The Annals of Statistics
1997, Vol. 25, No. 5, 2259–2272

SIMULTANEOUS ESTIMATION OF LOCATION PARAMETERS FOR SIGN-INVARIANT DISTRIBUTIONS

BY JIAN-LUN XU

University of Houston

Estimation of p, $p \geq 3$, location parameters of a distribution of a p-dimensional random vector $\mathbf{X}$ is considered under quadratic loss. Explicit estimators which are better than the best invariant one are given for a sign-invariantly distributed random vector $\mathbf{X}$. The results depend only on the second and the third moments of $||\mathbf{X} - \boldsymbol{\theta}||$. The generalizations to concave loss functions and to n observations are also considered. Additionally, if the scale is unknown, we investigate the estimators of the location parameters when the observation contains a residual vector.

1. Introduction. Since Stein (1956) first demonstrated the inadmissibility of the best invariant estimator of the p-dimensional ($p \geq 3$) normal mean under quadratic loss, there has been considerable interest in improving upon the best invariant estimator of a location vector. The ensuing development can be roughly classified in three major directions: considering different loss functions; considering more general estimates; relaxing the normality assumption.

Brown (1966) proved that the best invariant estimator of a location vector is inadmissible for a wide class of distributions and loss functions if the dimension is at least 3. James and Stein (1961) presented an explicit estimator $\{1-(p-2)/||\mathbf{X}||^2\}\mathbf{X}$ which is better than the best invariant estimator $\mathbf{X}$ under quadratic loss if $\mathbf{X}$ has a normal distribution with identity covariance matrix $\mathscr{I}$. This result remains true if the distribution of $\mathbf{X}$ is spherically symmetric and $p \geq 4$ as shown by Brandwein (1979), Brandwein and Strawderman (1978, 1980, 1991) and others; see the review article by Brandwein and Strawderman (1990). In another direction, James and Stein (1961) showed that the assumption of normality is unnecessary and suggested an estimator of the form

$$\boldsymbol{\delta}_{a,b}(\mathbf{X}) = \left\{1 - \frac{b}{a + ||\mathbf{X}||^2}\right\}\mathbf{X}, \tag{1.1}$$

which is better than $\mathbf{X}$ if a and b are suitably chosen. However, no explicit values for a and b were given. The determination of a and b was later studied by Shinozaki (1984), who investigated the ranges of a and b under the assumption that the components of $\mathbf{X}$ are independent, identically and symmetrically (iis) distributed about their respective means. Shinozaki's bounds

Received March 1994; revised January 1997.
AMS 1991 *subject classifications*. Primary 62C99; secondary 62F10, 62H12.
Key words and phrases. Simultaneous estimation, location parameters, sign-invariant distributions, quadratic loss, risk function.

for a and b involve the second and fourth moments of the component distributions. Shinozaki also used integration-by-parts to discuss the ranges of a and b by assuming that $\mathbf{X}$ has the uniform distribution, the t-distribution or the double exponential distribution.

In the direction of more general estimators, Miceli and Strawderman (1986) considered an estimator that is more general than (1.1), that is, they replace b by a function $br(X_1^2, \ldots, X_p^2)$. However, they restricted the distribution of $\mathbf{X}$ to the subclass of iis distributions called *independent component variance mixtures of normals*. Their loss function was nonquadratic. When $\mathbf{X}$ has a spherically symmetric distribution about its mean, Brandwein and Strawderman (1991) used the *divergence* theorem to obtain some beautiful results about the estimator of the form $\boldsymbol{\delta}_b(\mathbf{X}) = \mathbf{X} + b\mathbf{g}(\mathbf{X})$. Their loss function was either quadratic, or a concave function of quadratic loss or the general quadratic loss.

In this paper, we relax the distributional assumption. We assume that $\mathbf{X}$ has a sign-invariant distribution whose precise meaning is given by Definition 1 in Section 2. We obtain the ranges of a and b for the estimator given by (1.1) under the quadratic loss function or a nondecreasing concave function of the quadratic loss. These bounds depend only on the second and *third* moments of $W = ||\mathbf{X} - \boldsymbol{\theta}||$, which improves upon the previous results of Shinozaki (1984) and James and Stein (1961). In Shinozaki (1984), the second and *fourth* moments are used, while in James and Stein (1961), the fourth moment $EW_i^4 = E|X_i - \theta_i|^4$ is assumed to be finite. We also discuss the generalizations of the estimator (1.1). Furthermore, the class of sign-invariant distributions considered here is much larger than those referenced above. The class includes spherically symmetric distributions as well as iis distributions (symmetric about the mean). The basic properties of sign-invariant distributions can be found in Berman (1965a, b). Finally, we point out that our proofs utilize certain *positively associated* inequalities, which is considerably simpler than the integration-by-parts used by Shinozaki (1984).

The paper is organized as follows. In Section 2 we define sign-invariant distribution and present the main results about estimator (1.1). We obtain bounds for a and b with one observation $\mathbf{X}$. To illustrate the performance of bounds on a and b, two examples are presented after the main results. Extensions of the main results in three other directions are considered in Section 3. First, we consider loss functions which are monotone-concave functions of quadratic loss and show dominance results for estimator (1.1). Second, we consider dominance results for the more general estimator $[1 - br(||\mathbf{X}||^2)/(a + ||\mathbf{X}||^2)]\mathbf{X}$. Third, we consider the estimators of the location parameters when the scale is unknown and the observation $(\mathbf{X}', \mathbf{V}')'$ contains a residual vector $\mathbf{V}$. These extensions are given, respectively, by Theorems 2, 3 and 4. In Section 4 we consider the same problem with n iid observations $\mathbf{X}_1, \ldots, \mathbf{X}_n$. Under sign-invariance, any translation-invariant location estimator based on n observations also has a sign-invariant distribution. Thus the problem is reduced to the one treated in Sections 2 and 3. Section 5 is devoted to conclusions of this article along with some open research problems in this area. The last section consists of proofs of all results in Sections 2 and 3.

2. Main results. Let $\mathbf{Z} = (Z_1, \dots, Z_p)'$ be a $p \times 1$ random vector, and let $\mathscr{H}$ be a $p \times p$ diagonal matrix that belongs to $\mathscr{H}(p) = \{\mathrm{diag}(h_1, \dots, h_p), h_i^2 = 1, i = 1, \dots, p\}$. The notation $X \overset{d}{=} Y$ means that X and Y have the same distributions.

DEFINITION 1. The distribution of $\mathbf{Z}$ is said to be sign-invariant if

$$\mathscr{H}\mathbf{Z} \overset{d}{=} \mathbf{Z} \tag{2.1}$$

for every $\mathscr{H} \in \mathscr{H}(p)$. We write $\mathbf{Z} \sim \mathrm{SI}_p(\mathbf{0})$. The notation $\mathbf{X} \sim \mathrm{SI}_p(\boldsymbol{\theta})$ means that $\mathbf{X} - \boldsymbol{\theta} \sim \mathrm{SI}_p(\mathbf{0})$. If, in addition, the distribution of the random vector $(|X_1 - \theta_1|, \dots, |X_p - \theta_p|)$ is exchangeable and $P(X_i - \theta_i = 0) = 0$ for $1 \le i \le p$, we write $\mathbf{X} \sim \mathrm{SIE}_p(\boldsymbol{\theta})$. We focus our discussion on the family $\mathrm{SIE}_p(\boldsymbol{\theta})$.

Let $\boldsymbol{\delta} = (\delta_1, \dots, \delta_p)'$ be any estimator of $\boldsymbol{\theta}$. Consider the quadratic loss function

$$L(\boldsymbol{\delta}, \boldsymbol{\theta}) = ||\boldsymbol{\delta} - \boldsymbol{\theta}||^2 = \sum_{i=1}^{p} (\delta_i - \theta_i)^2. \tag{2.2}$$

Let $R(\boldsymbol{\delta}, \boldsymbol{\theta}) = E[L(\boldsymbol{\delta}, \boldsymbol{\theta})] = E||\boldsymbol{\delta} - \boldsymbol{\theta}||^2$ be the risk of $\boldsymbol{\delta}$. In this section we determine the ranges of a and b for which the risk $R(\boldsymbol{\delta}_{a,b}, \boldsymbol{\theta})$ of the estimator (1.1) is smaller than that of $R(\boldsymbol{\delta}_{a,0}, \boldsymbol{\theta}) = R(\mathbf{X}, \boldsymbol{\theta})$.

THEOREM 1. *Suppose that* $\mathbf{X} \sim \mathrm{SIE}_p(\boldsymbol{\theta})$ *and* $\boldsymbol{\delta}_{a,b}(\mathbf{X})$ *is defined by* (1.1). *Then, with respect to the quadratic loss* (2.2), $\boldsymbol{\delta}_{a,b}(\mathbf{X})$ *has smaller risk than* $\boldsymbol{\delta}_{a,0}(\mathbf{X}) = \mathbf{X}$ *provided* $a \in [a^*(p), \infty)$ *and* $b \in (0, b^*(a, p)]$, *where*

$$\begin{aligned} a^*(p) &= [2S/(p-2)]^2, \\ b^*(a, p) &= 2[(p-2)a - 2S\sqrt{a}\,]/(4 + pa\mu_{-2}), \end{aligned} \tag{2.3}$$

$\mu_i = EW^i$ *for* $i = -2, 2, 3$, $S = \mu_3/\mu_2$ *and* $W = ||\mathbf{X} - \boldsymbol{\theta}||$.

REMARK 1. The bounds for a and b given by Theorem 1 are by no means optimal, and they are considered as a guide post. If we know the distribution of $\mathscr{W}$, then it is possible to get much sharper bounds for a and b by using either the above methods or other methods in the literature. To illustrate the performance of the bounds in Theorem 1, we present two examples below.

EXAMPLE 1. Assume that the vector $\mathbf{Z}$ has a uniform distribution on the unit sphere with $p \ge 4$. Thus, $W = ||\mathbf{Z}||$ has density pw^{p-1}, $0 \le w \le 1$, and $\mu_{-2} = p/(p-2)$, $\mu_2 = p/(p+2)$ and $\mu_3 = p/(p+3)$. Applying Brandwein and Strawderman's (1991) result yields $a \ge 0$ and $0 < b \le b_{\mathrm{BS}} = 2[(p-2)/p]^2$. Applying Evans and Stark's (1996) result gives $a = 1$ and $0 < b \le b_{\mathrm{ES}}(\alpha^*) = 2[(p-2)/(p+2)][\alpha^*/(2+\alpha^*)]^6$, where α^* is the unique positive root of

$$(p-2)\alpha^6[1 + (2+\alpha)^2]^2 - p(2+\alpha)^4 = 0.$$

TABLE 1
Bounds for b ($a = 1$)

p	4	6	8	10	15	20	30	50	100
b_{BS}	0.500	0.889	1.125	1.280	1.502	1.620	1.742	1.843	1.921
$b_{\mathrm{ES}}(1.2)$	0.002	0.003	0.003	0.004	0.004	0.005	0.005	0.005	0.005
$b^*(1, p)$	0.048	0.342	0.570	0.746	1.043	1.227	1.442	1.642	1.811

It can be shown that $0 < \alpha^* < 1.2$. Thus, $b_{\mathrm{ES}}(\alpha^*) < b_{\mathrm{ES}}(1.2)$ because $\alpha^*/(2+\alpha^*)$ is nondecreasing in α^*. Applying $b^*(a, p)$ in Theorem 1 with $a = 1$ gives $b^*(1, p) = 2[(p-2-2(p+2)/(p+3))/(p^2/(p-2)+4)]$. Table 1 provides the values of three bounds for different p with $a = 1$.

It is not surprising that b_{BS} is the best among all three cases since the other two are obtained under the weaker assumptions. One can see from Table 1 that $b^*(1, p)$ is much better than $b_{\mathrm{ES}}(\alpha^*)$. Note that Shinozaki's (1984) result is not applicable to this example.

EXAMPLE 2. Assume that the vector $\mathbf{Z}$ has iid components with $P(Z_1 = \pm 1) = 1/2$. Applying Shinozaki's (1984) result yields $a \geq a_0(p)$, $0 < b \leq b_0(a, p) = [4(p-2)a - 3]/[2(a+1)]$ or $a \geq a_1(p)$, $0 < b \leq b_1(p) = p - 2$, where $a_0(p) = (p+2)I_{[p\in\{3, 4, 5\}]} + (p+1)I_{[p\in\{6, 7, 8\}]} + pI_{[p\in\{9, 10, 11\}]} + (p-1)I_{[p\in\{12, 13, 14\}]} + (p-2)I_{[p\geq 15]}$ and $a_1(p) = 4.5I_{[p=3]} + 4.2I_{[p=4]} + 4I_{[p\geq 5]}$ and $I_{[A]}$ is the indicator function of A. It can be checked from Theorem 1 that $b_0(a, p) \geq b^*(a, p)$ and $b_1(a, p) \geq b^*(a, p)$ for $a \leq [2(\sqrt{p} + \sqrt{p + (p-3)^3})/(p-2)]^2$. However, when $p \geq 4$, $a^*(p) < a_i(p)$, $i = 0, 1$, and $b^*(a, p) \geq b_1(a, p)$ for $a \geq [2(\sqrt{p} + \sqrt{p + (p-3)^3})/(p-2)]^2$. It is worth mentioning that the lower bound $a_0(p)$ is an increasing function of p while the bound $a^*(p)$ in Theorem 1 is a nonincreasing function in p. For instance, $a^*(15) \approx 0.391$ while $a_0(15) = 13$. Note that Brandwein and Strawderman's (1991) and Evans and Stark's (1996) results are not applicable to this example.

3. Extensions to other loss functions and to general estimators. In this section we show that Theorem 1 in Section 2 can be generalized to a larger class of loss functions and to a larger class of estimators. The loss function we consider is

$$L(\boldsymbol{\delta}, \boldsymbol{\theta}) = l(||\boldsymbol{\delta} - \boldsymbol{\theta}||^2), \tag{3.1}$$

where $l(\cdot)$ is a nonnegative and nondecreasing concave function. This loss function has been studied for the spherically symmetric distributions by Bock (1985) and Brandwein and Strawderman (1980, 1991).

THEOREM 2. *Let $F(\cdot)$ be the cdf of the random variable W^2 with*

$$0 < \int_0^\infty l^{(1)}(r)\, dF(r) < \infty,$$

where $l^{(1)}(r)$ is the first derivative of $l(r)$. Let $G(\cdot)$ be defined by

$$G(z) = \int_0^z l^{(1)}(r)\,dF(r) \Big/ \int_0^\infty l^{(1)}(r)\,dF(r)$$

for $z \geq 0$, that is, G is a weighted cdf of F with weight function $l^{(1)}$. If $\mathbf{X} \sim \mathrm{SIE}_p(\boldsymbol{\theta})$ and $\boldsymbol{\delta}_{a,b}(\mathbf{X})$ is defined by (1.1), then with respect to the loss (3.1), $\boldsymbol{\delta}_{a,b}(\mathbf{X})$ has smaller risk than $\boldsymbol{\delta}_{a,0}(\mathbf{X})$ provided

$$a \geq [2S_*^2/(p-2)]^2,$$

$$0 < b \leq 2[(p-2)a - 2S_*\sqrt{a}\,]/(4 + pa\mu_{-2}^*),$$

where $\mu_i^* = E_G W^i$ for $i = -2, 2, 3$, $S_* = \mu_3^*/\mu_2^*$ and $W = ||\mathbf{X} - \boldsymbol{\theta}||$.

Now we extend Theorem 1 to a larger class of estimators. The estimators we consider are of the form

$$\boldsymbol{\delta}_{a,b,r}(\mathbf{X}) = \left\{1 - \frac{br(||\mathbf{X}||^2)}{a + ||\mathbf{X}||^2}\right\}\mathbf{X}. \tag{3.2}$$

Clearly, $\boldsymbol{\delta}_{a,b,1}(\mathbf{X}) = \boldsymbol{\delta}_{a,b}(\mathbf{X})$. Estimators defined by (3.2) for spherically symmetric distributions have been studied by Brandwein and Strawderman (1980, 1991). Here we present a similar result for sign-invariant distributions.

THEOREM 3. *Let $\boldsymbol{\delta}_{a,b,r}(\mathbf{X})$ be defined by (3.2), where $\mathbf{X} \sim \mathrm{SIE}_p(\boldsymbol{\theta})$ and $p \geq 5$. Then with respect to the quadratic loss (2.2), $\boldsymbol{\delta}_{a,b,r}(\mathbf{X})$ has smaller risk than $\boldsymbol{\delta}_{a,0,r}(\mathbf{X}) = \mathbf{X}$ provided the following hold:*

(i) $0 < r(\cdot) \leq 1$;
(ii) *$r(||\mathbf{X}||^2)$ is nondecreasing in $||\mathbf{X}||^2$;*
(iii) *$a \geq a_1^*(p)$ and $0 < b \leq b_1^*(a, p)$, where*

$$a_1^* = \left\{2\left[S + \sqrt{S^2 + (p-4)T}\,\right]/(p-4)\right\}^2,$$

$$b_1^* = [(p-4)a - 4(T + \sqrt{a}\,S)]/[pa\mu_{-2} + 4]$$

and $\mu_i = E[W^i]$, $i = -2, 2, 3, 4$, $S = \mu_3/\mu_2$ and $T = \mu_4/\mu_2$.

REMARK 2. Although the ranges of a and b using the technique here require at least five dimensions instead of three, the conditions on the function $r(\cdot)$ are very weak. For spherically symmetric distributions, the function r in (3.2) has one more condition, $r(t)/t$ is nonincreasing in t. On the other hand, it is worth mentioning that Theorem 3 can be easily extended to loss function (3.1).

Now we consider the problem of estimating the mean vector $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)'$ when the observation $(\mathbf{X}', \mathbf{V}')'$ contains a residual vector $\mathbf{V}$ such that $\mathbf{X}^* = (1/\sigma)(\mathbf{X} - \boldsymbol{\theta}) \sim \mathrm{SIE}_p(\mathbf{0})$ and $\mathbf{V}^* = \mathbf{V}/\sigma \sim \mathrm{SIE}_m(\mathbf{0})$ are independent, where σ is an unknown scale. The improved estimators will be of the form

$$\boldsymbol{\delta}^*_{a,b}(\mathbf{X}, \mathbf{V}) = \left\{1 - \frac{b\,||\mathbf{V}||^2}{a||\mathbf{V}||^2 + ||\mathbf{X}||^2}\, r(||\mathbf{X}||^2)\right\}\mathbf{X}. \tag{3.3}$$

THEOREM 4. *Suppose* $\mathbf{X}$ *is a* $p \times 1$, $p \geq 5$, *vector and* $\mathbf{V}$ *is an* $m \times 1$ *vector such that* $(1/\sigma)(\mathbf{X} - \boldsymbol{\theta}) \sim \mathrm{SIE}_p(\mathbf{0})$ *and* $\mathbf{V}/\sigma \sim \mathrm{SIE}_m(\mathbf{0})$ *are independent, where* σ *is an unknown scale parameter. If* $\boldsymbol{\delta}^*_{a,\,b}(\mathbf{X}, \mathbf{V})$ *is defined by* (3.3), *then, with respect to scaled quadratic loss* $L(\boldsymbol{\delta}, \boldsymbol{\theta}) = ||\boldsymbol{\delta} - \boldsymbol{\theta}||^2/\sigma^2$, $\boldsymbol{\delta}^*_{a,\,b}(\mathbf{X}, \mathbf{V})$ *has smaller risk than* $\mathbf{X}$ *provided the following hold*:

(i) $0 < r(\cdot) \leq 1$;
(ii) $r(||\mathbf{X}||^2)$ *is nondecreasing in* $||\mathbf{X}||^2$;
(iii) $r(||\mathbf{X}||^2)/||\mathbf{X}||^2$ *is nonincreasing in* $||\mathbf{X}||^2$;
(iv) $a \geq a_2^*(p)$ *and* $0 < b \leq b_2^*(a, p)(\gamma_4/\gamma_6)$, *where*

$$\begin{aligned} a_2^*(p) &= \left\{2\left[S\gamma_{-1} + \sqrt{S^2\gamma_{-1}^2 + (p-4)T\gamma_{-2}}\,\right]/(p-4)\right\}^2, \\ b_2^*(a, p) &= \left[(p-4)a - 4(T\gamma_{-2} + \sqrt{a}\,S\gamma_{-1})\right]/[pa\nu_{-2} + 4\gamma_{-2}] \end{aligned} \tag{3.4}$$

and $\gamma_i = E_{\sigma=1,\,\theta=0}[||\mathbf{V}||^i]$, $i = -1, -2, 4, 6$, $\nu_i = E_{\sigma=1,\,\theta=0}[||\mathbf{X}||^i]$, $i = -2, 2, 3, 4$, $S = \nu_3/\nu_2$ *and* $T = \nu_4/\nu_2$.

REMARK 3. Theorem 4 is also true for the loss function (3.1).

4. Multiple observations. In practice, one would have a sample of n iid observations $\mathbf{X}_1, \dots, \mathbf{X}_n$ from a $\mathrm{SIE}_p(\boldsymbol{\theta})$, and the estimate of $\boldsymbol{\theta}$ depends on all n observations. The best invariant location estimator is Pitman's estimator defined by $\boldsymbol{\delta}(\mathbf{X}_1, \dots, \mathbf{X}_n) = \mathbf{X}_1 - E_0[\mathbf{X}_1|\mathbf{Y}_2, \dots, \mathbf{Y}_n]$, where $\mathbf{Y}_i = \mathbf{X}_i - \mathbf{X}_1$, $i = 2, \dots, n$. As pointed out by Brandwein (1979), this estimator is usually difficult to calculate and other estimators may be preferred. Let $\boldsymbol{\delta}$ be an estimator of $\boldsymbol{\theta}$ satisfying

$$\boldsymbol{\delta}(\mathscr{H}\mathbf{X}_1 - \boldsymbol{\gamma}, \dots, \mathscr{H}\mathbf{X}_n - \boldsymbol{\gamma}) = \mathscr{H}\boldsymbol{\delta}(\mathbf{X}_1, \dots, \mathbf{X}_n) - \boldsymbol{\gamma}, \tag{4.1}$$

where $\mathscr{H} \in \mathscr{H}(p)$ is defined in Section 2, and $\boldsymbol{\gamma}$ is any $p \times 1$ vector. It can be shown that $\boldsymbol{\delta}$ has a sign-invariant distribution about $\boldsymbol{\theta}$. Both Pitman's estimator and the sample mean $\overline{\mathbf{X}}$ satisfy (4.1). If $\boldsymbol{\delta}$ has a sign-invariant distribution, then the results of Sections 2 and 3 are applicable to $\boldsymbol{\delta}$.

5. Discussion. The ranges of a and b given by the theorems in Sections 2 and 3 are by no means optimal, and they are considered as a guide post. It is possible to get much sharper bounds for a and b by using other methods with specific distributions of $\mathscr{W}$. For example, the best bounds for a and b for estimator (1.1) with a normal distribution are $a \geq 0$ and $0 < b < 2(p-2)$ which are obtained by integration-by-parts. Generally, there is a reduction of the ranges of a and b without a normality assumption on the underlying distribution. The purpose of this paper is to generalize Shinozaki's (1984) result in two respects. One is to relax the distributional assumption from iis to $\mathrm{SIE}_p(\boldsymbol{\theta})$. The other is to use the second and the *third* moments of the components only while Shinozaki used the second and the *fourth* moments. On the other hand, three extensions have also been considered. The first is to extend the quadratic loss (2.2) to the loss function (3.1). The second is to extend the estimator (1.1) to estimator (3.2). The last is to investigate the estimator of the location parameters when the scale is unknown and the observation $(\mathbf{X}', \mathbf{V}')'$ contains a residual vector $\mathbf{V}$. As the referees mentioned, it would be interesting to investigate the more general estimator $\boldsymbol{\delta}_b(\mathbf{X}) = \mathbf{X} + b\mathbf{g}(\mathbf{X})$. Some nice results for this estimator have been obtained by Brandwein and Strawderman (1991) for the spherical symmetric distributions. It is known that the spherical symmetric distribution which is a mixture of the uniform distribution on the surface of the unit sphere depends on a scale $\sum_{i=1}^{p} |X_i - \theta_i|^2$. One can see from Definition 1 in Section 2 that the sign-invariant distributions $\mathbf{X} \sim \mathrm{SI}_p(\boldsymbol{\theta})$ depend on a vector $(|X_1 - \theta_1|, \ldots, |X_p - \theta_p|)$. It seems very difficult to obtain conditions on $\mathbf{g}(\mathbf{X})$ such that $\boldsymbol{\delta}_b(\mathbf{X})$ dominates $\mathbf{X}$ if $\mathbf{X} \sim \mathrm{SIE}_p(\boldsymbol{\theta})$. This difficulty can also be seen from a recent paper by Evans and Stark (1996). Under some moment assumptions on the underlying random vector $\mathbf{X}$, they also studied the estimator (1.1) with $a = 1$ and found the bound for b such that $\boldsymbol{\delta}_{1,b}(\mathbf{X})$ dominates $\boldsymbol{\delta}_{1,0}(\mathbf{X})$. It is worth pointing out that one can find a random vector $\mathbf{X}$ which is sign-invariant about its mean vector $\boldsymbol{\theta}$, but does not satisfy condition (1) of Evans and Stark's (1996) theorem, see Example 2.

6. Proofs.

PROOF OF THEOREM 1. Let $\mathbf{Z} = \mathbf{X} - \boldsymbol{\theta}$. Then $\mathbf{Z} \sim \mathrm{SIE}_p(\mathbf{0})$. By Lemma 1.2 of Berman (1965a), sign-invariant random variables are conditionally independent when their absolute values are given. Thus

$$\mathbf{Z} \stackrel{d}{=} \mathscr{W}\mathbf{U},$$

where the diagonal matrix $\mathscr{W} = \mathrm{diag}(W_1, \ldots, W_p)$ and the vector $\mathbf{U} = (U_1, \ldots, U_p)'$ are independent, $U_1, \ldots, U_p$ are independent, identically distributed with $P(U_1 = -1) = P(U_1 = 1) = 1/2$ and $W_i \stackrel{d}{=} |Z_i| = |X_i - \theta_i|$, $i = 1, \ldots, p$. The difference between the risks of two estimates $\boldsymbol{\delta}_{a,b}(\mathbf{X})$ and $\boldsymbol{\delta}_{a,0}(\mathbf{X}) = \mathbf{X}$ is given by

$$\begin{aligned}\Delta &= R(\boldsymbol{\delta}_{a,b}(\mathbf{X}), \boldsymbol{\theta}) - R(\boldsymbol{\delta}_{a,0}(\mathbf{X}), \boldsymbol{\theta}) \\ &= b^2 E_{\mathbf{X}}\left[\frac{\mathbf{X}'\mathbf{X}}{(a + \mathbf{X}'\mathbf{X})^2}\right] - 2bE_{\mathbf{X}}\left[\frac{(\mathbf{X} - \boldsymbol{\theta})'\mathbf{X}}{a + \mathbf{X}'\mathbf{X}}\right].\end{aligned}$$

We will show that, in the specified region (2.3), $\Delta \le 0$, or equivalently, $\Delta^* = -\Delta/b \ge 0$, where

$$\begin{aligned}\Delta^* &= E_{\mathbf{X}}\left[\frac{-b\mathbf{X}'\mathbf{X}}{(a+\mathbf{X}'\mathbf{X})^2}\right] + 2E_{\mathbf{X}}\left[\frac{(\mathbf{X}-\boldsymbol{\theta})'\mathbf{X}}{a+\mathbf{X}'\mathbf{X}}\right]\\ &\ge E_{\mathbf{X}}\left[\frac{-b+2(\mathbf{X}-\boldsymbol{\theta})'\mathbf{X}}{a+\mathbf{X}'\mathbf{X}}\right]\\ &= E_{\mathbf{Z}}\left[\frac{2\mathbf{Z}'\mathbf{Z}-b+2\mathbf{Z}'\boldsymbol{\theta}}{a+\mathbf{Z}'\mathbf{Z}+\boldsymbol{\theta}'\boldsymbol{\theta}+2\mathbf{Z}'\boldsymbol{\theta}}\right]\\ &= E_{(Y,W^2)}\left[\frac{2W^2-b+Y}{\beta+Y}\right].\end{aligned}$$

The last two equalities are obtained by applying $\mathbf{X}-\boldsymbol{\theta} = \mathbf{Z} \overset{d}{=} \mathscr{W}\mathbf{U}$ with $\theta^2 = \boldsymbol{\theta}'\boldsymbol{\theta}$, $W^2 = \sum_{i=1}^p W_i^2$, $\beta = a + W^2 + \theta^2$ and $Y = 2\sum_{i=1}^p W_i\theta_i U_i$. Now we use the following identity twice,

$$\frac{1}{\beta+y} = \frac{1}{\beta} - \frac{y}{\beta(\beta+y)}, \tag{6.1}$$

to obtain the conditional expectation of the ratio in the last equality given $\mathscr{W} = \operatorname{diag}(W_1, \dots, W_p)$ as follows:

$$E_Y\left[\frac{2W^2-b+Y}{\beta+Y}\,\middle|\,\mathscr{W}\right] = \frac{2W^2-b}{\beta} - \frac{4\mu}{\beta^2} + A,$$

where

$$A = E_Y\left[\frac{Y^2(2W^2-b+Y)}{\beta^2(\beta+Y)}\,\middle|\,\mathscr{W}\right],$$

$$\mu = \boldsymbol{\theta}'\mathscr{W}^2\boldsymbol{\theta}.$$

For simplicity, put $T \equiv 2W^2 - b - \beta = W^2 - (a+b+\theta^2)$. Let $I_{\mathscr{C}}$ be the indicator function of the event $\mathscr{C} = \{W^2\colon T \ge 0\}$. Noting the monotone property of the function $(2W^2-b+Y)/(\beta+Y)$ obtains

$$\begin{aligned}A &\ge \left(\frac{2W^2-b+2W\theta}{\beta+2W\theta}I_{\mathscr{C}} + \frac{2W^2-b-2W\theta}{\beta-2W\theta}I_{\mathscr{C}^c}\right)\left(\frac{4\mu}{\beta^2}\right)\\ &\ge \left(I_{\mathscr{C}} - \frac{b+\sqrt{a}\,W}{a}I_{\mathscr{C}^c}\right)\left(\frac{4\mu}{\beta^2}\right).\end{aligned}$$

The last inequality holds because the minimum value of $(2W^2-b-2W\theta)/(\beta-2W\theta)$ for all θ is $-(b+\sqrt{b^2+4aW^2})/(2a)$, which is greater than $-(b+\sqrt{a}W)/a$. Therefore, the conditional expectation of Δ^* given $\mathscr{W}$ satisfies

$$\begin{aligned}E(\Delta^*\,|\,\mathscr{W}) &\ge \frac{2W^2-b}{\beta} - \frac{4\mu}{\beta^2}\left(1+\frac{b+\sqrt{a}\,W}{a}\right)I_{\mathscr{C}^c}\\ &\ge \frac{2W^2-b}{\beta} - \frac{4\mu}{\beta^2}\left(1+\frac{b+\sqrt{a}\,W}{a}\right).\end{aligned}$$

It follows by the exchangeability assumption of $W_1, \dots, W_p$ that $E(W_i^2 \mid W^2) = W^2/p$, $i = 1, \dots, p$. Then

$$
\begin{aligned}
\Delta^* &= E_{\mathscr{W}}[E(\Delta^* \mid \mathscr{W})] \\
&\geq E_{\mathscr{W}}\left[\frac{2W^2 - b}{\beta} - \left(1 + \frac{b + \sqrt{a}W}{a}\right)\frac{4\mu}{\beta^2}\right] \\
&= E_{W^2}\left\{E_{\mathscr{W}}\left[\frac{2W^2 - b}{\beta} - \left(1 + \frac{b + \sqrt{a}W}{a}\right)\frac{4\mu}{\beta^2} \;\middle|\; W^2\right]\right\} \\
&= E_{W^2}\left[\frac{2W^2 - b}{\beta} - \frac{4W^2\theta^2}{p\beta^2}\left(1 + \frac{b + \sqrt{a}W}{a}\right)\right] \qquad (6.2)\\
&\geq E_{W^2}\left[\frac{2W^2 - b}{\beta} - \frac{4W^2}{p\beta}\left(1 + \frac{b + \sqrt{a}W}{a}\right)\right] \\
&= E_{W^2}\left\{\left[-b\left(\frac{1}{W^2} + \frac{4}{pa}\right) + \left(2\left(1 - \frac{2}{p}\right) - \frac{4W}{p\sqrt{a}}\right)\right]\left(\frac{W^2}{\beta}\right)\right\}.
\end{aligned}
$$

Note that

$$
E_{W^2}\left[\left(\frac{1}{W^2} + \frac{4}{pa}\right)\left(\frac{W^2}{\beta}\right)\right] \leq \left[\mu_{-2} + \frac{4}{pa}\right]E_{W^2}\left(\frac{W^2}{\beta}\right) \qquad (6.3)
$$

because $1/W^2 + 4/(pa)$ is nonincreasing and W^2/β is nondecreasing in W^2. On the other hand, applying Theorem 2 of Wijsman (1985) with $f_1(W^2) = 2(1 - 2/p) - 4W/(p\sqrt{a})$, $f_2(W^2) = 1$, $g_1(W^2) = W^2/\beta$ and $g_2(W^2) = W^2$ yields that

$$
\begin{aligned}
E_{W^2}\left[\left(2 - \frac{4}{p} - \frac{4W}{p\sqrt{a}}\right)\frac{W^2}{\beta}\right] &= E_{W^2}[f_1(W^2)g_1(W^2)] \\
&\geq \frac{E_{W^2}[f_1(W^2)g_2(W^2)]}{E_{W^2}[f_2(W^2)g_2(W^2)]}E_{W^2}[f_2(W^2)g_1(W^2)] \qquad (6.4)\\
&= \left[2\left(1 - \frac{2}{p}\right) - \frac{4S}{p\sqrt{a}}\right]E_{W^2}\left[\frac{W^2}{\beta}\right],
\end{aligned}
$$

where $S = \mu_3/\mu_2$ and $\mu_i = E[W^i]$, $i = 2, 3$. Thus, $\Delta^* \geq 0$ follows immediately by combining (6.3) and (6.4) and using the range (2.3). □

PROOF OF THEOREM 2. First assume that the W_i's are given. Using the same approach as in Brandwein and Strawderman (1980) we obtain that

$$
R(\mathbf{X}, \boldsymbol{\theta}) - R(\boldsymbol{\delta}_{a,b}(\mathbf{X}), \boldsymbol{\theta}) = El(W^2) - El(W^2 - \Delta_{a,b}(\mathbf{X}))
$$

where

$$
\Delta_{a,b}(\mathbf{X}) = \|\mathbf{X} - \boldsymbol{\theta}\|^2 - \|\boldsymbol{\delta}_{a,b}(\mathbf{X}) - \boldsymbol{\theta}\|^2.
$$

Since $l(\cdot)$ is a concave function,

$$
l(W^2 - \Delta_{a,b}(\mathbf{X})) < l(W^2) + l^{(1)}(W^2)[-\Delta_{a,b}(\mathbf{X})].
$$

Let $c = 1 - 2(a+b)/(pa)$ and $h(W^2) = [-b + 2cW^2 - 4W^3/(p\sqrt{a})]/\beta$. Then

$$
\begin{aligned}
&R(\mathbf{X}, \boldsymbol{\theta}) - R(\boldsymbol{\delta}_{a,b}(\mathbf{X}), \boldsymbol{\theta}) \\
&\quad \geq E_{\mathbf{X}}\,[l^{(1)}(W^2)\Delta_{a,b}(\mathbf{X})] \\
&\quad \geq E_{\mathscr{W}}\{l^{(1)}(W^2)\, E_{\mathbf{U}}[\Delta_{a,b}(\mathscr{W}\mathbf{U} + \boldsymbol{\theta}) \,|\, \mathscr{W}]\} \\
&\quad \geq E_{\mathscr{W}}\left\{l^{(1)}(W^2)\left(\frac{2W^2 - b}{\beta} - \left[1 + \frac{b + \sqrt{a}W}{a}\right]\frac{4\mu}{\beta^2}\right)\right\} \\
&\quad = E_{W^2}\left\{E_{\mathscr{W}}\left\{l^{(1)}(W^2)\left(\frac{2W^2 - b}{\beta} - \left[1 + \frac{b + \sqrt{a}W}{a}\right]\frac{4\mu}{\beta^2}\right) \,\middle|\, W^2\right\}\right\} \\
&\quad \geq E_{W^2}\left[l^{(1)}(W^2)\left(\frac{2W^2 - b}{\beta} - \left[1 + \frac{b + \sqrt{a}W}{a}\right]\frac{4W^2}{p\beta}\right)\right] \\
&\quad = E_{W^2}[l^{(1)}(W^2)h(W^2)] \\
&\quad = \int_0^\infty h(w^2)l^{(1)}(w^2)\,dF(w^2) \\
&\quad = \int_0^\infty h(w^2)\,dG(w^2) \int_0^\infty l^{(1)}(w^2)\,dF(w^2) \\
&\quad = E_G[h(W^2)]\,E_F[l^{(1)}(W^2)].
\end{aligned}
$$

The next-to-last equality above follows from the definition of the cdf G. The result follows immediately from the assumption that $E_F[l(W^2)] \in (0, \infty)$ and the proof of Theorem 1 except for changing the cdf F to G. □

PROOF OF THEOREM 3. Similarly to the proof of Theorem 1, it suffices to show $\Delta_1 \geq 0$, where

$$
\begin{aligned}
\Delta_1 &= [1/b]\,[R(\boldsymbol{\delta}_{a,0,r}(\mathbf{X}), \boldsymbol{\theta}) - R(\boldsymbol{\delta}_{a,b,r}(\mathbf{X}), \boldsymbol{\theta})] \\
&= E_{\mathbf{X}}\left[\frac{-b\,r^2(||\mathbf{X}||^2)||\mathbf{X}||^2}{(a + ||\mathbf{X}||^2)^2}\right] + 2E_{\mathbf{X}}\left[\frac{(\mathbf{X} - \boldsymbol{\theta})'\mathbf{X}\,r(||\mathbf{X}||^2)}{a + ||\mathbf{X}||^2}\right] \\
&\geq E_{\mathbf{X}}\left[\frac{-br(||\mathbf{X}||^2)}{a + ||\mathbf{X}||^2}\right] + 2E_{\mathbf{X}}\left[\frac{(||\mathbf{X}||^2 - ||\mathbf{X}||\,||\boldsymbol{\theta}||)\,r(||\mathbf{X}||^2)}{a + ||\mathbf{X}||^2}\right] \\
&= E_{\mathbf{X}}\left[\frac{2||\mathbf{X}||^2 - 2||\mathbf{X}||\,||\boldsymbol{\theta}|| - b}{a + ||\mathbf{X}||^2}r(||\mathbf{X}||^2)\right] \\
&\geq r(\eta^2)\,E_{\mathbf{X}}\left[\frac{2||\mathbf{X}||^2 - 2||\mathbf{X}||\,||\boldsymbol{\theta}|| - b}{a + ||\mathbf{X}||^2}\right],
\end{aligned}
\tag{6.5}
$$

where $\eta = (\theta + \sqrt{\theta^2 + 2b}\,)/2$, $\theta^2 = \boldsymbol{\theta}'\boldsymbol{\theta}$. The last inequality of (6.5) follows from the facts that the function $r(t^2)$ is nondecreasing in t^2 and that the function $2t^2 - 2t\theta - b < 0$ if $t < \eta$ and $2t^2 - 2t\theta - b > 0$ for $t > \eta$ and crosses 0 when

$t = \eta$. Thus, it suffices to show that

$$\Delta_1^* = E_{\mathbf{X}}\left[\frac{2||\mathbf{X}||^2 - 2||\mathbf{X}||\,||\boldsymbol{\theta}|| - b}{a + ||\mathbf{X}||^2}\right] \geq 0$$

because of condition (i) of Theorem 3. Using the identity (6.1) twice and $\mathbf{X}-\boldsymbol{\theta} = \mathbf{Z} \stackrel{d}{=} \mathscr{W}\mathbf{U}$ with $W^2 = \sum_{i=1}^p W_i^2$, $\beta = a + W^2 + \theta^2$, $Y = 2\sum_{i=1}^p W_i \theta_i U_i$ we obtain the conditional expectation of Δ_1^* given $\mathscr{W} = \operatorname{diag}(W_1, \dots, W_p)$ as follows:

$$E_{\mathbf{X}}\left[\frac{2||\mathbf{X}||^2 - 2||\mathbf{X}||\,||\boldsymbol{\theta}|| - b}{a + ||\mathbf{X}||^2} \;\middle|\; \mathscr{W}\right] \geq E_Y\left[\frac{W^2 - b + Y}{\beta + Y} \;\middle|\; \mathscr{W}\right]$$
$$= \frac{W^2 - b}{\beta} - \frac{4\mu}{\beta^2} + A_1,$$

where

$$A_1 = E_Y\left[\frac{Y^2(W^2 - b + Y)}{\beta^2(\beta + Y)} \;\middle|\; \mathscr{W}\right],$$
$$\mu = \boldsymbol{\theta}'\mathscr{W}^2\boldsymbol{\theta}.$$

Since $(W^2 - b + Y)/(\beta + Y)$ is a strictly increasing function of Y,

$$\begin{aligned} A_1 &\geq \left(\frac{W^2 - b - 2W\theta}{\beta - 2W\theta}\right)\left(\frac{4\mu}{\beta^2}\right) \\ &\geq -\left(\frac{1}{2a}\right)\left(W^2 + b + \sqrt{(W^2 + b)^2 + 4aW^2}\right)\left(\frac{4\mu}{\beta^2}\right) \\ &\geq -\left(\frac{1}{a}\right)(W^2 + b + \sqrt{a}\,W)\left(\frac{4\mu}{\beta^2}\right). \end{aligned}$$

The last two inequalities hold because the minimum value of $(W^2 - b - 2W\theta)/(\beta - 2W\theta)$ for all θ is $-(W^2 + b + \sqrt{(W^2 + b)^2 + 4aW^2})/(2a)$, which is greater than $-(W^2 + b + \sqrt{a}W)/a$. Therefore, the conditional expectation of Δ_1^* given $\mathscr{W}$ satisfies

$$E(\Delta_1^* \,|\, \mathscr{W}) \geq \frac{W^2 - b}{\beta} - \frac{4\mu}{\beta^2}\left(1 + \frac{W^2 + b + \sqrt{a}\,W}{a}\right),$$

which implies that

$$\begin{aligned} \Delta_1^* &= E_{\mathscr{W}}[E(\Delta_1^* \,|\, \mathscr{W})] \\ &\geq E_{\mathscr{W}}\left[\frac{W^2 - b}{\beta} - \frac{4\mu}{\beta^2}\left(1 + \frac{W^2 + b + \sqrt{a}\,W}{a}\right)\right] \\ &= E_{W^2}\left[\frac{W^2 - b}{\beta} - \frac{4W^2\theta^2}{p\beta^2}\left(1 + \frac{W^2 + b + \sqrt{a}\,W}{a}\right)\right] \\ &\geq E_{W^2}\left[\frac{W^2 - b}{\beta} - \frac{4W^2}{p\beta}\left(1 + \frac{W^2 + b + \sqrt{a}\,W}{a}\right)\right] \qquad (6.6) \end{aligned}$$

$$= -b\,E_{W^2}\left[\left(1+\frac{4W^2}{pa}\right)\frac{1}{\beta}\right] + E_{W^2}\left[\left(1-\frac{4}{p}-\frac{4W}{p\sqrt{a}}-\frac{4W^2}{pa}\right)\frac{W^2}{\beta}\right]$$

$$\geq -b\left[\mu_{-2}+\frac{4}{pa}\right]E_{W^2}\left(\frac{W^2}{\beta}\right)$$

$$+ E_{W^2}\left[\left(1-\frac{4}{p}-\frac{4W}{p\sqrt{a}}-\frac{4W^2}{pa}\right)\frac{W^2}{\beta}\right].$$

Applying Theorem 2 of Wijsman (1985) with $f_1(W^2) = 1-4/p-4W/(p\sqrt{a})-4W^2/(pa)$, $f_2(W^2) = 1$, $g_1(W^2) = W^2/\beta$ and $g_2(W^2) = W^2$ yields that

$$E_{W^2}\left[\left(1-\frac{4}{p}-\frac{4W}{p\sqrt{a}}-\frac{4W^2}{pa}\right)\frac{W^2}{\beta}\right] \geq \left[1-\frac{4}{p}-\frac{4S}{p\sqrt{a}}-\frac{4T}{pa}\right]E_{W^2}\left(\frac{W^2}{\beta}\right), \tag{6.7}$$

where $\mu_i = E[W^i]$, $i = 2, 3, 4$. Combining (6.6) and (6.7) we conclude that $\Delta_1^* \geq 0$ under condition (iii) of Theorem 3. □

PROOF OF THEOREM 4. The difference between the risks of two estimates $\boldsymbol{\delta}^*_{a,b}(\mathbf{X},\mathbf{V})$ and $\mathbf{X}$ is given by

$$\Delta_2 = R(\boldsymbol{\delta}^*_{a,b}(\mathbf{X},\mathbf{V}),\boldsymbol{\theta}) - R(\mathbf{X},\boldsymbol{\theta})$$

$$= \left(\frac{1}{\sigma^2}\right)\left\{b^2 E_{(\mathbf{X},\mathbf{V})}\left[\frac{||\mathbf{V}||^4\,||\mathbf{X}||^2}{(a||\mathbf{V}||^2+||\mathbf{X}||^2)^2}\,r^2(||\mathbf{X}||^2)\right]\right.$$

$$\left. - 2bE_{(\mathbf{X},\mathbf{V})}\left[\frac{(\mathbf{X}-\boldsymbol{\theta})'\mathbf{X}||\mathbf{V}||^2}{a||\mathbf{V}||^2+||\mathbf{X}||^2}r(||\mathbf{X}||^2)\right]\right\}.$$

We will show that, in the specified region (3.4), $\Delta_2 \leq 0$ or equivalently $\Delta_2^* = -\sigma^2\Delta_2/b \geq 0$, where

$$\Delta_2^* = E_{(\mathbf{X},\mathbf{V})}\left(\frac{-b||\mathbf{X}||^2\,||\mathbf{V}||^4}{(a||\mathbf{V}||^2+||\mathbf{X}||^2)^2}\,r^2(||\mathbf{X}||^2)\right)$$

$$+ 2\,E_{(\mathbf{X},\mathbf{V})}\left(\frac{(\mathbf{X}-\boldsymbol{\theta})'\mathbf{X}\,||\mathbf{V}||^2}{a||\mathbf{V}||^2+||\mathbf{X}||^2}\,r(||\mathbf{X}||^2)\right)$$

$$\geq E_{(\mathbf{X},\mathbf{V})}\left[\frac{-b||\mathbf{V}||^4\,r(||\mathbf{X}||^2)}{a||\mathbf{V}||^2+||\mathbf{X}||^2}\right] \tag{6.8}$$

$$+ 2E_{(\mathbf{X},\mathbf{V})}\left[\frac{(||\mathbf{X}||^2-||\mathbf{X}||\,||\boldsymbol{\theta}||)\,||\mathbf{V}||^2\,r(||\mathbf{X}||^2)}{a||\mathbf{V}||^2+||\mathbf{X}||^2}\right]$$

$$= E_{(\mathbf{X},\mathbf{V})}\left[\frac{2||\mathbf{X}||^2-2||\mathbf{X}||\,||\boldsymbol{\theta}||-b||\mathbf{V}||^2}{a||\mathbf{V}||^2+||\mathbf{X}||^2}||\mathbf{V}||^2\,r(||\mathbf{X}||^2)\right]$$

$$\geq E_{(\mathbf{X},\mathbf{V})}\left[\frac{2||\mathbf{X}||^2-2||\mathbf{X}||\,||\boldsymbol{\theta}||-b||\mathbf{V}||^2}{a||\mathbf{V}||^2+||\mathbf{X}||^2}||\mathbf{V}||^2\,r(\rho^2)\right],$$

where $\rho = (\theta + \sqrt{\theta^2 + 2b||\mathbf{V}||^2})/2$, $\theta = ||\boldsymbol{\theta}||$. The last inequality of (6.8) follows from the facts that the function $r(t^2)$ is nondecreasing in t^2 and that the function $2t^2 - 2t\theta - b||\mathbf{V}||^2 < 0$ if $t < \rho$ and $2t^2 - 2t\theta - b||\mathbf{V}||^2 > 0$ for $t > \rho$ and crosses 0 when $t = \rho$. Applying $\mathbf{X}_* = (1/\sigma)(\mathbf{X} - \boldsymbol{\theta}) \sim \mathrm{SIE}_p(\mathbf{0})$ and $\mathbf{V}_* = \mathbf{V}/\sigma \sim \mathrm{SIE}_m(\mathbf{0})$ and using the fact that $2||\mathbf{X}||\,||\boldsymbol{\theta}|| \le ||\mathbf{X}||^2 + ||\boldsymbol{\theta}||^2$ yield that

$$\begin{aligned}\Delta_2^{**} &= E_{\mathbf{X}}\left(\frac{2||\mathbf{X}||^2 - 2||\mathbf{X}||\,||\boldsymbol{\theta}|| - b||\mathbf{V}||^2}{a||\mathbf{V}||^2 + ||\mathbf{X}||^2}\,\middle|\, ||\mathbf{V}||^2\right)\\ &= E_{\mathbf{X}_*}\left(\frac{||\mathbf{X}_*||^2 - b||\mathbf{V}_*||^2 + Y_*}{a||\mathbf{V}_*||^2 + ||\mathbf{X}_*||^2 + \theta_*^2 + Y_*}\,\middle|\, ||\mathbf{V}_*||^2\right),\end{aligned}$$

where $\theta_*^2 = \theta^2/\sigma^2$ and $Y_* = Y/\sigma$. Note that we have $(a||\mathbf{V}_*||^2, b||\mathbf{V}_*||^2)$ here instead of (a, b) in the proof of Theorem 3. Using the same argument as in the proof of Theorem 3 we obtain that

$$\begin{aligned}\Delta_2^{**} \ge{}& -b||\mathbf{V}_*||^2\left[\nu_{-2} + \frac{4}{pa||\mathbf{V}_*||^2}\right]E\left(\frac{W_*^2}{\beta_*}\,\middle|\, ||\mathbf{V}_*||^2\right)\\ &+\left[1 - \frac{4}{p} - \frac{4\nu_3/\nu_2}{p\sqrt{a}\,||\mathbf{V}_*||} - \frac{4\nu_4/\nu_2}{pa||\mathbf{V}_*||^2}\right]E\left(\frac{W_*^2}{\beta_*}\,\middle|\, ||\mathbf{V}_*||^2\right),\end{aligned}$$

where $\nu_i = E(W_*^i \mid ||\mathbf{V}_*||^2)$, $i = -2, 2, 3, 4$, $W_* = ||\mathbf{X}_*||$ and $\beta_* = a||\mathbf{V}_*||^2 + W_*^2 + \theta_*^2 + Y_*$. Let $H(||\mathbf{V}_*||^2) = ||\mathbf{V}_*||^2\, E_{\mathbf{V}_*}(W_*^2/\beta_* \mid ||\mathbf{V}_*||^2)$. Then $H(||\mathbf{V}_*||^2)$ is nondecreasing in $||\mathbf{V}_*||^2$. Noting that the ν_i, $i = -2, 2, 3, 4$, are independent of $\mathbf{V}_*$ and using the assumptions of the theorem yield that

$$\begin{aligned}\left(\frac{1}{\sigma^2}\right)\Delta_2^{*} &= E[\Delta_2^{**}\; ||\mathbf{V}_*||^2\; r(\rho^2)]\\ &\ge -bE_{\mathbf{V}_*}\left\{\left[\nu_{-2} + \frac{4}{pa||\mathbf{V}_*||^2}\right]\left[||\mathbf{V}_*||^2\; r(\rho^2)\; H(||\mathbf{V}_*||^2)\right]\right\}\\ &\quad + E_{\mathbf{V}_*}\left\{\left[1 - \frac{4}{p} - \frac{4\nu_3/\nu_2}{p\sqrt{a}\,||\mathbf{V}_*||} - \frac{4\nu_4/\nu_2}{pa||\mathbf{V}_*||^2}\right]\left[r(\rho^2)\; H(||\mathbf{V}_*||^2)\right]\right\}\\ &\ge -b\left[\nu_{-2} + \frac{4\gamma_{-2}}{pa}\right]E_{\mathbf{V}_*}\left[||\mathbf{V}_*||^2\, r(\rho^2)\, H(||\mathbf{V}_*||^2)\right]\\ &\quad + \left[1 - \frac{4}{p} - \frac{4\nu_3\gamma_{-1}}{p\sqrt{a}\,\nu_2} - \frac{4\nu_4\gamma_{-2}}{pa\nu_2}\right]E_{\mathbf{V}_*}\left[r(\rho^2)H(||\mathbf{V}_*||^2)\right]\\ &\ge -b\left[\nu_{-2} + \frac{4\gamma_{-2}}{pa}\right]E_{\mathbf{V}_*}\left[||\mathbf{V}_*||^2\, r(\rho^2)\, H(||\mathbf{V}_*||^2)\right]\\ &\quad + \left[1 - \frac{4}{p} - \frac{4\nu_3\gamma_{-1}}{p\sqrt{a}\,\nu_2} - \frac{4\nu_4\gamma_{-2}}{pa\nu_2}\right]\left(\frac{\gamma_4}{\gamma_6}\right)E_{\mathbf{V}_*}\left[||\mathbf{V}_*||^2\, r(\rho^2)\, H(||\mathbf{V}_*||^2)\right]\\ &\ge 0.\end{aligned} \tag{6.9}$$

The second term of the second to last inequality of (6.9) is obtained by applying Theorem 2 of Wijsman (1985) with $f_1(||\mathbf{V}_*||^2) = 1$, $f_2(||\mathbf{V}_*||^2) =$

$||\mathbf{V}_*||^2$, $g_1(||\mathbf{V}_*||^2) = r(\rho^2)H(||\mathbf{V}_*||^2)$ and $g_2(||\mathbf{V}_*||^2) = ||\mathbf{V}_*||^4$ with the facts that $r(\rho^2)/||\mathbf{V}_*||^2$ and $H(||\mathbf{V}_*||^2)/||\mathbf{V}_*||^2$ are nonnegative and nonincreasing in $||\mathbf{V}_*||^2$. The last inequality of (6.9) follows from (3.4). The proof is complete. □

Acknowledgments. The author wishes to thank an Associate Editor, the referees and Professor Grace Yang for many valuable suggestions and comments, having led to substantial improvements of the paper.

REFERENCES

BERMAN, S. M. (1965a). Sign-invariant random variable and stochastic processes with sign-invariant increments. *Trans. Amer. Math. Soc.* **119** 216–243.

BERMAN, S. M. (1965b). Sign-invariant random elements in topological groups. *Proc. Fifth Berkeley Symp. Math. Statist. Probab.* **2** 1–11. Univ. California Press, Berkeley.

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. (1979). Minimax estimation of mean of spherically symmetric distributions under general quadratic loss. *J. Multivariate Anal.* **9** 579–588.

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. (1980). Minimax estimation of location parameters for spherically symmetric distributions with concave loss. *Ann. Statist.* **8** 279–284.

BRANDWEIN, A. C. and STRAWDERMAN, W. E. (1990). Stein estimation, the spherically symmetric case. *Statist. Sci.* **5** 356–369.

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

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

EVANS, S. N. and STARK, P. B. (1996). Shrinkage estimators, Skorokhod's problem and stochastic integration by parts. *Ann. Statist.* **24** 809–815.

JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* **1** 361–379. Univ. California Press, Berkeley.

MICELI, R. J. and STRAWDERMAN, W. E. (1986). Minimax estimation for certain independent component distributions under weighted squared error loss. *Comm. Statist. Theory Methods* **15** 2191–2200.

SHINOZAKI, N. (1984). Simultaneous estimation of location parameters under quadratic loss. *Ann. Statist.* **12** 322–335.

STEIN, C. (1956). Inadmissibility of the usual estimator for the mean vector of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 197–206. Univ. California Press, Berkeley.

WIJSMAN, R. A. (1985). A useful inequality on ratios of integrals, with application to maximum likelihood estimation. *J. Amer. Statist. Assoc.* **80** 472–475.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF HOUSTON
HOUSTON, TEXAS 77204-3476
E-MAIL: jlx@math.uh.edu