Theorem 2.3 will be satisfied.

Note. If k_1 is twice differentiable and satisfies (2.18), then k_1 is superharmonic.

If h is not spherically symmetric, we can apply the following remark to find improved minimax estimates using Theorem 2.3.

Remark 2.5. If k is superharmonic and homogeneous of degree -2 (i.e. $k(W) = \frac{1}{S^2}k(\frac{W}{S})$ and $W \sim \mathcal{U}\{\|W-\theta\|^2 = R^2\}$ then $E_\theta k(W)$ is nondecreasing in R and $R^2 E_\theta k(W)$ is nondecreasing in R.

By restricting our attention to special classes of spherically symmetric distributions in Theorems 2.1, 2.2 and 2.3, we were able to find better minimax estimators than the ones for the general s.s. case found by Brandwein and Strawderman (1991).

Ralescu *et al.* (1991) define another special s.s. distribution, namely a "controlled" spherically symmetric unimodal (s.s.u.) distribution as follows:

DEFINITION. We say that X has a controlled s.s.u. density $f(\|x-\theta\|^2)$ if there exists a pair $(f_1(\cdot)\downarrow, m>0)$ such that for all $t>0$, $f_1(t)\le f(t)\le mf_1(t)$.

For f, s.s.u. controlled, the following theorem finds an improved bound for a, than the general spherically symmetric case domination bound found in Brandwein and Strawderman (1991).

THEOREM 2.4. *Let X be a p-dimensional observation from a controlled s.s.u. density $f(\|x-\theta\|^2)$, where $(f_1(t),m)$ is a pair controlling f. Then $\delta_{a,g}(X)$* (1.1) *has smaller risk than X with respect to quadratic loss provided conditions* (i) *and* (ii) *in Theorem* 2.3 *hold and*

$$0<a\le\frac{p}{(p^2-4)m^2E_0(\|X\|^{-2})}.$$

PROOF. From (2.1) and definition of s.s.u. controlled

$$\begin{aligned}\xi_f(\theta)&=\frac{\int -\mathrm{div}g(x)(\int_{\|x-\theta\|^2}^{\infty}f(t)dt)dx}{\int\|g(x)\|^2f(\|x-\theta\|^2)dx}\\&\ge\frac{1}{m}\frac{\int -\mathrm{div}g(x)\int_{\|x-\theta\|^2}^{\infty}\tilde f_1(t)dt)dx}{\int\|g(x)\|^2\tilde f_1(\|x-\theta\|^2)dx}\end{aligned}\tag{2.19}$$

where $\tilde f_1(t)=f_1(t)/c$ and $c=\int f_1(\|x\|^2)dx\le\int f(\|x\|^2)dx=1$. Now, if $Y=(Y_1,Y_2,\ldots,Y_p)\sim\tilde f_1(\|y-\theta\|^2)$, then $Y\sim$ s.s.u.(θ) and (2.19) and Theorem 2.3 together imply

$$\begin{aligned}\xi_f(\theta)\ge\frac{p}{m(p^2-4)}1/E_0(1/\|Y\|^2)&=\frac{p}{m(p^2-4)}\frac{\int f_1(\|x-\theta\|^2)dx}{\int(1/\|x\|^2)f_1(\|x-\theta\|^2)dx}\\&\ge\frac{1}{m^2}\frac{p}{(p^2-4)}1/E_0(1/\|X\|^2).\end{aligned}\qquad\square$$

Remark 2.6. (Robustness) A strong motivation for the interest in s.s.d.'s derives from the desire to broaden the scope of the linear regression model $X=A\theta+\epsilon$ to allow for more robustness in the treatment of the random error ϵ. For this model, it was traditionally assumed in the literature that ϵ is normally distributed. However, very often this normality assumption has been criticized as being too restrictive since one of the main concerns of the statistician is related to his uncertainty about the true functional form of the distribution. It can be argued that, since in practice it is rare that the actual form of the distribution of ϵ is known, instead of the normality assumption a much weaker assumption of spherical symmetry is considered more plausible. By appropriate transformations it can be assumed that $X = \theta + \epsilon$ and the distribution of X has a s.s. density. In this context one needs to cope with the degree of uncertainty associated with the lack of advanced knowledge of f and an important aspect is the robustness of the Stein effect (see Cellier *et al.* (1989)) i.e. the domination of the lse by shrinkage estimators uniformly over a class of s.s.d.'s. Here our aim is to use our theorems and establish that, for a given class $\mathcal{F}$ of s.s.d.'s there exists $a_0 > 0$ such that for all $a \in (0, a_0]$ and all $g \in \mathcal{G}$ (= a class of "shrinkers") the usual lse X can be dominated by $\delta_{a,g}$ no matter which s.s.d. in the class $\mathcal{F}$ is sampled. The fact that one does not have to specify precisely the distribution of the error but only to assume that it belongs to a certain class of s.s.d.'s (thus removing the much restrictive normality assumptions) has deep practical consequences allowing the experimenter a wider degree of flexibility. Before we present our examples, it is noteworthy that from condition (H4) in Cellier *et al.* ((1989), p. 48) there is a restriction on the s.s.d.'s to which their Proposition 5.2 applies. In fact, their condition implies that X is sampled from a class of s.s.d.'s which have tails that are flatter than the normal (see also Berger (1975)).

Our examples below provide a step in the robustness directed at generalized Stein estimators.

Example 2.3. Let

$$\mathcal{G} = \left\{ g : \|g\|^2 + 2\text{div}g \le 0 \text{ and } (\exists)\ h \right.$$
$$\left. \text{with} \left\{ \begin{array}{l} (1) - \|g\|^2/2 \le -h \le \text{div}g, \text{ and} \\ (2)\ E_\theta[h(W)] \text{ is nondecreasing in } R \\ \qquad \text{when } W \sim \mathcal{U}\{\|W-\theta\|^2 = R^2\} \end{array} \right\} \right\}$$

and for $k > 0$

$$\mathcal{F}_k = \left\{ f : q_f(\cdot) \text{ is} \downarrow \text{ and } \int_0^\infty r^{p+1} f(r^2)dr \ge k > 0 \right\}.$$

Then by Theorem 2.1, $(\exists)\ a_0 (= ck/p) > 0\ s.t. (\forall) a \in (0, a_0]$ and $(\forall)\, g \in \mathcal{G}$, $\delta_{a,g}$ dominates X uniformly over $\mathcal{F}_k$.

Note. Some simple conditions which insure the monotonicity of $q_f(\cdot)$ in $\mathcal{F}_k$ are given in Bock ((1985), Lemma 3).

It is worth mentioning that $\mathcal{F}_k$ contains all s.s. uniform on balls with radii bounded from below by a level $\gamma = \gamma_k > 0$. Therefore Proposition 5.2 in Cellier *et al.* (1989) cannot be applied to show dominance in this case.

Example 2.4. Let $\alpha(\cdot)$ be a nondecreasing function and for $m > 0$, set $\mathcal{F}_{(\alpha,m)} = \{\text{all } f = \text{s.s. unimodal controlled by } (\alpha, m)\}$. Set $\mathcal{G}_1 = \{g \in \mathcal{G} : \text{the } h\text{-}$

function associated with g satisfies also condition (3) $R^2E[h(W)]$ is nonincreasing when $W \sim \mathcal{U}\{\|W-\theta\|^2 = R^2\}\}$. Then there exists $a_0 > 0$ such that $(\forall)a \in (0, a_0]$ and $(\forall)g \in \mathcal{G}$, $\delta_{a,g}$ dominates X uniformly over $\mathcal{F}_{(\alpha,m)}$. Indeed, according to the proof of Theorem 2.3 and in view of Lemma 2.1 it is enough to take $a_0 = \frac{p}{m(p^2-4)} \inf_\theta \xi_\alpha(\theta) > 0$.

Many other examples along similar lines could be considered including the class of s.s.d.'s with compact support $\mathcal{F}_{c_1,c_2,d} = \{f : c_1 I_{\{f(t)\le d\}} \le f(t) \le c_2 I_{\{f(t)\le d\}}\}$ where I_A denotes the indicator function and $c_1, c_2, d > 0$. For the sake of brevity we omit the details.

3. Improved estimators for spherically symmetric distributions via integration by parts technique

Consider a $p\times 1$ random vector $X = (X_1, X_2, \ldots, X_p)'$ having a p-dimensional s.s.d. (θ). We seek general conditions for $\delta_{a,g}(X)$ (1.1) to be minimax. Our result is essentially that of Brandwein and Strawderman (1991), although there are some interesting remarks concerning the specific conditions which lead to some special examples.

THEOREM 3.1. *If the $p \times 1$ random vector $X \sim$ s.s.d.(θ) then with respect to quadratic loss, $\delta_{a,g}(X)$ (1.1) has smaller risk than X provided:*

(i) $-\text{div}g \ge -h$ *where h is a function such that if $W \sim \mathcal{U}\{\|W-\theta\|^2 = R^2\}$, $E_\theta[h(W)]$ is nondecreasing in R and $E_\theta[R^2h(W)]$ is nonincreasing in R,*

(ii) $\|g\|^2 + 2h \le 0$, *and*

(iii) $0 < a \le \frac{1}{pE_0(1/\|X\|^2)}$, *and $E_0(1/\|X\|^2)$ exists and is finite.*

PROOF. The numerator of $\xi_f(\theta)$, defined by (2.1), equals

$$\text{(3.1)} \qquad \int_0^\infty \left[\int_{\|x-\theta\|^2\le t} (-\text{div}g(x))dx\right] f(t)dt$$

$$= 2\int_0^\infty \left(\int_{\|x-\theta\|^2\le r^2} (-\text{div}g(x))dx\right) rf(r^2)dr$$

by change of variables

$$\ge 2\int_0^\infty \left(\int_{\|x-\theta\|^2\le r^2} (-h(x))dx\right) rf(r^2)dr$$

by assumption (i).

If $f_R(r)$ is the density of $R = \|x-\theta\|$ as defined in (2.5), then by (3.1),

$$\text{(3.2)} \qquad \text{numerator of } \xi_f(\theta) \ge \frac{2}{p}\int_0^\infty r^2 B_{\theta,r}(-h) f_R(r)dr$$

where

$$\text{(3.3)} \qquad B_{\theta,r}(-h) = \frac{p}{cr^p}\int_{\|x-\theta\|^2\le r^2} -h(x)dx.$$

By (2.16), $B_{\theta,r}(-h) \geq E_\theta[-h(X) \mid \|x-\theta\| = r]$. This together with (3.2) implies

$$(3.4) \quad \xi_f(\theta) \geq \frac{\frac{2}{p}\int_0^\infty r^2 E_\theta[-h(X) \mid \|X-\theta\| = r] f_R(r)dr}{E_\theta[\|g(X)\|^2]}$$

$$= \frac{2}{pE_{f_R}(1/R^2)} \frac{E_{f_R}\left(\frac{1}{R^2}\right) E_{f_R}[R^2 E_\theta[-h(X) \mid \|X-\theta\| = R]]}{E_\theta[\|g(X)\|^2]}.$$

By assumption (i), $[R^2 E_\theta[-h(X) \mid \|X-\theta\| = R]]$ is nondecreasing in R and $1/R^2$ is nonincreasing in R. Therefore, these facts together with assumption (ii) imply (3.5)

$$E_{f_R}\left[\frac{1}{R^2}\right] E_{f_R}[R^2 E_\theta[-h(X) \mid \|x-\theta\| = R]] \geq E_\theta[-h(X)] \geq \frac{1}{2} E_\theta[\|g(X)\|^2].$$

Hence, by (3.4) and (3.5), $\xi_f(\theta) \geq \frac{1}{pE_0(1/\|X\|^2)}$ and so $\inf_\theta \xi_f(\theta) \geq \frac{1}{pE_0(1/\|X\|^2)}$. □

Note that the conditions in Theorem 3.1 are for uniform distributions on a shell of a sphere and those in Brandwein and Strawderman (1991) relate to uniform distributions on a ball. In many of the examples in Brandwein and Strawderman (1991), it was necessary to show the unimodality of h and apply Anderson's theorem to show $E_\theta[R^2 h(W)]$ is nonincreasing in R for W uniform on the ball of radius R. By having conditions on the shell, the integration by parts technique bypasses the unimodality requirement.

The remarks and examples in Section 2 can also be applied to Theorem 3.1 and the bounds adjusted accordingly.

4. Minimax estimators for other loss functions

Consider X a $p \times 1$ random vector having a s.s.d.(θ) or another one of the special s.s. distributions in Section 2. With respect to quadratic loss, for $p \geq 3$, $\delta_{a,g}(X)$ (1.1) is better than X under certain conditions stated in the theorems of Sections 2 and 3.

Let us now consider generalizing these results for quadratic loss to concave loss functions of quadratic loss and general quadratic loss.

Specifically, consider the nonquadratic loss of the form

$$(4.1) \qquad L(\delta,\theta) = \ell(\|\delta-\theta\|^2)$$

where $\ell(\cdot)$ is a nondecreasing concave function and

$$(4.2) \qquad L(\delta,\theta) = (\delta-\theta)' D(\delta-\theta)$$

where D is a given $p \times p$ positive definite matrix.

THEOREM 4.1. *Let the $p\times 1$ random vector X have a p-dimensional, s.s.d.(θ) and the loss be concave loss, $\ell(\|X-\theta\|^2)$, given by (4.1). Then, $\delta_{a,g}(X)$ (1.1) is* *better than* X *provided conditions* (i)–(ii) *of Theorem* 3.1 *are satisfied and* $0 < a \leq (1/p)(E_0\ell'(\|X\|^2)/E_0(\ell'(\|X\|^2)/\|X\|^2))$.

PROOF. Since $\ell(\cdot)$ is a nondecreasing concave function, $\ell'(t)$ is nonnegative and nonincreasing, so $\ell''(t) \leq 0$. Moreover, for any points s and t, $\ell(s) < \ell(t) + \ell'(t)(s-t)$. If $\Delta_{\delta_{a,g}}(X) = \|X-\theta\|^2 - \|\delta_a(X)-\theta\|^2$, then

$$\ell(\|X-\theta\|^2 - \Delta_{\delta_{a,g}}(X)) < \ell(\|X-\theta\|^2) + \ell'(\|X-\theta\|^2)(-\Delta_{\delta_{a,g}}(X)). \tag{4.3}$$

Hence,

$$\begin{aligned} R(\delta_0,\theta) - R(\delta_{a,g},\theta) &\geq E_\theta(\ell'(\|X-\theta\|^2)\Delta_{\delta_a}(X)) \\ &= \left[\int \ell'(\|x-\theta\|^2)f(\|x-\theta\|^2)dx\right] E_{f^*}\Delta_{\delta_{a,g}}(X), \end{aligned} \tag{4.4}$$

where

$$f^*(\|x-\theta\|^2) = \frac{\ell'(\|x-\theta\|^2)f(\|x-\theta\|^2)}{\int \ell'(\|x-\theta\|^2)f(\|x-\theta\|^2)dx} \tag{4.5}$$

is a density since $\ell'(\|x-\theta\|^2)$ is nonnegative.

With respect to this density, X has a s.s. distribution and from the results of Theorem 3.1, $E_{f^*}\Delta_{\delta_{a,g}}(X) \geq 0$ for $0 < a \leq (1/p)(1/E_{f^*}(1/\|X-\theta\|^2)) = (2/p) \cdot (E_0\ell'(\|X\|^2)/E_0(\ell'(\|X\|^2)/\|X\|^2))$ and so by (4.4), $R(\delta_0,\theta) - R(\delta_{a,g},\theta) \geq 0$ for these values of a as well. □

Define

$$q_\ell(t) = \frac{\int_t^\infty \ell'(u)f(u)du}{\ell'(t)f(t)}, \qquad \text{for} \quad f(t) > 0. \tag{4.6}$$

When $\ell(\|\delta-\theta\|^2)$ is quadratic loss, $\ell'(t) = 1$, and $q_\ell(t)$ coincides with $q(t)$ defined by (2.2).

In the following theorems, we consider spherically symmetric distributions for which $q_\ell(t)$ is nonincreasing and $q_\ell(t)$ is nondecreasing when estimating θ with respect to concave loss functions of quadratic loss. Since the proofs following directly from the proofs of Theorems 2.1, 2.2 and 4.1, we omit them.

THEOREM 4.2. *If the* $p \times 1$ *random vector* $X \sim$ *s.s.d.*(θ), *then with respect to loss* (4.1), $\delta_{a,g}(X)$ *dominates* X *provided* $q_\ell(t)$ *is nonincreasing, conditions* (ii) *and* (iii) *of Theorem* 2.1 *hold, and*

$$0 < a \leq \frac{1}{p}\frac{E_0(\ell'(\|X\|^2)\|X\|^2)}{E_0(\ell'(\|X\|^2))}.$$

THEOREM 4.3. *Suppose the* $p \times 1$, $p \geq 3$, *random vector* $X \sim$ *s.s.d.*(θ) *and* $q_\ell(t)$ *defined by* (4.6) *is nondecreasing. If assumptions* (ii) *and* (iii) *of Theorem* 2.2 *hold, then for*

$$0 < a \leq \frac{1}{(p-2)} \frac{E_0 \ell'(\|X\|^2)}{E_0(\ell'(\|X\|^2)/\|X\|^2)},$$

$\delta_{a,g}(X)$ (1.1) *has (uniformly) smaller risk than* X.

Remark 4.1. If the $p \times 1$ random vector $X \sim$ s.s.u.(θ) under the same conditions as in Theorem 2.3, $\delta_{a,g}(X)$ (1.1) is better than X with respect to concave loss (4.1) for

$$0 < a \leq \frac{p}{(p^2-4)} \frac{E_0 \ell'(\|X\|^2)}{E_0(\ell'(\|X\|^2)/\|X\|^2)}.$$

This follows directly from (4.4) since $\ell(\|X-\theta\|^2)$ is concave loss, density $f^*(\|x-\theta\|^2)$ defined by (4.5) is nonincreasing and so the random variable X has a s.s.u. distribution.

Consider now the problem of estimating θ with respect to general quadratic loss given by (4.2). With respect to this loss we can restate Lemma 2.1 as follows.

LEMMA 4.1. *Let the* $p \times 1$ *random vector* X *have a spherically symmetric distribution about* θ *with density* $f(\|x-\theta\|^2)$ *with respect to Lebesgue measure. Then, provided divergence of* Dg *is nonpositive, the risk of the estimator* $\delta_{a,g}(X)$ (1.1) *dominates the risk of* $\delta_0(X)$ *with respect to general quadratic loss* (4.2) *if and only if* $0 < a \leq \inf_\theta \xi_{f,D}(\theta)$ *where*

$$\xi_{f,D}(\theta) = \frac{\int -(\text{div} Dg(x))(\int_{\|x-\theta\|^2}^{\infty} f(t)dt)dx}{\int g'(x)Dg(x)f(\|x-\theta\|^2)dx}.$$

Thus, in Theorems 2.1, 2.2, 2.3, 2.4 *and* 3.1, *if we change the conditions that* $-\text{div} g \geq -h$ *to* $-\text{div} Dg \geq -h$ *and* $\|g\|^2 + 2h \leq 0$ *to* $g'Dg + 2h \leq 0$, *we will obtain the same improved minimax estimators with respect to general quadratic loss.*

5. Minimax estimators for spherically symmetric distributions with an unknown scale

Suppose the $p \times 1$ random vector X has a density $(1/\sigma^p)f(\|x-\theta\|^2/\sigma^2)$ where σ is unknown, and consider the random variable V, with density $(1/\sigma^2)f(v/\sigma^2)$, independent of X. Consider first estimating θ with respect to scale quadratic loss

$$L(\delta, \theta) = \|\delta - \theta\|^2/\sigma^2. \tag{5.1}$$

It is straightforward to show that Lemma 2.1 holds for estimators $\delta_{a,V,g}(X) = X + aVg(X)$, provided $0 < a \leq \inf_\theta \xi_f^*(\theta)$ where $\xi_f^*(\theta) = (E_{\sigma=1}V/E_{\sigma=1}V^2)\xi_f(\theta)$ and $\xi_f(\theta)$ is defined by (2.1).

The results of Sections 2, 3 and 4 need only be modified for the unknown variance case by multiplying the upper bounds of a by $E_{\sigma=1}V/E_{\sigma=1}V^2$. Thus we have improved minimax estimators for all s.s. distributions and for some special ones with respect to scale quadratic loss, concave loss functions of scale quadratic loss and general scaled quadratic loss when the variance is unknown. Because of Lemma 2.1 and integration by parts, this is a straightforward extension of the known variance case.

Acknowledgement

The authors would like to thank the referee for a careful reading and valuable suggestions which helped to improve the original version of the paper.

References

Akai, T. (1986). Simultaneous estimation of location parameters of the distributions with finite support, *Ann. Inst. Statist. Math.*, **38**, 85–99.

Berger, J. (1975). Minimax estimation of location vectors for a wide class of densities, *Ann. Statist.*, **3**, 1318–1328.

Bock, M. E. (1985). Minimax estimators that shift towards a hypersphere for location vectors of spherically symmetric distributions, *J. Multivariate Anal.*, **17**, 127–147.

Brandwein, A. C. and Strawderman, W. E. (1978). Minimax estimation of location parameters for spherically symmetric unimodal distributions under quadratic loss, *Ann. Statist.*, **6**, 377–416.

Brandwein, A. C. and Strawderman, W. E. (1991). Generalizations of James-Stein estimators under spherical symmetry, *Ann. Statist.*, **19**, 1639–1650.

Cellier, D., Fourdrinier, D. and Robert, C. (1989). Robust shrinkage estimators of the location parameter for elliptically symmetric distributions, *J. Multivariate Anal.*, **29**, 39–52.

Chou, J. and Strawderman, W. E. (1990). Minimax estimation of the mean of multivariate normal mixtures, *J. Multivariate Anal.*, **35**, 141–150.

Ralescu, S., Brandwein, A. C. and Strawderman, W. E. (1992). Stein estimation for non-normal, spherically symmetric location families in three dimensions, *J. Multivariate Anal.*, **42**, 35–50.

Stein, C. (1973). Estimation of the mean of a multivariate normal distribution, *Proc. Prague Symp. on Asymptotic Statistics*, 345–381.

Stein, C. (1981). Estimation of the mean of a multivariate normal distribution, *Ann. Statist.*, **9**, 1135–1151.