# Minimax Estimators that Shift Towards a Hypersphere for Location Vectors of Spherically Symmetric Distributions

by

M. E. Bock*

Technical Report #81-51

Department of Statistics
Purdue University

revised June 1983

*This research was supported by NSF grant #MCS-81-01670.

# Minimax Estimators that Shift Towards a Hypersphere for Location Vectors of Spherically Symmetric Distributions

by

M. E. Bock *

Department of Statistics

Purdue University

*This research was supported by NSF grant MCS-81-01670.

Running Title: Minimax Estimators to a Hypersphere

Address: M. E. Bock
Department of Statistics
Mathematical Sciences Building
Purdue University
West Lafayette, IN 47907

ABSTRACT

Let X be a p-dimensional random vector with density $f(||X-\theta||)$ where $\theta$ is an unknown location vector. For $p \geq 3$, conditions on f are given for which there exist minimax estimators $\hat{\theta}(X)$ satisfying $||X|| \cdot ||\hat{\theta}(X)-X|| \leq C$, where C is a known constant depending on f. (The positive part estimator is among them.) The loss function is a nondecreasing concave function of $||\hat{\theta}-\theta||^2$. If $\theta$ is assumed likely to lie in a ball in $\mathbb{R}^p$, then minimax estimators are given which shrink from the observation X in the direction of P(X) the closest point on the surface of the ball. The amount of shrinkage depends on the distance of X from the ball.

Key Words and Phrases. Minimax Estimation, spherically symmetric, multivariate, shrinkage estimator, location vector, positive part estimator.

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

## Section 1. Introduction

The problem considered is that of estimating the p-dimensional location vector $\theta$ for a p-dimensional random vector X under the nondecreasing non-negative concave loss function $\ell$, i.e.,

$$L(\theta,\hat{\theta}) = \ell(||\hat{\theta}-\theta||^2),$$

where "$||\cdot||$" denotes the Euclidean norm. It is assumed that p is three or more and that the distribution of X is spherically symmetric about $\theta$. Assume that $E_{\theta=0}[||X||^{-2}]$ and $E_{\theta=0}[||X||^2]$ and $E_{\theta=0}[\ell(||X||^2)]$ exist and are finite.

Because the minimax invariant estimator X has constant risk for all values of $\theta$, a reasonable guess as to the location of $\theta$ may be very useful in specifying a better minimax estimator. That is, one may use a minimax estimator which has the property that the smallest values of the risk function occur for values of $\theta$ close to the "guessed" location of $\theta$; yet all the values of the risk function are less than or equal to the constant risk of the estimator X. When the "guess" is the specification of a single vector, the minimax estimators of Section 2 have this property. (In subsection b it is assumed that the "guess" is that $\theta$ is the zero vector but it can clearly be any vector.) Consider the situation where the "guess" is that $\theta$ is likely to lie in a ball of radius G centered at the vector $\beta$ (i.e., $||\theta-\beta||\leq G$). For values of $\theta$ outside the ball the estimators of Section 3 have smaller risk when $\theta$ is closer to the ball. For values of $\theta$ inside the ball, the estimators in

Section 3 may be specified so that the actual loss (rather than the expected loss or risk) is less than or equal to that of X no matter what the distribution of X. It should be noted that results for the closed convex polyhedron like those for the ball in Section 3 have been obtained for the normal distribution in [2] for squared error loss. (In the case that the "guess" is that $\theta$ is likely to satisfy a finite system of linear inequalities, then $\theta$ is thought likely to lie in the closed convex polyhedron that satisfies that finite system of linear inequalities.)

It is assumed that the density of the random vector X may be represented by the function $f(||X-\theta||)$. In Section 2 various conditions on the density function f of X and the loss function $\ell$ are presented under which there exist minimax estimators $\hat{\theta}(X)$ satisfying $||X||\cdot||\hat{\theta}(X)-X||\leq C$ where C is a known constant depending on f. In particular, $\hat{\theta}(X)$ has the form $(1-r(||X||^2)/||X||^2)X$ where r is a nondecreasing function such that $\{r(t)/t\}$ is nonincreasing, and $0\leq r(t)\leq C$. The conditions on f allow larger values of C than were previously known. (See Brandwein [3], Brandwein and Strawderman [4], Berger [1] and Strawderman [6].) For instance, under squared error loss a larger class of minimax estimators (i.e. a larger value of C) is given when the derivative of $\log f(t^{1/2})$ is monotone as a function of t.

The estimators in subsection a of Section 2 are "positive part" estimators analyzed under squared error loss and shift the estimate from X in the direction of a previously chosen vector $\beta$. The estimators in subsection b are more general and so is the loss. These estimators may be chosen so that they differ from the invariant estimator X in that they shift the estimate from X in the direction towards the vector zero. In Section 3 any chosen sphere or ball may play the role of the vector zero. For values of X outside the chosen sphere the estimators in Section 3 may be specified to shift the estimate from X in the direction towards the closest vector P(X) on the surface of that sphere. All the estimators of Section 3 dominate estimators which take the value X outside the sphere. Similar results for any chosen closed convex polyhedron are obtained for the normal distribution in [2] under squared error loss.

Although the discussions in this paper are presented only for a single observation vector from the spherically symmetric distribution, Brandwein [3] has noted that such results apply to spherically symmetric translation invariant estimators in the multiple observation case.

## Section 2: An enlarged class of minimax estimators for nondecreasing concave loss functions.

The theorems and corollaries in this section give general conditions on the density function f for the domination of the estimator X by a family of estimators whose distance from X may be as great as $C/||X||$, i.e. $||\hat{\theta}(X)-X||\leq C/||X||$. Because the estimator X is minimax with constant risk for all values of $\theta$, the estimators $\delta$ given in the theorems and corollaries are also minimax.

A useful example of such a dominating estimator is given by the positive part estimator analyzed in subsection a for the simple case of squared error loss. The theorems in subsection b give the most general results.

### Subsection a: Enlarged families of minimax positive part estimators

Let X be a p-dimensional random vector with density $f(||X-\theta||)$. If it is considered likely that the location vector $\theta$ for the distribution of X is a particular vector $\beta$, then a simple estimator which takes account of this "vague prior information" is the positive part estimator $\hat{\theta}_c^+$. It is an appealing aspect of the positive part estimator that for values of X close to $\beta$ the estimator is equal to $\beta$.

Define

$$\hat{\theta}_c^+(X) = \begin{cases} (1-c/||X-\beta||^2)(X-\beta) + \beta & \text{if } ||X-\beta||^2 > c \\ \beta & \text{if } ||X-\beta||^2 \leq c \end{cases}$$

where c is a fixed nonnegative constant depending on the density f. The constant c that appears in the definition of $\hat{\theta}_c^+$ may be chosen so that $\hat{\theta}_c^+$ dominates the constant risk minimax estimator $\hat{\theta}(X) \equiv X$ under the squared error loss. (Recall that the estimator $\hat{\theta}_0(X) \equiv X$ is the generalized Bayes estimator for the uninformative prior on $\theta$.) The larger c is, the more values of X for which $\hat{\theta}_c^+(X) = \beta$, the likely candidate for $\theta$. Yet it is also desirable to choose c sufficiently small so that $\hat{\theta}_c^+$ still dominates $\hat{\theta}_0(X)$. Brandwein [3] has shown that as long as

$$c \leq \left(\frac{2(p-2)}{p}\right) \Big/ E_{\theta=0}[||X||^{-2}] \text{ ,}$$

$\hat{\theta}_c^+(X)$ dominates X for any spherically symmetric density f with $p \geq 4$. The theorems in the following subsection b enlarge that upper bound of Brandwein's for c under certain conditions on f while retaining the dominance of $\hat{\theta}_c^+$ over X. (For convenience $\beta$ is assumed to be the zero vector in the theorems in the next subsection b.) In particular the conditions which follow involve the function

$$q(R) = \int_R^{\infty} uf(u)du \ / \ f(R)$$

defined where $R \geq 0$ and $f(R) > 0$.

If $q(R)$ is nondecreasing then the upper bound for c is

$$2 \;/\; E_{\theta=0}[||X||^{-2}].$$

If $q(R)$ is nonincreasing, then the upper bound for c is

$$\left(\frac{2(p-2)}{p}\right) \; E_{\theta=0}[||X||^{2}] \quad .$$

(The monotonicity of the derivative of $\log f(R^{1/2})$ with respect to R implies the monotonicity of q (Lemma 4) and this may be easier to check.) These improvements in the upper bound for c may be contrasted with those developed by other authors. In the case that the density f is nonincreasing Brandwein and Strawderman [4] have given the upper bound $\left(\frac{2p}{(p+2)}\right) / E_{\theta=0}[||X||^{-2}]$ for c. When f is a mixture of normals then the upper bound for c has been shown to be $2 / E_{\theta=0}[||X||^{-2}]$ by Strawderman [6] and Berger [1]. (This is a special case of the situation where $q(R)$ is nondecreasing.) Note that $q(R)$ nondecreasing implies that f is nonincreasing. However for $q(R)$ nonincreasing there may be values of R for which $f(R)$ is nondecreasing. (If f is the normal density then q is a constant function and thus both nondecreasing and non-increasing. In that case the two upper bounds for c which depend on the monotonicity of q agree.) Berger [1] also gave upper bounds for c which depended on the function q. In the case that $\inf q(R)>0$ where inf is taken over those values of R such that $f(R)>0$, Berger's upper bound for c is $2(p-2) \inf q(R)$. The upper bounds for c given in the situation where q is monotone are larger than this bound.

## Subsection 2b: A general loss function

The positive part estimator given in subsection 2a is not admissible for squared error loss but is robust in the sense that it depends on the density f only through the constant c. For squared error loss, the general minimax estimators in Theorem 1 include admissible estimators and Theorem 2 gives a minimax positive part estimator under certain conditions on f. Theorem 3 generalizes these theorems to nondecreasing concave loss functions. Without loss of generality, these estimators shift towards the vector zero, rather than $\beta$.

Theorem 1. Define X to be a p-dimensional random vector with density $f(||X-\theta||)$, where the set of points in $[0,\infty]$ for which f is discontinuous has Lebesgue measure zero. Assume $E_{\theta=0}[||X||^{-2}]$ and $E_{\theta=0}[||X||^{2}]$ are finite. For $p \geq 4$, under squared error loss, the estimator

$$\hat{\theta}(X) = \left(1 - \frac{r(||X||^2)}{||X||^2}\right) X$$

dominates the estimator $\hat{\theta}_0(X)=X$ provided that r is a nondecreasing real-valued function such that $r(t)/t$ is nonincreasing in t and

(1) $0 < r(t) < 2/E_{\theta=0}[||X||^{-2}]$; and

(2) $q(t) = \{\int_t^\infty uf(u)du \,/f(t)\}$ is finite nondecreasing on $\{t \geq 0 : f(t)>0\}$.

Unless X is a normal random variable, the result holds for $r(t) \leq 2/E_{\theta=0}[||X||^{-2}]$.

Remark 1. The assumption (2) that q is nondecreasing in Theorem 1 implies that the density f is nonincreasing. The upper bound for r(t) in assumption (2) can be multiplied by p/(p+2) for general f nonincreasing where q is not necessarily nondecreasing according to Brandwein and Strawderman [4].

In Theorem 2 of this paper q is assumed to be nonincreasing. Some simple conditions which insure the monotonicity of q are given after Theorem 2. Also some examples of distributions satisfying such conditions are given after Theorem 3.

Proof of Theorem 1. Note r is differentiable almost everwhere. It suffices to show that $\Delta<0$ where

$$\Delta = E[||\hat{\theta}(X)-\theta||^2] - E[||X-\theta||^2].$$

Using integration by parts (see [1] Berger, P.1325),

$$\Delta = E[\phi^2(||X||^2)||X||^2] - 2E[q(||X-\theta||)\{p\phi(||X||^2 + 2||X||^2\phi'(||X||^2)\}]$$

Recall that $r(t)=t\phi(t)$ implies $t\phi'(t) = \frac{-r(t)}{t} + r'(t)$.

If $r'(t)\geq 0$ (since r is nondecreasing), then

$$\Delta \leq E\left[\frac{r^2(||X||^2)}{||X||^2}\right] - 2(p-2)E\left[q(||X-\theta||)\ \frac{r(||X||^2)}{||X||^2}\right] .$$

Because $E_{\theta=0}[||X||^{-2}]=E_\theta[||X-\theta||^{-2}]$, assumption (1) of the theorem implies that

$$\Delta < E\left[\frac{r(||X||^2)}{||X||^2}\left\{c - 2(p-2)q(||X-\theta||)\right\}\right]$$

where $c=2/E_\theta[||X-\theta||^{-2}]$.

This last bound for $\Delta$ may be written as

$$c\,E\,[G(R)] + \int_0^\infty 2(p-2)\{-G(R)\}\{q(R)\}R^{p-1}f(R)dR\;\alpha_p$$

where $G(R) = E\left[\frac{r(||X||^2)}{||X||^2}\;\middle|\;||X-\theta||=R\right]$, and $\alpha_p$ is the surface area of the p-dimensional unit sphere.

The integral in the bound given above for $\Delta$ may be written as

$$2(p-2)d^*\left[\int_0^\infty \{-R^2G(R)\}\{q(R)\}\frac{R^{p-3}f(R)\alpha_p}{d^*}dR\right]$$

where $d^* = E_\theta[||X-\theta||^{-2}]$. Note that $\alpha_p R^{p-3}f(R)/d^*$ is a density for R.

Because $\{-R^2G(R)\}$ is nonincreasing in R for $p\geq 4$ (see Brandwein [3]) and $\{q(R)\}$ is nondecreasing in R, we may apply Remark 1 and Lemma 1 of the Appendix to find an upper bound for the integral which is

$$2(p-2)d^*\left[\int_0^\infty\{-R^2G(R)\}\frac{R^{p-3}\alpha_p f(R)dR}{d^*}\right]\left[\int_0^\infty\frac{\{q(R)\;R^{p-3}\alpha_p f(R)dR}{d^*}\right].$$

This may be rewritten as

$$2(p-2)E[-G(R)]\left[\frac{1}{E_\theta[||X-\theta||^{-2}]}\;\left(\frac{1}{p-2}\right)\right]$$

because $\int_0^\infty q(R)f(R)R^{p-3}dR\alpha_p = \frac{1}{(p-2)}$ by Lemma 2 in the appendix.

This upper bound implies that $\Delta < cE[G(R)] - \frac{2E[G(R)]}{E_\theta[||X-\theta||^{-2}]}$, which is zero by definition of c.

q.e.d.

Theorem 2. Let X be a p-dimensional random vector with density $f(||X-\theta||)$, where the set of points in $[0,\infty)$ for which f is discontinuous has Lebesgue measure zero. For $p \geq 4$, under squared error loss the estimator

$$\hat{\theta}(X) = (1 - \frac{c}{||X||^2})X$$

dominates the estimator $\hat{\theta}(X) = X$ provided that

$$q(t) = \left\{\int_t^\infty uf(u)du \,/\, f(t)\right\}$$

is finite nonincreasing on $\{t \geq 0 : f(t) > 0\}$ and

$$c < \left(\frac{2(p-2)}{p}\right)E[||X-\theta||^2].$$

Unless X is a normal random vector, the result holds for

$$c = \left(\frac{2(p-2)}{p}\right)E[||X-\theta||^2].$$

In the case that X is normal $\hat{\theta}$ and $\hat{\theta}_0$ have the same risk for this value of c.

Corollary 1. Under the assumptions of the Theorem 2, the positive part estimator

$$\hat{\theta}_c^+(X) = \begin{cases} \left(1 - \frac{c}{||X||^2}\right) X & \text{if } ||X||^2 > c \\ 0 & \text{if } ||X||^2 \leq c \end{cases}$$

dominates the estimator $\hat{\theta}_0(X) = X$ for $c \leq \frac{2(p-2)}{p} E_{\theta=0}[||X||^2]$.

Proof of Theorem 2. The proof of Theorem 1 shows that (setting $r(t) \equiv c$)

$$\Delta = E[||\hat{\theta}(X)-\theta||^2] - E[||X-\theta||^2]$$

$$= c\, E[||X||^{-2} \{c-2(p-2)q(||X-\theta||)\}]$$

$$= c\, E[G(R) \{c-2(p-2)q(R)\}]$$

where $R = ||X-\theta||$ and $G(R) = E[||X||^{-2} \mid ||X-\theta|| = R]$.

Brandwein [3] has shown that $R^2G(R) = \psi\left(\frac{R}{||\theta||}\right)$ where $\psi$ is nondecreasing in the argument $\frac{R}{||\theta||}$. It is clear that $||\theta||^2G(R) = \psi\left(\frac{||\theta||}{R}\right)$ which is thus nonincreasing as a function of $\frac{R}{||\theta||}$. For fixed $||\theta||$, this shows that $G(R)$ is nonincreasing in $R$. The assumption that $q(R)$ is nonincreasing in $R$ implies that

$$E[G(R) \cdot q(R)] \geq E[G(R)] \cdot E[q(R)].$$

(Unless X is normal this inequality is actually strict.) Thus

$$\Delta \le cE[G(R)](c-2(p-2)E[q(R)]).$$

Because (by Lemma 2B of the Appendix) $E[q(R)] = \frac{E[||X-\theta||]}{p}$,

we have $\Delta \le 0$ for $c \le \frac{2(p-2)}{p} E[||X-\theta||^2]$.

This inequality is actually strict unless $q(R)\equiv 1$ (i.e. X is a normal random vector.) q.e.d.

Proof of Corollary 1.

The corollary follows directly from the theorem and the fact that the positive part estimator $\hat{\theta}_c^+$ dominates the estimator $\hat{\theta}$. q.e.d.

The following lemmas proved in the appendix give simple assumptions that insure the monotonicity of q. Note that $q'(t)<0$ if $f'(t)\ge 0$.

Lemma 3: On $\{t\ge 0:f(t)>0\}$, q(t) is nondecreasing if and only if on $\{t\ge 0:f(t)>0\}$ we have $f'(t)<0$ and $q(t)\ge tf(t)/(-f'(t))$. On $\{t\ge 0:f(t)>0\}$, q(t) is nonincreasing if and only if on $\{t\ge 0:f(t)>0\}$ we have

$$q(t)\le tf(t)/(-f'(t)) \text{ if } f'(t)<0.$$

Lemma 4: On $\{t\ge 0:f(t)>0\}$, q(t) is nondecreasing if $f'(t)<0$ and $\{f'(t)t^{-1}/f(t)\}$ is nondecreasing in t there; on $\{t\ge 0:f(t)>0\}$, q(t) is nonincreasing if $\{f'(t)t^{-1}/f(t)\}$ is nonincreasing in $\{t\ge 0:f(t)>0$ and $f'(t)<0\}$.

The following corollary employs simple assumption to insure that q is monotone.

Corollary 2. Let X be a p-dimensional random vector with density $f(||X-\theta||)$.

Let $g(t) = \log f(t^{1/2})$, and $q(t) = \int_t^\infty uf(u)du/\{f(t)\}$.

Then q is nondecreasing on $\{t\geq 0: f(t)>0\}$ if $f'(t)<0$ there and $g'(t)$ is nondecreasing there. Also, q is nonincreasing there if $g'(t)$ is nonincreasing on $\{t\geq 0: f'(t)<0$ and $f(t)>0\}$.

Proof. Lemma 4 of the appendix implies that if f is differentiable and nonincreasing, q is nondecreasing because $g'(t)$ nondecreasing is equivalent to the condition that $\{f'(t)t^{-1}/f(t)\}$ be nondecreasing: In the case that g is twice differentiable we have

$$\frac{d}{du}[g'(u)] = \frac{d^{(2)}}{du^2}[g(u)] = \frac{d^{(2)}}{du^2}[\log f(u^{1/2})]$$

$$= \left\{\frac{d}{dt}\left[\frac{f'(t)}{tf(t)}\right]_{t=u^{1/2}}\right\} u^{-1/2}/4.$$

A similar argument works for q nonincreasing. q.e.d.

Observe that the condition (2) that the function q be nondecreasing in Theorem 1 is satisfied for distributions which are mixtures of normal distributions. Thus the results of the theorem agree with those obtained by Strawderman [6] and Berger [1] for this class of distributions which includes the normal as well as the multivariate t-distribution.

Example. The following density is not a mixture of normals yet $q(t)$ is nondecreasing. Let $g(t) = t-(2+t)e^{-t}$. Define $f(t) = Kg'(t^2)\exp(-g(t^2))$.

Then $$q(t) = \int_t^\infty uf(u)du/f(t) = \frac{K}{2}\int_t^\infty 2ug'(u^2)\exp(-g(u^2))du/f(t)$$

$$= \frac{K}{2}\exp(-g(t^2))/f(t) = \tfrac{1}{2}/g'(t^2).$$

Then $q'(t) = -tg^{(2)}(t^2)/[g'(t)]^2$.

Note that $g'(u) = 1 + (1+u)e^{-u}$, and $g^{(2)}(u) = -ue^{-u}$, and $g^{(3)}(u) = (u-1)e^{-u}$, and $g^{(4)}(u) = (2-u)e^{-u}$. Thus $q'(t)>0$ and $q(t)$ is nondecreasing.

By Theorem 2 of Berger [1], f is a mixture of normals if and only if $\rho(u)$ is completely monotonic for u in $(0,\infty)$ where $\rho(u) = g'(u)\exp(-g(u))$ i.e.,

$$\rho^{(j)}(u)(-1)^j \geq 0 \text{ for all } j \geq 0.$$

But
$$\rho^{(3)}(u) = e^{-g(u)}\{g^{(4)}(u) - 3[g^{(2)}(u)]^2 + 6g^{(2)}(u)[g'(u)]^2$$
$$- 4g'(u)g^{(3)}(u) - [g'(u)]^4\}.$$

Setting u=1, we see $(-1)^3\rho^{(3)}(u)<0$. Thus $\rho$ is not completely monotonic and f is not a mixture of normals. q.e.d.

In the case that q is not nondecreasing, but f is nonincreasing, results of Brandwein and Strawderman [4] insure that the $\hat{\theta}$ in Theorem 1 of this section is minimax for $p\geq 4$ when (2) is replaced by "(2') f is nonincreasing" and (1) is replaced by

$$\text{"(1') } 0 \leq r(t) \leq \{p/(p+2)\}2/E_{\theta=0}[||X||^{-2}]."$$

(For p=3, the ratio $p/(p+2)$ is replaced by 3/8.)

If f is not necessarily nonincreasing, Brandwein [3] has shown that $\hat{\theta}$ is minimax for $p\geq 4$ when (2) is deleted and (1) is replaced by

$$\text{"(1") } 0 \leq r(t) \leq \{(p-2)/p\}2/E_{\theta=0}[||X||^{-2}]."$$

Note that if f is differentiable and there is a value $t_0$ such that $f'(t_0)\geq 0$, then $q'(t_0)<0$ (using definition of q) so that q will not be nondecreasing at $t_0$.

The following theorem extends the result of Theorems 1 and Corollary 1 as well as the result of Brandwein and Strawderman [4] for density f to a general nondecreasing concave loss function. It is an extension of a similar result by Brandwein and Strawderman [5] for all spherically symmetric distributions.

Theorem 3. Let X have a p-dimensional spherically symmetric distribution about $\theta$ with density $f(||X-\theta||)$. Define the nondecreasing concave loss function $\ell$ by $L(\theta,\hat{\theta}) = \ell(||\hat{\theta}-\theta||^2)$. Then the estimator is better than X and minimax where

$$\delta(X) = \left[1 - \frac{r(||X||^2)}{||X||^2}\right] X$$

provided

(i) $r(t)$ is nondecreasing in t;

(ii) $r(t)/t$ is nonincreasing in t;

(iii) $0<r(t)<c^*2E_{\theta=0}[\ell'(||X||^2)]/E_{\theta=0}[||X||^{-2}\ell'(||X||^2)]$

in the following three cases:

Case A: If the density f(t) is nonincreasing in t, set $c^*=p/(p+2)$ for $p\geq 4$ and $c^*=3/8$ for $p=3$.

Case B: If $\ell'(t^2)f(t)$ has a negative derivative and $Q(t)$ is nondecreasing on $\{t\geq 0: \ell'(t^2)f(t)>0\}$, set $c^* = 1$ for $p \geq 4$ where

$$Q(t) = \int_t^\infty u\ell'(u^2)f(u)du/[\ell'(t^2)f(t)].$$

Case C. If the function $Q(t)$ is nonincreasing on $\{t\geq 0: \ell'(t^2)f(t)$ is positive with a negative derivative$\}$, set $r(t) = t$ for $t \leq c$ and set $r(t) = c$ for $t>c$ where $c < \frac{2(p-2)}{p} E_{\theta=0}[||X||^2\ell'(||X||^2)] / E_{\theta=0}[\ell'(||X||^2)]$.

Remarks. (1) For $p\geq 4$ and $c^*=(p-2)/p$, the result of the theorem was obtained by Brandwein and Strawderman [5] with no restrictions on f.

(2) If the strict inequalities in (iii) are relaxed to "$\leq$", then $\delta$ is minimax and at least as good as X, but not necessarily better than X; however if Q is strictly increasing on a Lebesgue set of positive measure where $\ell'(t^2)f(t)$ is positive, then $\delta$ is better than X.

(3) Note that the proof of Corollary 2 shows that $Q(t)$ is nondecreasing if $\log(\ell'(t)f(t^{1/2}))$ is nonincreasing in t with a nondecreasing derivative in t.

(4) The condition (ii) of Theorem 3 that $\{r(t)/t\}$ be nondecreasing in t may be replaced by the weaker condition

(ii) $h(R) = R^2E[r(||X||^2)||X||^{-2} \mid ||X-\theta|| = R]$ is nondecreasing in R.

Examples

In each case assume the loss $L(\theta,\hat{\theta}) = ||\hat{\theta}-\theta||^b$ where $0<b\leq 2$ and $p\geq 4$.

(1) If X has the p-dimensional uniform distribution on a sphere, i.e., $||X-\theta||^2\leq S^2$, then case A applies, but not case B. So $\delta(X)$ is better than X if

$$0 < r(\cdot) \leq \frac{2p}{(p+2)} E_{\theta=0}[||X||^{b-2}]/E_{\theta=0}[||X||^{b-4}] = \frac{2p}{(p+2)} S^2\frac{(p+b-4)}{(p+b-2)}$$

(2) Let X have a spherically symmetric distribution about $\theta$ which is a mixture of normals. Assume that p+b>4. Then Case B applies since log $(\frac{b}{2} t^{\frac{b}{2}-1} f(t^{1/2}))$ is nonincreasing with a nondecreasing derivative. Mixtures of normals include the multivariate t and normal distributions. For the normal distribution, $\delta$ is minimax if

$$0 \leq r(\cdot) \leq 2E_{\theta=0}[||X||^{b-2}]/E_{\theta=0}[||X||^{b-4}] = 2(b+p-4).$$

(3) Let X have a density $f(||X-\theta||)$ of a form considered by Berger [1] where

$$f(t) = kt^{2n} \exp(-t^2/2),$$

for $n \geq 0$. Assume b/2 + n>1. Then Case C applies since log $(\frac{b}{2} t^{\frac{b}{2}-1} f(t^{1/2}))$ has a nonincreasing derivative. Then the positive part estimator $\delta$ is minimax where $r(t) = t$ for $t \leq c^*$ and $r(t) = c^*$ for $t>c^*$ provided

$$c^* = \frac{2(p-2)}{p} E_{\theta=0}[||X||^b] / E_{\theta=0}[||X||^{b-2}] = \frac{2(p-2)}{p} (b+2n+p-2).$$

Proof of Theorem 2.

It follows immediately from Theorem 1 and Corollary 2 of this paper and from Theorem 2.1 of Brandwein and Strawderman [5] and Theorem 3.3.1 of Brandwein and Strawderman [4]. q.e.d.

## Section 3. Estimators that shift towards a hypersphere

In the situation where $\theta$ is deemed "likely" to lie in a certain ball or sphere, $K_{\beta,G}$, of radius G centered at the vector $\beta$, one has the opportunity to make use of this "vague" information when the minimax invariant estimator X does not happen to fall in the sphere $K_{\beta,G}$. The estimators given in this section may be defined so they are minimax; they may also be chosen so that the estimate of $\theta$ (for values of X outside $K_{\beta,G}$) is shifted in a direction from X towards P(X), the closest vector on the sphere to X. Note that these estimators avoid the problem that occurs when $\theta$ is deemed likely to lie in a closed convex polyhedron and the estimate of $\theta$ is to be shifted from X towards the closest vector on the polyhedron (for values of X outside the polyhedron). For the polyhedron situation the estimators considered in [2] were equal to X if the closest vector on the polyhedron to X lay on a very high-dimensional face, in which case no "shifting" took place.

The notation $K^c_{\beta,G}$ denotes the complement of $K_{\beta,G}$ in the following theorem which gives a class of estimators containing estimators which shift to the ball $K_{\beta,G}$.

<u>Theorem 4</u>. Let $K_{\beta,G}$ be the ball of radius G in $\mathbb{R}^p$ centered at the vector $\beta$, i.e.,

$$K_{\beta,G} = \{Y \text{ in } \mathbb{R}^p: \ ||Y-\beta|| \leq G\}.$$

Let X and $\theta$ be p-dimensional vectors in $\mathbb{R}^p$ and assume that X is a random vector with density $f(||X-\theta||)$. Assume $p \geq 5$ and define the nondecreasing concave loss function $\ell$ by

$$L(\theta,\hat{\theta}) = \ell(||\theta-\hat{\theta}||^2).$$

Define P(X) to be the closest vector to X on the surface of $K_{\beta,G}$. Let $\delta_0$ be an estimator of $\theta$ which equals X if X is not in $K_{\beta,G}$. Define the estimator $\delta(X)$ equal to $\delta_0(X)$ if X is not in $K_{\beta,G}$; if X is in $K_{\beta,G}$, let $\delta(X)=X-r(||X-P(X)||^2)(X-P(X))/\{||X-P(X)||(||X-P(X)||+G)\}$

where

(a) $r$ is a real-valued nondecreasing function such that $\{r([t^{1/2}-G]^2)/t\}$ is nonincreasing in $t$ for $t > G^2$;

(b) $0<r(t)<2c^*E_{\theta=0}[\ell'||X||^2)]/E_{\theta=0}[||X||^{-2}\ell'(||X||^2)]$.

Then $\delta$ dominates $\delta_0$ for the following cases:

Case 0: With no restriction on $f$, set $c^*=(p-2)/p$.

Case A: If the density $f(t)$ is nondecreasing in $t$, set $c=(p/(p+2)$.

Case B: Define the function $Q(t)$ to be

$$Q(t) = \int_t^\infty u\ell'(u^2)f(u)\,du/\ell'[(t^2)f(t)].$$

If $\ell'(t^2)f(t)$ has a negative derivative and $Q(t)$ is nondecreasing on $\{t \geq 0: \ell'(t^2)f(t) \gg 0\}$, set $c^* = 1$ for $p \geq 4$.

The estimators $\delta$ given in the theorem are minimax if the estimator $\delta_0$ is minimax. The values that $\delta_0$ takes when $X$ falls in $K_{\beta,G}$ were not specified and $\delta$ was defined to agree with $\delta_0$ for those values of $X$.

Remark. If the strict inequalities in (b) of Theorem 4 are relaxed to "$\leq$", then $\delta$ is as good as $\delta_0$ and minimax if $\delta_0$ is minimax.

Example: The following estimator is a simple shrinkage estimator to $K_{\beta,G}$ which belongs to the class of estimators given in Theorem 4. It is minimax and better than $X$.

Define $$c = 2c^*E_{\theta=0}[\ell'(||X||^2)]/E_{\theta=0}[||X||^{-2}\ell'(||X||^2)].$$

Then let $$\delta(X) = P(X) + \left[1 - \frac{c}{||X-P(X)||(||X-P(X)||+G)}\right]^+(X-PX)).$$

(The "+" indicates that the quantity within the square brackets should be replaced by zero if it is negative.)

For this estimator, a value of X outside the ball $K_{\beta,G}$ that satisfies

$$G < ||X-\beta|| \leq \frac{G}{2} + \sqrt{(\frac{G}{2})^2+c}$$

is shrunk to P(X), the point closest to X on the surface of the ball. Values of X further away from the ball are also shifted closer to the ball, but not onto the surface of the ball. Clearly, if c* and thus c is larger, then more values are shrunk all the way to the ball. Consider the case for squared error loss. In the case that p=5, knowing that the density $f(||X-\theta||)$ is nonincreasing allows one to set c*-5/7 rather than c*=3/5. Knowing that $q(t)=\int_t^\infty uf(u)du/f(t)$ is nondecreasing allows one to choose c*=1.

Proof of Theorem 4: Observe that $I_{K^c_{\beta,G}}(X) = I_{(G,\infty)}(||X-\beta||)$ and the projection of X to the ball is $P(X) = I_{K_{\beta,G}}(X)X + I_{K^c_{\beta,G}}(X)(\beta+(X-\beta)G/||X-\beta||)$.

Since $I_{K^c_{\beta,G}}(X)(X-P(X)) = I_{(G,\infty)}(||X-\beta||)(X-\beta)(1-G/||X-\beta||)$

and $I_{K^c_{\beta,G}}(X)||X-P(X)|| = I_{(G,\infty)}(||X-\beta||)(||X-\beta||-G)$, we have

$$I_{K^c_{\beta,G}}(X)\delta(X) = I_{K^c_{\beta,G}}(X)[X-r([||X-\beta||-G]^2)(X-\beta)/||X-\beta||^2].$$

Thus

$$\Delta = E[||\delta(X)-\theta||^2] - E[||\delta_0(X)-\theta||^2]$$

$$= E[I_{K^c_{\beta,G}}(X)\{||\beta+(1-r([||X-\beta||-G]^2)/||X-\beta||^2)(X-\beta)-\theta||^2-||X-\theta||^2\}].$$

Now define $\delta_0^*(X)=X$ and $r^*(||X-\beta||^2) = I_{K^c_{\beta,G}}(X)r([||X-\beta||-G]^2)$, and

$$\delta^*(X) = \beta + (1-r^*(||X-\beta||^2)/||X-\beta||^2)(X-\beta)$$

$$= X - r^*(||X-\beta||^2)(X-\beta)/||X-\beta||^2.$$

Then $I_{K^c_{\beta,G}}(X)\delta(X) = I_{K^c_{\beta,G}}(X)\delta^*(X)$ and $I_{K^c_{\beta,G}}(X)\delta_0(X) = I_{K^c_{\beta,G}}(X)\delta^*_0(X)$.

Thus the difference in risks between $\delta$ and $\delta_0$ is equal to the difference in risks between $\delta^*$ and and $\delta^*_0$, i.e.,

$$E[\ell(||\delta(X)-\theta||^2)] - E[\ell(||\delta_0(X) - \theta||^2)]$$

$$= E[I_{K^c_{\beta,G}}(X)\{\ell(||\delta(X) - \theta||^2) - \ell(||\delta_0(X) - \theta||^2)\}]$$

$$= E[I_{K^c_{\beta,G}}(X)\{\ell(||\delta^*(X) - \theta||^2) - \ell(||\delta^*_0(X) - \theta||^2)\}]$$

$$= E[\ell(||\delta^*(X) - \theta||^2)] - E[\ell(||\delta^*_0(X) - \theta||^2)].$$

Therefore, it suffices to show that $\delta^*$ dominates $\delta^*_0$.

The conditions a) and b) of Theorem 4 imply that $r^*$ is a real-valued nondecreasing function such that $r^*(t)/t$ is nonincreasing for $t > G^2$ and

$$0 < r^*(t) < 2c^* E_{\theta=0}[\ell'(||X||^2)]/E_{\theta=0}[||X||^{-2}\ell'(||X||^2)].$$

However, $\{r^*(t)/t\} = 0$ for $t \le G^2$; it is not true that $\{r^*(t)/t\}$ is nonincreasing for all $t > 0$. The condition that $\{r^*(t)/t\}$ is nonincreasing is used only to show that $h(R)$ is a nondecreasing function of $R$ where

$$h(R) = R^2 E[r^*(||X||^2)||X||^{-2} \big| ||X-\theta|| = R],$$

by Brandwein [3], Brandwein and Strawderman [4], [5] and Theorem 3 of this paper. Lemma 6 of the Appendix shows that $h(R)$ is nondecreasing in $R$ for $p \geq 5$.

Thus, for case 0, the dominance of $\delta^*$ over $\delta_0^*$ is shown by Brandwein and Strawderman [5] and Lemma 6 of the Appendix. For cases A and B, the dominance of $\delta^*$ over $\delta_0^*$ is given by Theorem 3 of this paper and Lemma 6 of the Appendix.

q.e.d.

## Appendix

Lemma 1. Let T be a non-negative random variable. Assume $f_1$ and $f_2$ are real-valued functions defined on $[0,\infty]$ such that $\mu_i=E(f_i(T))$, i=1,2, are finite. Let $f_2$ be nonincreasing and let there be a value $f_1^{-1}(\mu_1)$ such that

$$f_1(t) \leq \mu_1 \text{ for } t \leq f_1^{-1}(\mu_1) \text{ and}$$

$$f_1(t) \geq \mu_1 \text{ for } t \geq f_1^{-1}(\mu_1)$$

Then $E[f_1(T)\cdot f_2(T)] \leq E[f_1(T)]\cdot E[f_2(T)]$.

Proof. It suffices to show that $(*)\leq 0$ where $(*)=E[(f_1(T)-\mu_1)f_2(T)]$. Without loss of generality, assume that $f_1^{-1}(\mu_1)\geq 0$. Let G be the distribution function of T. Then

$$(*) = \int_0^\infty (f_1(t)-\mu_1)f_2(t)dG(t)$$

$$= \int_0^{f_1^{-1}(\mu_1)} (f_1(t)-\mu_1)f_2(t)dG(t) + \int_{f_1^{-1}(\mu_1)}^\infty (f_1(t)-\mu_1)f_2(t)dG(t)$$

For $0\leq t\leq f_1^{-1}(\mu_1)$, $f_1(t) \leq \mu_1$ and $f_2(t) \geq f_2(f_1^{-1}(\mu_1))$ since $f_2$ is nonincreasing. This implies that for $0\leq t\leq f_1^{-1}(\mu_1)$, we have $(f_1(t)-\mu_1)$ nonpositive and

$$(f_1(t)-\mu_1)f_2(t)\leq(f_1(t)-\mu_1)f_2(f_1^{-1}(\mu_1))$$

and the first integral in the latest representation for (*) is bounded above by

$$A = \int_0^{f_1^{-1}(\mu_1)} (f_1(t)-\mu_1)f_2(f_1^{-1}(\mu_1))dG(t).$$

For $t\geq f_1^{-1}(\mu_1)$, we have $(f_1(t)-\mu_1)\geq 0$ and $f_2(t)\leq f_2(f_1^{-1}(\mu_1))$ since $f_2$ is nonincreasing. This implies that for $t \geq f_1^{-1}(\mu_1)$,

$$(f_1(t)-\mu_1)f_2(t) \leq (f_1(t)-\mu_1)f_2(f_1^{-1}(\mu_1)).$$

So the second integral in the last representation for (*) is bounded above by

$$B = \int_{f_1^{-1}(\mu_1)}^{\infty} (f_1(t)-\mu_1)f_2(f_1^{-1}(\mu_1))dG(t).$$

The sum of A and B gives an upper bound for (*) which is zero.

q.e.d.

Remark 1. If $f_1$ is nondecreasing it satisfies the following conditions of Lemma 1: there exists a value $f_1^{-1}(\mu_1)$ such that

$$f_1(t) \leq \mu_1 \text{ for } t \leq f_1^{-1}(\mu_1)$$

and $\quad f_1(t) \geq \mu_1$ for $t \geq f_1^{-1}(\mu_1)$.

Remark 2. If there is an interval (where the density of T is positive) such that $f_1$ is strictly increasing and $f_2$ is strictly decreasing, the conclusion of Lemma 1 may be strengthened to a strict inequality.

Lemma 2A Let X be a p-dimensional random vector with density $f(||X-\theta||)$. Define on $\{R \geq 0: f(R)>0\}$ the function

$$q(R) = \int_R^{\infty} uf(u)du/f(R)$$

and assume that $q(R) < \infty$. Then $E[q(||X-\theta||)||X-\theta||^{-2}] = \frac{1}{(p-2)}$.

Proof: Let $\alpha_p$ be the surface area of the p-dimensional unit sphere.

Then

$$E[q(||X-\theta||)||X-\theta||^{-2}] = \int_0^\infty \{R^{-2}q(R)\}f(R)R^{p-1}dR\alpha_p.$$

Because $q(R)f(R)=\int_0^\infty uf(u)du$, a change in the order of integration implies

$\int_0^\infty R^{p-3}q(R)f(R)dR\alpha_p = \int_0^\infty \frac{u^{p-2}}{(p-2)} uf(u)du\alpha_p$. Since $\int_0^\infty u^{p-1}f(u)du\alpha_p=1$, the last expression equals $1/(p-2)$. q.e.d.

<u>Lemma 2B</u> $E[q(||X-\theta||)] = E[||X-\theta||^2]/p$.

<u>Proof</u>:

Note $E[q(||X-\theta||)] = \int_0^\infty q(R)R^{p-1}f(R)dR\alpha_p$. Because $q(R)f(R) = \int_R^\infty uf(u)du$, we have (by a change in the order of integration) that

$$E[q(||X-\theta||)] = \int_0^\infty \frac{u^p}{p} uf(u)du\alpha_p$$

$$= \frac{E[||X-\theta||^{+2}]}{p}.$$

q.e.d.

<u>Lemma 3</u>. Let f be a differentiable function such that q(t) is finite where for $\{t\geq 0:f(t)>0\}$, we have $q(t) = \int_t^\infty uf(u)du/f(t)$.

Then q is nondecreasing on $\{t\geq 0:f(t)>0\}$ if and only if on $\{t\geq 0:f(t)>0\}$, we have $q(t) \geq tf(t)/(-f'(t))$ and $f'(t)<0$.

Also, q is nonincreasing on $\{t\geq 0:f(t)>0\}$ if and only if on $\{t\geq 0:f(t)>0\}$ we have

$q(t) \leq tf(t)/(-f'(t))$, when $f'(t)<0$.

Proof. The derivative of q is

$$q'(t) = -t - f'(t)q(t)/f(t) = (-f'(t)/f(t))(q(t) - tf(t)/(-f'(t))).$$

Note that $q'(t)<0$ if $f'(t)\geq 0$. For $f'(t)<0$, we have $(-f'(t)/f(t))\geq 0$, and the statements of the lemma follow.

q.e.d.

Lemma 4. Let f,q be given as in Lemma 3. Then

(a) q(t) is nondecreasing on $\{t:f(t)>0\}$ if on $\{t:f(t)>0\}$ the function $f'(t)t^{-1}/f(t)$ is nondecreasing in t and $f'(t)<0$;

(b) q(t) is nonincreasing on $\{t:f(t)>0\}$ if on $\{t:f(t)>0\}$ the function $f'(t)t^{-1}/f(t)$ is nonincreasing for t in $\{t\geq f(t)>0$ and $f'(t)<0\}$.

Proof. Note that

$$\begin{aligned} f(t) &= \int_t^\infty [-f'(s)]ds = \int_t^\infty (-f'(s)s^{-1}/f(s))sf(s)ds \\ &\leq (-f'(t)t^{-1}/f(t)) \int_t^\infty sf(s)ds \end{aligned}$$

if $(-f'(t)t^{-1}/f(t))$ is nonincreasing. Multiplying both sides of the above inequality by $t/(-f'(t))$ implies $tf(t)/(-f'(t))\leq q(t)$, using the definition of q. Lemma 3 implies that this last inequality insures that q is nondecreasing. A similar argument gives (b).

q.e.d.

Lemma 5. Assume $p \geq 5$. Let $r(t)$ be a bounded nonnegative function for $t \geq 0$ with $r'(t) \geq 0$ and $\frac{d}{dt}\{\frac{r(t)}{t}\} \leq 0$, for $t \geq 0$. For $-(R/||\theta||) \geq u \geq c_R$, and $r>0$, $M(u)$ is a nondecreasing function of u where

$$M(u) = \{r(||X||^2)||X||^{-2}(1-u^2)^{(p-3)/2}\}$$

and

$$||X||^2 = ||\theta||^2 + R^2 + 2R||\theta||u,$$

and $c_R = (2R||\theta||)^{-1}(G^2 - R^2 - ||\theta||^2) > -1.$

Proof. It suffices to show that $\frac{d}{du}(M(u)) \geq 0$. Because $r'(||X||^2) \geq 0$,

$$\frac{d}{du}(M(u)) = (r'(||X||^2)||X||^{-2} - r(||X||^2)||X||^{-4})2R||\theta||$$

$$\cdot (1-u^2)^{(p-3)/2} + r(||X||^2)||X||^{-2}(p-3)(1-u^2)^{(p-5)/2}(-u)$$

$$\geq (1-u^2)^{(p-5)/2} r(||X||^2)||X||^{-4}((-u)||X||^2(p-3) - 2R||\theta||(1-u^2)).$$

Since $(-u) \geq (R/||\theta||)$ and $(p-3) \geq 2$,

$$\frac{d}{du}(M(u)) \geq (1-u^2)^{(p-5)/2} r(||X||^2)||X||^{-4}(2R||\theta||^{-1}||X||^2 - 2R||\theta||(1-u^2))$$

$$= 2R||\theta||^{-1} r(||X||^2)||X||^{-4}(R + ||\theta||u)^2 \geq 0.$$

q.e.d.

Lemma 6. Assume that the p-dimensional random vector X has a spherically symmetric distribution about the vector $\theta$. Let $r(t)$ be a nonnegative function of t such that $r(t)$ is nondecreasing in t and $\{r(t)/t\}$ is nonincreasing in t for $t > 0$. If $p \geq 5$, then $h(R)$ is nondecreasing in R where

$$h(R) = R^2 E[I_{(G^2,\infty)}(||X||^2) r(||X||^2) ||X||^{-2} \,\big|\, ||X-\theta|| = R].$$

Proof. Note that for fixed $\theta$ and R $(=||X-\theta||)$, the distribution of $||X||^2$ may be chosen to be that of

$$R^2 + 2R||\theta||u + ||\theta||^2$$

where u is a random variable with density $I_{[-1,+1]}(u)(1-u^2)^{(p-3)/2}M^*$. (The $M^*$ is a normalizing constant.) We will write

$$||X||^2 = R^2 + 2R||\theta||u + ||\theta||^2$$

in the sense that their distributions are alike for fixed R and $\theta$.

Define $c_R = (G^2 - R^2 - ||\theta||^2)/(2R\,||\theta||)$ for $R > 0$ and $||\theta|| > 0$. It is clear that $||X|| > G$ corresponds to $u > c_R$.

Case 1: Assume R satisfies $G^2 > (R + ||\theta||)^2$. Then $c_R > 1$. Since $u \leq 1$, we have

$$I_{(G^2,\infty)}(||X||^2) = I_{(c_R,\infty)}(u) = 0$$

and $h(R) = 0$. Thus $h(R)$ is trivially nondecreasing for these values of R.

Case 2: Assume R satisfies $G^2 < (R - ||\theta||)^2$. Then $c_R < -1$. Since $u \geq -1$, we have

$$I_{(G^2,\infty)}(||X||^2) = I_{(c_R,\infty)}(u) = 1$$

and

$$h(R) = R^2 E[r(||X||^2)||X||^{-2} \big| ||X-\theta|| = R].$$

This has been shown to be nondecreasing in R by Brandwein [3].

Case 3: Assume R satisfies $(R-||\theta||)^2 \leq G^2 \leq (R + ||\theta||)^2$. We may write

$$h(R) = R^2 \int_{c_R}^{1} r(||X||^2)||X||^{-2}(1-u^2)^{(p-3)/2} M^* du.$$

Without loss of generality, assume that r is differentiable. Then

$$\frac{d}{dR}(h(R)) = A + B + C$$

where

$$A = [G^2 + R^2 - ||\theta||^2] r(G^2)(2||\theta||G^2)^{-1}(1-c_R^2)^{(p-3)/2} M^*$$

and

$$B = \int_{c_R}^{1} R^2 r'(||X||^2)||X||^{-2}(2R + 2||\theta||u)(1-u^2)^{(p-3)/2} M^* du$$

and

$$C = \int_{c_R}^{1} r(||X||^2)||X||^{-4}(2Ru + 2||\theta||)R||\theta||(1-u^2)^{(p-3)/2}M^*du.$$

Subcase 1: Assume R satisfies $R^2 + ||\theta||^2 \leq G^2$. Then $c_R \geq 0$ and A, B and C are nonnegative which implies that $\frac{d}{dR}(h(R)) \geq 0$.

Subcase 2: Assume R satisfies $G^2 < R^2 + ||\theta||^2$. Then $c_R < 0$ since $c_R = (G^2 - R^2 - ||\theta||^2)/(2||\theta||R)$.

Subcase a: Assume R satisfies $G^2 > ||\theta||^2 - R^2$. This implies that

$$Ru + ||\theta|| > Rc_R + ||\theta|| = (G^2 - R^2 + ||\theta||^2)/(2||\theta||).$$

If $Rc_R + ||\theta|| \geq 0$, then $Ru + ||\theta|| \geq 0$ and $c \geq 0$. If $Rc_R + ||\theta|| < 0$, then

$$C \geq \int_{-1}^{+1} r(||X||^2)||X||^{-2}R||\theta||(2Ru + 2||\theta||)(1-u^2)^{(p-3)/2}M^*du$$

since $(2Ru + 2||\theta||) < 0$ for $-1 \leq u \leq c_R$. This lower bound for C was shown to be nonnegative by Brandwein [3]. Because $R + ||\theta||u > R + ||\theta||c_R = (G^2 + R^2 - ||\theta||^2)/(2R)$, we have $A \geq 0$ and $B \geq 0$. Thus

$$\frac{d}{dR}(h(R)) = A + B + C \geq 0.$$

Subcase b: Assume R satisfies $G^2 \leq ||\theta||^2 - R^2$.

Since $||\theta|| > R$ and since $(R + ||\theta||u) > 0$ for $u > -(R/||\theta||)$ and since $r'(t) > 0$, we have

$$B \geq 2R^2 \int_{c_R}^{-(R/||\theta||)} r'(||X||^2)||X||^{-2}(R + ||\theta||u)(1-u^2)^{(p-3)/2}M^*du.$$

Because $\{r(t)/t\}$ nonincreasing implies $r'(t)/t \leq r(t)/t^2$ and because $(R + ||\theta||u) \leq 0$ for $c_R \leq u \leq -(R/||\theta||)$, we have

$$B \geq 2R^2 \int_{c_R}^{-(R/||\theta||)} r(||X||^2)||X||^{-4}(R + ||\theta||u)(1-u^2)^{(p-3)/2}M^*du.$$

Thus

$$C + B \geq 2R \int_{c_R}^{-(R/||\theta||)} r(||X||^2)||X||^{-4}[(R||\theta||u + ||\theta||^2) + (R^2 + R||\theta||u)].$$

$$\cdot\ (1-u^2)^{(p-3)/2}M^*du$$

$$= 2R \int_{c_R}^{-(R/||\theta||)} r(||X||^2)||X||^{-2}(1-u^2)^{(p-3)/2}M^*du.$$

Because $r(||X||^2)||X||^{-2}(1-u^2)^{(p-3)/2}$ is nondecreasing in $u$ for $p \geq 5$ by Lemma 5 of the Appendix if $-(R/||\theta||) \geq u \geq c_R$, we have

$$C + B \geq [2R[r(||X||^2)||X||^{-2}(1-u^2)^{(p-3)/2}M^*]_{u=c_R} \ [\int_{c_R}^{-(R/||\theta||)} du]$$

$$= 2Rr(G^2)G^{-2}(1-c_R^2)^{(p-3)/2}M^*[-(R/||\theta||) - c_R] = -2A.$$

Thus $A + B + C \geq -A \geq 0$. q.e.d.

# References


[1] Berger, James (1975) "Minimax estimation of location vectors for a wide class of densities". The Annals of Statistics. Vol.3, 1318-1328.

[2] Bock, M.E. (1981) "Employing Vague Inequality Information in the Estimation of Normal Mean Vectors (Estimators that shrink to closed convex polyhedra)" Purdue University Statistics Department Technical Report #81-42.

[3] Brandwein, Ann Cohen (1979) "Minimax estimation of the mean of spherically symmetric distributions under general quadratic loss". Journal of Multivariate Analysis Vol.9, 579-588.

[4] Brandwein, Ann Cohen and Strawderman, William E. (1978) "Minimax estimation of location parameters for spherically symmetric unimodal distributions." The Annals of Statistics 6, 377-416.

[5] Brandwein, Ann Cohen and Strawderman, William E. (1980) "Minimax estimation of location parameters for spherically symmetric distributions with concave loss". The Annals of Statistics Vol.8, 279-284.

[6] Strawderman, W.E. (1974). Minimax estimation of location parameters for certain spherically symmetric distributions. Journal of Multivariate Analysis Vol.4, 255-264.