JOURNAL OF MULTIVARIATE ANALYSIS **59**, 109–140 (1996)
ARTICLE NO. 0056

A Paradox Concerning Shrinkage Estimators: Should a Known Scale Parameter Be Replaced by an Estimated Value in the Shrinkage Factor?

Dominique Fourdrinier

Université de Rouen, Rouen, France

and

William E. Strawderman*

Rutgers University

When estimating, under quadratic loss, the location parameter θ of a spherically symmetric distribution with known scale parameter, we show that it may be that the common practice of utilizing the residual vector as an estimate of the variance is preferable to using the known value of the variance. In the context of Stein-like shrinkage estimators, we exhibit sufficient conditions on the spherical distributions for which this paradox occurs. In particular, we show that it occurs for t-distributions when the dimension of the residual vector is sufficiently large. The main tools in the development are upper and lower bounds on the risks of the James–Stein estimators which are exact at $\theta = 0$.

1. INTRODUCTION

We study the problem of estimating the mean vector θ of a spherically symmetric distribution when the scale parameter is known but when a residual vector U is available: more precisely, let (X, U) be a random vector around $(\theta, 0)$. The loss function is assumed to be $\|\delta - \theta\|^2$.

This problem has recently been considered by Brandwein and Strawderman [2, 3], Brandwein *et al.* [4], Cellier *et al.* [7], and Cellier and Fourdrinier [6]. These papers study classes of estimators which improve on the usual

Received March 19, 1996.

AMS 1991 subject classification: 62F10, 62H12, 62C20.

Key words and phrases: spherical symmetry, quadratic loss, James–Stein estimation, location parameter, minimaxity, robustness.

* Research supported by National Science Foundation Grants DMS-9307727 and DMS-9400476.

0047-259X/96 $18.00

minimax estimator X. A particularly important class of such estimators is the class of James–Stein estimators $(1-a/X'X)X$. An alternative class, when a residual vector U is available, is the class of robust James–Stein estimators $(1-aU'U/X'X)X$. This latter class was shown in [7] and [6] to have the important property that, for proper choices of a, they dominate X simultaneously for all spherically symmetric distributions, with appropriate mild moment conditions, hence the term "robust" James–Stein estimators.

We show in the context of Stein-like shrinkage estimation that use of the residual vector to estimate the variance in the shrinkage constant may be superior to using the known value of the variance.

This phenomenon seems paradoxical in the sense that risk behavior of an estimator may be improved by substituting an estimate for a known quantity. Furthermore, it adds, at least somewhat (and perhaps substantially), to the attractiveness of the robust James–Stein class by demonstrating not only domination of the usual estimator X simultaneously for all spherically symmetric distributions, but also domination of the usual James–Stein estimator in many cases.

We give, in Section 2, expressions for the risk of the James–Stein estimator $(1-a/X'X)X$ and of the robust James–Stein estimators $(1-aU'U/X'X)X$ (see [7] and [6] for the robustness property). As the risks depend on expectations of the form $E[(U'U)^q/X'X]$, Section 3 is devoted to upper and lower bounds on these expectations. Then Section 4 derives upper and lower bounds on the risk of James–Stein and robust James–Stein estimators. These bounds are similar in spirit to those developed in Casella and Hwang [5] in the normal case.

Using these bounds in Section 5, we are able to give sufficient conditions for domination of the robust James–Stein estimator over the usual James–Stein estimator. The main result consists in a sufficient condition for domination of the robust James–Stein estimator over the class of the James–Stein estimators. We also give conditions for domination when θ is in a neighborhood of 0 and in a neighborhood of infinity.

In Section 6, various examples of distributions illustrate the phenomenon. A basic example is the t-distribution when the dimension of the residual vector used for estimating the variance is large enough. We also consider the context where such a phenomenon cannot arise: one such context is the normal case. The reason is that an application of the Rao–Blackwell result to any robust James–Stein estimator gives rise to a better estimate in the James–Stein class.

Section 7 gives concluding remarks. We indicate that the paradox is likely to occur in situations other James–Stein estimation.

Finally, we present an appendix which contains technical lemmas helpful for risk calculations.

2. JAMES–STEIN AND ROBUST JAMES–STEIN ESTIMATORS

Assume that the random vector $(X, U)'$ has a spherically symmetric distribution with location parameter $(\theta, 0)'$. The dimension of the subvectors X and θ is p while that of U and 0 is k.

We are interested in the estimation of the unknown parameter θ under the usual quadratic loss $\|\theta-\delta\|^2$. Classical James–Stein estimators are of the form

$$\delta_{\text{JS}}^a = \left(1 - \frac{a}{X'X}\right) X, \tag{1}$$

where a is a positive constant and it is well known that such estimators improve on the usual estimator X when $p \geqslant 3$ and $0 \leqslant a \leqslant 2(p-2)$ in the particular case of a normal distribution with identity covariance matrix.

Other competing estimators, improving on X, are those which take into account the residual vector U, that is,

$$\delta_{\text{RJS}}^a = \left(1 - \frac{aU'U}{X'X}\right) X. \tag{2}$$

More general estimators of this form were proved to have in addition a nice robustness property (cf. [7] and [6]) in the sense that the improvement on X does not depend on the specification of the distribution. We will refer to the estimators of the form (2) as robust James–Stein estimators. For these last estimators, it is easy to see (cf. formula (4) of Section 4) that improvement on X occurs when $0 < a < 2(p-2)/(k+2)$.

In the following, for any θ, we denote by E_θ the expectation with respect to the underlying distribution spherically symmetric around θ, so that the risk of any estimator δ is given by

$$\mathscr{R}(\theta, \delta) = E_\theta[\|\delta - \theta\|^2].$$

The risks of the estimators (1) and (2) can be calculated simultaneously, noting that they belong to the family of estimators

$$\delta_\alpha^a = \left(1 - a\frac{(U'U)^\alpha}{X'X}\right) X.$$

Finiteness of such risks is guaranteed as soon as the second moment of the spherical distribution exists and $E_\theta[(U'U)^{2\alpha}/X'X] < \infty$, which is what we suppose in the following (see [6]).

The expression of the risk of δ_α^a is given by the following proposition.

PROPOSITION 1. *For any $\theta \in \mathbb{R}^p$, the risk of δ_α^a equals*

$$\mathscr{R}(\theta, \delta_\alpha^a) = E_0[X'X] + a^2 E_\theta\left[\frac{(U'U)^{2\alpha}}{X'X}\right] - 2a\,\frac{p-2}{k+2\alpha}\,E_\theta\left[\frac{(U'U)^{\alpha+1}}{X'X}\right].$$

Proof. Let θ be fixed in $\mathbb{R}^p$. It is easy to see that

$$\mathscr{R}(\theta, \delta_\alpha^a) = E_\theta[\,\|X-\theta\|^2] + a^2 E_\theta\left[\frac{(U'U)^{2\alpha}}{X'X}\right] - 2aE_\theta\left[\frac{(U'U)^{\alpha}\,(X-\theta)'\,X}{X'X}\right].$$

Now $E_\theta[\,\|X-\theta\|^2] = E_0[X'X]$ and the Corollary of Lemma 2 in the Appendix applied with $g(X) = X/X'X$ gives

$$E_\theta\left[\frac{(U'U)^{\alpha}\,(X-\theta)'\,X}{X'X}\right] = \frac{p-2}{k+2\alpha}\,E_\theta\left[\frac{(U'U)^{\alpha+1}}{X'X}\right],$$

recalling that $\operatorname{div}(X/X'X) = (p-2)/X'X$.

Therefore the above risk expression is proved. ∎

It is easy to deduce from the risk expression of δ_α^a that, for any $\theta \in \mathbb{R}^p$, the constant a for which the risk is minimum is

$$a(\theta) = \frac{p-2}{k+2\alpha}\,\frac{E_\theta\left[\dfrac{(U'U)^{\alpha+1}}{X'X}\right]}{E_\theta\left[\dfrac{(U'U)^{2\alpha}}{X'X}\right]}$$

and the corresponding risk is

$$E_0[X'X] - \left(\frac{p-2}{k+2\alpha}\right)^2 \frac{\left(E_\theta\left[\dfrac{(U'U)^{\alpha+1}}{X'X}\right]\right)^2}{E_\theta\left[\dfrac{(U'U)^{2\alpha}}{X'X}\right]}.$$

It is worth noting that, for the James–Stein estimator (i.e., for $\alpha = 0$), the optimal constant typically depends on θ and equals

$$a(\theta) = \frac{p-2}{k}\,\frac{E_\theta\left[\dfrac{(U'U)}{X'X}\right]}{E_\theta\left[\dfrac{1}{X'X}\right]}$$

while, for the robust James–Stein estimator (i.e., for $\alpha = 1$), the optimal constant does not depend on θ anymore and equals $(p-2)/(k+2)$.

However, note that, in the normal case $\mathcal{N}(\theta, \sigma^2 I)$, the optimal constant $a(\theta)$ does not depend on θ and is equal to $((p-2)/k)\, E_0[U'U] = (p-2)\,\sigma^2$. In general, for independence of $a(\theta)$ on θ, it would be sufficient that $U'U$ and $1/X'X$ are uncorrelated for all θ but we conjecture that this only occurs in the normal case (of course, the independence of X and U is a characterization of normality in the spherical case).

3. BOUNDS FOR $E_\theta[(U'U)^q/X'X]$

Proposition 1 indicates that bounds for the risks of the estimators (1) and (2) rest on bounds on expectations of the type $E_\theta[(U'U)^q/X'X]$ where q is an integer. The following propositions yield such upper and lower bounds. These bounds are expressed, for any fixed $R \geqslant 0$, conditionally on the radius $R = (\|X-\theta\|^2 + \|U\|^2)^{1/2}$ and we will denote by $E_{R,\theta}$ the corresponding expectation (that is, the expectation with respect to the uniform distribution $U_{R,\theta}$ on the sphere $S_{R,\theta} = \{y \in \mathbb{R}^{p+k}/\|y-\theta\| = R\}$ of radius R and centered at θ). Thus we can write $E_\theta[(U'U)^q/X'X] = E[E_{R,\theta}[(U'U)^q/X'X]]$ where E denotes the expectation with respect to the radial distribution (i.e., the distribution of R).

First we give an expression of $E_{R,\theta}[(U'U)^q/X'X]$ in terms of integrals with respect to a beta distribution.

For notational convenience we often use $B(\alpha, \beta, dv)$ for the density of the beta distribution with parameters $\alpha > 0$ and $\beta > 0$.

PROPOSITION 2. *For $p \geqslant 3$, any $R \geqslant 0$, any $\theta \in \mathbb{R}^p$, and any integer q such that $-k/2 < q$, the expectation of $(U'U)^q/X'X$ conditionally on the radius R is equal to*

$$E_{R,\theta}\left[\frac{(U'U)^q}{X'X}\right] = \frac{\Gamma\left(\frac{p+k}{2}\right)\Gamma\left(\frac{k}{2}+q\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{p+k}{2}+q\right)} R^{2q}$$

$$\times \int_0^1 \int_0^1 \frac{R^2 u + \|\theta\|^2}{(R^2 u + \|\theta\|^2)^2 - 4\,\|\theta\|^2\, R^2 u v}$$

$$\times B\left(\frac{1}{2}, \frac{p-1}{2}, dv\right) B\left(\frac{p}{2}, \frac{k}{2} + q, du\right).$$

Proof. Lemma 1 applied with $g(X'X) = 1/X'X$ and $h(U'U) = (U'U)^q$ gives, for $R \geqslant 0$ and $\theta \in \mathbb{R}^p$,

$$E_{R,\theta}\left[\frac{(U'U)^q}{X'X}\right] = R^{2q}\int_0^1 (1-u)^q \times\left[\int_0^1 \frac{R^2u+\theta^2}{(R^2u+\|\theta\|^2)^2-4\theta^2R^2uv}\, B\left(\frac{1}{2},\frac{p-1}{2}, dv\right)\right] \times B\left(\frac{p}{2},\frac{k}{2}, du\right).$$

Inserting $(1-u)^q$ into the beta distribution $B(p/2, k/2, du)$ yields the desired result. ∎

At $\theta=0$, Proposition 2 simplifies greatly in the following corollary.

COROLLARY 1. *Under the conditions of Proposition 2, we have*

$$E_{R,0}\left[\frac{(U'U)^q}{X'X}\right] = \frac{\Gamma\left(\frac{p+k}{2}\right)\Gamma\left(\frac{k}{2}+q\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{p+k}{2}+q\right)}\,\frac{p+k+2q-2}{p-2}\,R^{2q-2}$$

and, in the particular cases where $q=0$, 1, *and* 2, *we have*

$$E_{R,0}\left[\frac{1}{X'X}\right] = \frac{p+k-2}{p-2}\,\frac{1}{R^2}$$

$$E_{R,0}\left[\frac{U'U}{X'X}\right] = \frac{k}{p-2}$$

$$E_{R,0}\left[\frac{(U'U)^2}{X'X}\right] = \frac{k(k+2)}{(p+k)(p-2)}\,R^2.$$

The following proposition gives a lower bound for $E_\theta[(U'U)^q/X'X]$.

PROPOSITION 3. *Let q be an integer such that $-k/2<q$. If $p\geqslant 5$ then, for any $\theta\in\mathbb{R}^p$, we have*

$$E_\theta\left[\frac{(U'U)^q}{X'X}\right] \geqslant \frac{\Gamma\left(\frac{p+k}{2}\right)\Gamma\left(\frac{k}{2}+q\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{p+k}{2}+q\right)}\,\frac{p+k+2q-2}{p-2} \times E\left[\frac{R^{2q}}{R^2+\frac{p+k+2q-4}{p-4}\|\theta\|^2}\right].$$

With q equal to 0, 1, and 2 we directly deduce the following corollary.

COROLLARY 2. *If $p \geqslant 5$ then, for any $\theta \in \mathbb{R}^p$, we have*

$$E_\theta\left[\frac{1}{X'X}\right] \geqslant \frac{p+k-2}{p-2}\left[E\frac{1}{R^2+\dfrac{p+k-4}{p-4}\|\theta\|^2}\right]$$

$$E_\theta\left[\frac{U'U}{X'X}\right] \geqslant \frac{k}{p-2}\, E\left[\frac{R^2}{R^2+\dfrac{p+k-2}{p-4}\|\theta\|^2}\right]$$

$$E_\theta\left[\frac{(U'U)^2}{X'X}\right] \geqslant \frac{k(k+2)}{(p+k)(p-2)}\, E\left[\frac{R^4}{R^2+\dfrac{p+k}{p-4}\|\theta\|^2}\right].$$

Proof of Proposition 3. Let $R \geqslant 0$ and $\theta \in \mathbb{R}^p$ fixed. From the expression of $E_\theta[(U'U)^q/X'X]$ given by Proposition 1, it is clear that

$$E_{R,\theta}\left[\frac{(U'U)^q}{X'X}\right] \geqslant \frac{\Gamma\left(\dfrac{p+k}{2}\right)\Gamma\left(\dfrac{k}{2}+q\right)}{\Gamma\left(\dfrac{k}{2}\right)\Gamma\left(\dfrac{p+k}{2}+q\right)} R^{2q} \int_0^1 \frac{1}{R^2u+\|\theta\|^2}\, B\left(\frac{p}{2}, \frac{k}{2}+q, du\right)$$

$$= \frac{\Gamma\left(\dfrac{p+k}{2}\right)\Gamma\left(\dfrac{k}{2}+q\right)}{\Gamma\left(\dfrac{k}{2}\right)\Gamma\left(\dfrac{p+k}{2}+q-2\right)} \frac{4}{(p-2)(p-4}\, R^{2q}$$

$$\times \int_0^1 \frac{u^2}{R^2u+\|\theta\|^2}\, B\left(\frac{p}{2}-2, \frac{k}{2}+q, du\right).$$

As the function $u \to u^2/(R^2u+\|\theta\|^2)$ is convex, we have by the Jensen inequality

$$E_{R,\theta}\left[\frac{(U'U)^q}{X'X}\right]$$

$$\geqslant \frac{\Gamma\left(\dfrac{p+k}{2}\right)\Gamma\left(\dfrac{k}{2}+q\right)}{\Gamma\left(\dfrac{k}{2}\right)\Gamma\left(\dfrac{p+k}{2}+q-2\right)} \frac{4}{(p-2)(p-4)}$$

$$\times \frac{\left(\dfrac{p-4}{p+k+2q-4}\right)^2}{\dfrac{p-4}{p+k+2q-4}R^2+\|\theta\|^2} R^{2q}$$

$$= \frac{\Gamma\left(\dfrac{p+k}{2}\right)\Gamma\left(\dfrac{k}{2}+q\right)}{\Gamma\left(\dfrac{k}{2}\right)\Gamma\left(\dfrac{p+k}{2}+q-1\right)} \frac{2}{p-2} \frac{R^{2q}}{R^2+\dfrac{p+k+2q-4}{p-4}\|\theta\|^2}$$

$$= \frac{\Gamma\left(\dfrac{p+k}{2}\right)\Gamma\left(\dfrac{k}{2}+q\right)}{\Gamma\left(\dfrac{k}{2}\right)\Gamma\left(\dfrac{p+k}{2}+q\right)} \frac{p+k+2q-2}{p-2} \frac{R^{2q}}{R^2+\dfrac{p+k+2q-4}{p-4}\|\theta\|^2}.$$

Then, when we uncondition, we obtain the desired result. ▮

An upper bound for $E_\theta[(U'U)^q/X'X]$ is given by the following proposition.

PROPOSITION 4. *Let q be an integer such that $-k/2<q$. If $p\geqslant 6$ then, for any $\theta\in\mathbb{R}^p$, an upper bound for $E_\theta[(U'U)^q/X'X]$ is given by*

$$E_\theta\left[\frac{(U'U)^q}{X'X}\right] \leqslant \frac{\Gamma\left(\dfrac{p+k}{2}\right)\Gamma\left(\dfrac{k}{2}+q\right)}{\Gamma\left(\dfrac{k}{2}\right)\Gamma\left(\dfrac{p+k}{2}+q\right)} \frac{p+k+2q-2}{p-2}$$

$$\times E\left[\frac{R^{2q}}{R^2+\dfrac{(p+k+2q-2)(p-6)}{(p-2)^2}\|\theta\|^2}\right].$$

As above we gather the particular cases $q=0$, 1, and 2.

COROLLARY 3. *If $p\geqslant 6$ then, for any $\theta\in\mathbb{R}^p$, we have*

$$E_\theta\left[\frac{1}{X'X}\right] \leqslant \frac{p+k-2}{p-2} E\left[\frac{1}{R^2+\dfrac{(p+k-2)(p-6)}{(p-2)^2}\|\theta\|^2}\right]$$

$$E_\theta\left[\frac{U'U}{X'X}\right] \leqslant \frac{k}{p-2}\, E\left[\frac{R^2}{R^2+\dfrac{(p+k)(p-6)}{p-2)^2}\,\|\theta\|^2}\right]$$

$$E_\theta\left[\frac{(U'U)^2}{X'X}\right] \leqslant \frac{k(k+2)}{(p+k)(p-2)}\, E\left[\frac{R^4}{R^2+\dfrac{(p+k+2)(p-6)}{(p-2)^2}\,\|\theta\|^2}\right].$$

Proof of Proposition 4. Let $R \geqslant 0$ and $\theta \in \mathbb{R}^p$ be fixed. In the expression of $E_{R,\theta}[(U'U)^q/X'X]$ given by Proposition 2, the inner integral can be written as

$$\int_0^1 \frac{R^2u+\|\theta\|^2}{(R^2u+\|\theta\|^2)^2-4\,\|\theta\|^2\,R^2v}\, B\left(\frac{1}{2},\frac{p-1}{2},dv\right)$$

$$=\frac{p-2}{p-3}\int_0^1 \frac{(R^2u+\|\theta\|^2)(1-v)}{(R^2u+\|\theta\|^2)^2-4\,\|\theta\|^2\,R^2u+4\,\|\theta\|^2\,R^2u(1-v)}$$

$$\times B\left(\frac{1}{2},\frac{p-3}{2},dv\right)$$

$$\leqslant\frac{p-2}{p-3}\,\frac{(R^2u+\|\theta\|^2)\,\dfrac{p-3}{p-2}}{(R^2u+\|\theta\|^2)^2-4\,\|\theta\|^2\,R^2u+4\,\|\theta\|^2\,R^2u\,\dfrac{p-3}{p-2}}$$

(by Jensen inequality)

$$\leqslant\frac{1}{R^2u+\left(1-\dfrac{4}{p-2}\,\dfrac{R^2u}{R^2u+\|\theta\|^2}\right)\|\theta\|^2}$$

$$\leqslant\frac{1}{R^2u+\dfrac{p-6}{p-2}\,\|\theta\|^2}$$

since $p \geqslant 6$.

Therefore $E_{R,\theta}[(U'U)^q/X'X]$ is bounded from above by

$$\frac{\Gamma\left(\frac{p+k}{2}\right)\Gamma\left(\frac{k}{2}+q\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{p+k}{2}+q\right)} R^{2q} \int_0^1 \frac{1}{R^2 u+\frac{p-6}{p-2}\|\theta\|^2} B\left(\frac{p}{2}, \frac{k}{2}+q, du\right)$$

$$=\frac{\Gamma\left(\frac{p+k}{2}\right)\Gamma\left(\frac{k}{2}+q\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{p+k}{2}+q-1\right)} \frac{2}{p-2} R^{2q}$$

$$\times \int_0^1 \frac{u}{R^2 u+\left(1-\frac{4}{p-2}\right)\|\theta\|^2} B\left(\frac{p}{2}-1, \frac{k}{2}+q, du\right).$$

As the integrand is a concave function of u we have by Jensen inequality

$$E_{R,\theta}\left[\frac{(U'U)^q}{X'X}\right] \leqslant \frac{\Gamma\left(\frac{p+k}{2}\right)\Gamma\left(\frac{k}{2}+q\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{p+k}{2}+q-1\right)} \frac{2}{p-2}$$

$$\times R^{2q} \frac{\frac{p-2}{p+k+2q-2}}{\frac{p-2}{p+k+2q-2} R^2+\frac{p-6}{p-2}\|\theta\|^2}$$

$$=\frac{\Gamma\left(\frac{p+k}{2}\right)\Gamma\left(\frac{k}{2}+q\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{p+k}{2}+q\right)} \frac{p+k+2q-2}{p-2}$$

$$\times \frac{R^{2q}}{R^2+\frac{(p+k+2q-2)(p-6)}{(p-2)^2}\|\theta\|^2}.$$

Finally when we uncondition the desired result follows. ∎

Remark 1. All bounds given in this section are exact at $\theta=0$. This follows easily from Corollary 1 and expressions of the bounds.

4. BOUNDS FOR THE RISKS OF THE JAMES–STEIN AND THE ROBUST JAMES–STEIN ESTIMATORS

In this section, we give lower and upper bounds for the risks of the James–Stein estimator (1) and the robust James–Stein estimator (2). Their risks correspond respectively to the cases $\alpha = 0$ and $\alpha = 1$ in Proposition 1. See Casella and Hwang [5] for similar bounds in the normal case. Thus we have

$$\mathscr{R}(\theta, \delta^a_{\mathrm{JS}}) = E_0[X'X] + a^2 E_\theta \left[\frac{1}{X'X} \right] - 2a \, \frac{p-2}{k} \, E_\theta \left[\frac{U'U}{X'X} \right] \tag{3}$$

and

$$\mathscr{R}(\theta, \delta^a_{\mathrm{RJS}}) = E_0[X'X] + \left(a^2 - 2a \, \frac{p-2}{k+2} \right) E_\theta \left[\frac{(U'U)^2}{X'X} \right]. \tag{4}$$

It is clear, from formula 4, that domination of δ^a_{RJS} over X occurs when $0 < a < 2(p-2)/(k+2)$. In this context, bounds for $\mathscr{R}(\theta, \delta^a_{\mathrm{JS}})$ and $\mathscr{R}(\theta, \delta^a_{\mathrm{RJS}})$ are immediately deduced from the corollaries of Propositions 3 and 4.

PROPOSITION 5. *If $p \geqslant 6$ then, for any $\theta \in \mathbb{R}^p$, we have*

$$E_0[X'X] + a^2 \, \frac{p+k-2}{p-2} \, E \left[\frac{1}{R^2 + \frac{p+k-4}{p-4} \, \|\theta\|^2} \right]$$
$$- 2aE \left[\frac{R^2}{R^2 + \frac{(p+k)(p-6)}{(p-2)^2} \, \|\theta\|^2} \right]$$
$$\leqslant \mathscr{R}(\theta, \delta^a_{\mathrm{JS}})$$
$$\leqslant E_0[X'X] + a^2 \, \frac{p+k-2}{p-2} \, E \left[\frac{1}{R^2 + \frac{(p+k-2)(p-6)}{(p-2)^2} \, \|\theta\|^2} \right]$$
$$- 2aE \left[\frac{R^2}{R^2 + \frac{p+k-2}{p-4} \, \|\theta\|^2} \right]$$

and

$$E_0[X'X] + \left(a^2 - 2a\,\frac{p-2}{k+2}\right)\frac{k(k+2)}{(p+k)(p-2)}$$

$$\times E\left[\frac{R^4}{R^2 + \dfrac{(p+k+2)(p-6)}{(p-2)^2}\,\|\theta\|^2}\right]$$

$$\leqslant \mathscr{R}(\theta, \delta^a_{\text{RJS}})$$

$$\leqslant E_0[X'X] + \left(a^2 - 2a\,\frac{p-2}{k+2}\right)\frac{k(k+2)}{(p+k)(p-2)}$$

$$\times E\left[\frac{R^4}{R^2 + \dfrac{p+k}{p-4}\,\|\theta\|^2}\right].$$

Remark 2. All the bounds given above are exact at 0 since they are deduced from bounds of $E_\theta[(U'U)^q/X'X]$ which are also exact at 0. However, it is often desirable to have bounds in terms of moments of R^2. By applications of the Jensen inequality to the function $R^{2q}/(R^2+A)$, where q is a fixed integer and A is a fixed non-negative constant, it can be shown that

$$\frac{(E[R^{2q-2}])^2}{E[R^{2q-2}] + AE[R^{2q-4}]} \leqslant E\left[\frac{R^{2q}}{R^2+A}\right] \leqslant \frac{E[R^{2q}]\,E[R^{2q-2}]}{E[R^{2q}] + AE[R^{2q-2}]}.$$

Then, from the corollaries of Propositions 3 and 4, we deduce the following bounds when $p \geqslant 6$:

$$\frac{p+k-2}{p-2}\,\frac{\left(E\left[\dfrac{1}{R^2}\right]\right)^2}{E\left[\dfrac{1}{R^2}\right] + \dfrac{p+k-4}{p-4}\,\|\theta\|^2}$$

$$\leqslant E_\theta\left[\frac{1}{X'X}\right] \leqslant \frac{p+k-2}{p-2}\,\frac{E\left[\dfrac{1}{R^2}\right]}{1 + \dfrac{(p+k-2)(p-6)}{(p-2)^2}\,\|\theta\|^2\,E\left[\dfrac{1}{R^2}\right]}$$

$$\frac{k}{p-2}\,\frac{1}{1 + \dfrac{p+k-2}{p-4}\,\|\theta\|^2\,E\left[\dfrac{1}{R^2}\right]}$$

$$\leqslant E_\theta\left[\frac{U'U}{X'X}\right] \leqslant \frac{k}{p-2}\,\frac{E[R^2]}{E[R^2]+\frac{(p+k)(p-6)}{(p-2)^2}\,\|\theta\|^2}$$

$$\frac{k(k+2)}{(p+k)(p-2)}\,\frac{(E[R^2])^2}{E[R^2]+\frac{p+k}{p-4}\,\|\theta\|^2}$$

$$\leqslant E_\theta\left[\frac{(U'U)^2}{X'X}\right]$$

$$\leqslant \frac{k(k+2)}{(p+k)(p-2)}\,\frac{E[R^4]\,E[R^2]}{E[R^4]+\frac{(p+k+2)(p-6)}{(p-2)^2}\,\|\theta\|^2\,E[R^2]}.$$

It worth noting that these new bounds are also exact at 0. Then it is easy to derive bounds for the risks of δ_{JS} and δ_{JRS} in terms of moments of R^2 which are exact at 0 (but note that they are less sharp than those of Proposition 5).

Proposition 6. *If $p\geqslant 6$ then, for any $\theta\in\mathbb{R}^p$, we have*

$$E_0[X'X]+a^2\,\frac{p+k-2}{p-2}\,\frac{\left(E\left[\frac{1}{R^2}\right]\right)^2}{E\left[\frac{1}{R^2}\right]+\frac{p+k-4}{p-4}\,\|\theta\|^2}$$

$$-\,2a\,\frac{E[R^2]}{E[R^2]+\frac{(p+k)(p-6)}{(p-2)^2}\,\|\theta\|^2}$$

$$\leqslant \mathscr{R}(\theta,\delta^a_{\mathrm{JS}})$$

$$\leqslant E_0[X'X]+a^2\,\frac{p+k-2}{p-2}\,\frac{E\left[\frac{1}{R^2}\right]}{1+\frac{(p+k-2)(p-6)}{(p-2)^2}\,\|\theta\|^2\,E\left[\frac{1}{R^2}\right]}$$

$$-\,2a\,\frac{1}{1+\frac{p+k-2}{p-4}\,\|\theta\|^2\,E\left[\frac{1}{R^2}\right]}$$

and

$$E_0[X(X] + \left(a^2 - 2a\,\frac{p-2}{k+2}\right)\frac{k(k+2)}{(p+k)(p-2)}$$

$$\times \frac{E[R^4]\,E[R^2]}{E[R^4] + \dfrac{(p+k+2)(p-6)}{(p-2)^2}\,\|\theta\|^2\,E[R^2]}$$

$$\leqslant \mathscr{R}(\theta, \delta^a_{\mathrm{RJS}})$$

$$\leqslant E_0[X'X] + \left(a^2 - 2a\,\frac{p-2}{k+2}\right)\frac{k(k+2)}{(p+k)(p-2)}\,\frac{(E[R^2])^2}{E[R^2] + \dfrac{p+k}{p-4}\,\|\theta\|^2}.$$

5. DOMINATION OF THE JAMES–STEIN ESTIMATOR BY THE ROBUST JAMES–STEIN ESTIMATOR

From (4) it is easy to see that the optimal constant a for the risk of the robust James–Stein estimator is $(p-2)/(k+2)$ (that is, the constant for which the risk of δ_{RJS} is minimum).

Then the corresponding estimator $\delta^{\mathrm{opt}}_{\mathrm{RJS}}$ has risk

$$\mathscr{R}(\theta, \delta^{\mathrm{opt}}_{\mathrm{RJS}}) = E_0[X'X] - \left(\frac{p-2}{k+2}\right)^2 E_\theta\left[\frac{(U'U)^2}{X'X}\right]. \tag{5}$$

The main result of this section is Theorem 1 which yields a sufficient condition for the optimal robust James–Stein estimator to dominate any James–Stein estimator.

THEOREM 1. *The optimal robust James–Stein estimator $\delta^{\mathrm{opt}}_{\mathrm{RJS}}$ uniformly (in θ) dominates all the James–Stein estimators δ^a_{JS} provided, for any $\theta \in \mathbb{R}^p$,*

$$\frac{\left(E_\theta\left[\dfrac{U'U}{X'X}\right]\right)^2}{E_\theta\left[\dfrac{(U'U)^2}{X'X}\right] E_\theta\left[\dfrac{1}{X'X}\right]} < \left(\frac{k}{k+2}\right)^2. \tag{6}$$

Proof. According to (3), the optimal choice of the constant a leading to a minimum risk for the James–Stein estimator δ_{JS}^a depends on θ and equals

$$a(\theta) = \frac{p-2}{k} \frac{E_\theta\left[\dfrac{U'U}{X'X}\right]}{E_\theta\left[\dfrac{1}{X'X}\right]}.$$

Then the corresponding risk is

$$\mathscr{R}(\theta, \delta_{\mathrm{JS}}^{a(\theta)}) = E_0[X'X] - \left(\frac{p-2}{k}\right)^2 \frac{\left(E_\theta\left[\dfrac{U'U}{X'X}\right]\right)^2}{E_\theta\left[\dfrac{1}{X'X}\right]}.$$

Therefore, according to (5), the difference in risk between $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ and δ_{JS}^a is bounded from above by

$$\left(\frac{p-2}{k}\right)^2 \frac{\left(E_\theta\left[\dfrac{U'U}{X'X}\right]\right)^2}{E_\theta\left[\dfrac{1}{X'X}\right]} - \left(\frac{p-2}{k+2}\right)^2 E_\theta\left[\frac{(U'U)^2}{X'X}\right].$$

Thus $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ uniformly dominates all the estimators δ_{JS}^a (for any a) if this last quantity is negative, that is, if

$$\frac{\left(E_\theta\left[\dfrac{U'U}{X'X}\right]\right)^2}{E_\theta\left[\dfrac{(U'U)^2}{X'X}\right] E_\theta\left[\dfrac{1}{X'X}\right]} < \left(\frac{k}{k+2}\right)^2$$

which is the desired result. ∎

Condition (6) may be difficult to verify directly and a convenient way is to express an upper bound of the left-hand side of (6) using the bounds obtained in Section 3. This easily leads to be the following corollary.

COROLLARY 4. *For $p \geqslant 6$, a sufficient condition for which $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ dominates uniformly (in θ) all the estimators δ_{JS}^a is*

$$\frac{\left(E\left[\dfrac{R^2}{R^2+((p+k)(p-6))/((p-2)^2)\,\|\theta\|^2}\right]\right)^2}{E\left[\dfrac{R^4}{R^2+(p+k)/(p-4)\,\|\theta\|^2}\right]\,E\left[\dfrac{1}{R^2+(p+k-4)/(p-4)\,\|\theta\|^2}\right]} < \frac{k}{k+2}\,\frac{p+k-2}{p+k}. \tag{7}$$

However, it is interesting to first consider condition (7) for $\theta=0$ and for $\|\theta\|$ at infinity since there it reduces to simple conditions that are easy to check. These are the subject of the following corollaries.

COROLLARY 5. *The optimal James–Stein estimator at* $\theta=0$ *is* $\delta_{\mathrm{JS}}^{a_0}$ *where* $a_0=(p-2)/(p+k-2)(1/E[1/R^2]$. *Then the optimal robust estimator* $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ *dominates* $\delta_{\mathrm{JS}}^{a_0}$ *at* $\theta=0$ *if and only if*

$$E[R^2]\,E\left[\frac{1}{R^2}\right] > \frac{k+2}{k}\,\frac{p+k}{p+k-2}. \tag{8}$$

Proof. The result is a straightforward application of Corollary 4 for $\theta=0$. The "only if" part follows from the fact that the left-hand side of (6) equals the left-hand side of (7) at $\theta=0$. ∎

The dominance conditions of $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ over δ_{JS}^{a}, for $\|\theta\|$ at infinity, are deduced from (7) by dividing the numerator and denominator of the left-hand side of (7) by $\|\theta\|^4$.

COROLLARY 6. *For* $p\geqslant 7$, *the optimal robust James–Stein estimator* $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ *dominates all the estimators* δ_{JS}^{a} *for* $\|\theta\|$ *at infinity if*

$$\frac{(E[R^2])^2}{E[R^4]} < \frac{k}{k+2}\,\frac{(p-4)^2\,(p-6)^2}{(p-2)^4}\,\frac{p+k-2}{p+k-4}. \tag{9}$$

Remark 3. Straightforward calculations show that the right-hand side of (9) is less than 1 (for $p\geqslant 7$).

6. EXAMPLES AND COUNTEREXAMPLES

This section is devoted to examples (and counterexamples) of domination of the optimal robust James–Stein $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ estimator over the James–Stein estimators δ_{JS}^{a}.

6.1. *Condition at* $\theta = 0$

We start with three counterexamples for which domination of $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ over δ_{JS}^{a} cannot happen. Then we give two examples of such a domination.

1. If $(X, U)'$ is normal with the identity covariance matrix, then the square of the radius has a χ^2_{p+k} distribution and condition 8 becomes

$$\frac{p+k}{p+k-2} > \frac{k+2}{k} \frac{p+k}{p+k-2}$$

which is never satisfied. Of course, in this situation, as noted earlier, $\delta_{\mathrm{JS}}^{p-2}$ uniformly dominates $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$.

2. If $(X, U)'$ has a distribution which is uniform on the sphere $S_{R_0} = \{ y \in \mathbb{R}^{p+k} / \|y\| = R_0 \}$ of radius R_0 and centered at 0, then condition (8) becomes

$$1 > \frac{k+2}{k} \frac{p+k}{p+k-2}$$

which is not satisfied for any k and p.

3. If $(X, U)'$ has a distribution which is uniform on the ball $B_{R_0} = \{ y \in \mathbb{R}^{p+k} / \|y\| \leqslant R_0 \}$ of radius R_0 and is centered at 0, then condition (8) becomes

$$\frac{p+k}{p+k-2} > \frac{k+2}{k}$$

which is not satisfied for any k and $p \geqslant 3$.

4. If the square of the radius has a gamma density $\gamma(\alpha, \beta)$ where $\alpha > 0$ and $\beta > 0$, then condition (8) is equivalent to

$$1 < \alpha < \frac{(k+2)(p+k}{2(p+2k)}.$$

Hence this is an example of domination, at $\theta = 0$, of $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ over δ_{JS}^{a} (note that $(k+2)(p+k)/(2p+4k) > 1$ for $k > 2$ and any value of p). Note that this condition is independent of β.

5. Suppose that $(X, U)'$ has a Student distribution with m degrees of freedom. Straightforward calculations show that the density of the radial distribution is given by

$$\frac{2\Gamma\left(\frac{m+p+k}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{p+k}{2}\right)m^{(p+k)/2}}\left(1+\frac{R^2}{m}\right)^{-((m+p+k)/2)} R^{p+k-1}.$$

hence $E[R^2]=m(p+k)/(m-2)$, if $m>2$, and $E[1/R^2]=1/(p+k-2)$, if $p+k>2$. Thus condition (8) is equivalent to $m/(m-2)>(k+2)/k$, that is, to $m<k+2$.

Remark 4. It follows easily from Corollary 5 that, in the general case of variance mixture of normals, condition 8 becomes

$$E[\sigma^2]\,E\left[\frac{1}{\sigma^2}\right]>\frac{k+2}{k}.$$

6.2. *Condition at Infinity for* $\|\theta\|$

We will see that domination of $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ at infinity for $\|\theta\|$ is obtained in the cases of the distributions considered in 6.1(4) and 6.1(5).

1. When the square of the radius has a gamma density $\gamma(\alpha, \beta)$, condition (9) is equivalent to

$$\frac{\alpha}{\alpha+1}<\frac{k}{k+2}\,\frac{(p-4)^2\,(p-6)^2}{(p-2)^4}\,\frac{p+k-2}{p+k-4}.$$

It is easy to check that the left-hand side of this inequality increases from 0 to 1 when α goes from 0 to infinity. Then domination of $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ is assured for small values of α.

More precisely it can be shown that the corresponding range of values of α is the interval $]0, \lambda/(1-\lambda)[$ where λ denotes the right-hand side of the previous inequality.

Note that, since the upper bound of this interval is an increasing function of $\lambda\in\,]0,1[$ and its value at $\lambda=1$ equals infinity, there exists a $\lambda\in\,]0,1[$ such that this bound is greater than 1. As a consequence we can find a range of values of α for which the condition of domination of $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ is satisfied at 0 and infinity for $\|\theta\|$.

2. In the context of 6.1 (5) we have $E[R^4]=(m^2/(m-2)(m-4))((p+k)(p+k+2))$ and then, when $m>4$, condition (9) is equivalent to

$$\frac{m-4}{m-2}<\frac{k}{k+2}\,\frac{(p-4)^2\,(p-6)^2}{(p-2)^4}\,\frac{p+k+2}{p+k}\,\frac{p+k-2}{p+k-4}.$$

It is clear that, for $p\geqslant 7$ and for any k,

$$\lambda=\frac{(p-4)^2\,(p-6)^2}{(p-2)^4}\,\frac{p+k+2}{p+k}\,\frac{p+k-2}{p+k-4}<\frac{(p-4)^2\,(p-6)^2}{(p-2)^4}<1.$$

Hence condition (9) is satisfied as soon as

$$m<\frac{2}{(1-\lambda)\,k+2}\,(k+2)+2.$$

Lastly it is clear, since $0<\lambda<1$, that

$$k+2<\frac{2}{(1-\lambda)\,k+2}\,(k+2)+2.$$

Thus the domination condition at 0 implies the domination condition at infinity whenever the fourth moment exists.

6.3. *Uniform Domination in θ*

It is of course more difficult to obtain uniform domination of $\delta^{\text{opt}}_{\text{RJS}}$ over δ^{a}_{JS} uniformly in $\theta\in\mathbb{R}^p$. The case where the radial distribution is concentrated on two points is a first example. Of course, this example may not be a reasonable model for data likely to be encountered in practice. However, it is the simplest class of distributions beyond the class where the radial distributions is degenerate. It is therefore particularly interesting that the paradox occurs for this simple example.

Thus we assume that

$$P[R^2=\lambda]=\alpha=1-P[R^2=H\lambda]$$

for fixed $\lambda>0$ and $H>0$ and we have the following result.

Proposition 7. *For $\alpha=\frac{1}{2}$, any $\lambda>0$, and sufficiently large p and k, there exists H_0 such that for all $H>H_0$ the robust estimator $\delta^{\text{opt}}_{\text{RJS}}$ uniformly and simultaneously dominates all estimators δ^{a}_{JS}.*

Proof. Condition 6 upon setting $A=(p+k)(p-6)/(p-2)^2$, $B=(p+k)/(p-4)$, and $C=(p+k-4)/(p-4)$ becomes

$$\frac{\left(\dfrac{\lambda}{\lambda+A\,\|\theta\|^2}+\dfrac{H\lambda}{H\lambda+A\,\|\theta\|^2}\right)^2}{\left(\dfrac{\lambda^2}{\lambda+B\,\|\theta\|^2}+\dfrac{H^2\lambda^2}{H\lambda+B\,\|\theta\|^2}\right)\left(\dfrac{1}{\lambda+C\,\|\theta\|^2}+\dfrac{1}{H\lambda+C\,\|\theta\|^2}\right)}\leqslant\frac{k(p+k-2)}{(k+2)(p+k)}$$

or equivalently

$$\left\{\frac{(\lambda+B\,\|\theta\|^2)(H\lambda+B\,\|\theta\|^2)(\lambda+C\,\|\theta\|^2)(H\lambda+C\,\|\theta\|^2)}{[(\lambda+A\,\|\theta\|^2)(H\lambda+A\,\|\theta\|^2)]^2}\right\}$$
$$\times\left\{\frac{[\lambda(H\lambda+A\,\|\theta\|^2)+H\lambda(\lambda+A\,\|\theta\|^2)]^2}{[\lambda^2(H\lambda+B\,\|\theta\|^2)+H^2\lambda^2(\lambda+B\,\|\theta\|^2)][(H\lambda+C\,\|\theta\|^2)+(\lambda+C\,\|\theta\|^2)]}\right\}$$
$$\leqslant\frac{k(p+k-2)}{(k+2)(p+k)}. \tag{10}$$

For $p\geqslant 4$ and $p+k\geqslant 6$, $A\leqslant C\leqslant B$ and the first term in brackets on the left-hand side of (10) is bounded above by

$$\frac{B^2C^2}{A^4}=\left(\frac{p+k-4}{p+k}\right)^2\left(\frac{p-2}{p-4}\right)^4\left(\frac{p-2}{p-6}\right)^4. \tag{11}$$

The second term in brackets is bounded above (for $A\leqslant C\leqslant B$) by replacing A with C and B with C. Hence by letting $X=\lambda/\|\theta\|^2$ condition (10) will be satisfied, provided

$$\sup_X G_H(X)\leqslant\left(\frac{k}{k+2}\right)\left(\frac{p+k}{p+k-4}\right)\left(\frac{p+k-2}{p+k-4}\right)\left(\frac{p-6}{p-2}\right)^2\left(\frac{p-4}{p-2}\right)^4 \tag{12}$$

when

$$G_H(X)=\frac{[(HX+C)+H(X+C)]^2}{[HX+C+H^2(X+C)][(HX+C)+(X+C)]}. \tag{13}$$

The right-hand side of (12) approaches 1 as p and k get large. In fact, for any $k>2$, p may be chosen sufficiently large so that it is strictly greater than $1/2+\varepsilon$ for any $1/2>\varepsilon>0$. It therefore suffices to show that we may choose H_0 so that, for $H>H_0$,

$$\sup_X G(X)<\tfrac{1}{2}+\varepsilon. \tag{14}$$

A detailed calculation (checked with Maple) gives the derivative of $G_H(X)$ as

$$\frac{[-2HC^2(1-4H+6H^2-4H^3+H^4)(1+X)]}{Q(X)},$$

where $Q(X)$ is a positive fourth degree polynomial. Hence for H sufficiently large $G(X)$ is decreasing, and its supremum occurs at $X=0$. Therefore it suffices to choose H_0 so that, for $H>H_0$, $G(X)$ is decreasing and

$$G_H(0)=\frac{1}{2}\,\frac{(1+H)^2}{1+H^2}>\frac{1}{2}+\varepsilon.$$

But $\lim_{H\to 0} G_H(0)=\frac{1}{2}$ and hence such H_0 exists. The theorem follows.

We note that a similar result can be proven for any fixed $\alpha>0$. Additionally, numerical calculations indicate that p and H need not be nearly as large as the proof of the theorem indicates for domination to hold. ∎

Remark 5. From the point of view of practical modelling of multivariate data, the t-distribution is perhaps, next to the normal distribution, the most important and widely used. It is therefore interesting and potentially important that the paradox holds when the dimension k of the residual vector is sufficiently large, whatever the value of p ($\geqslant 7$) and the degree of freedom m ($\geqslant 5$).

Precisely, we have the following result:

PROPOSITION 8. *Assume that $(X, U)'$ has a Student distribution with m degrees of freedom. Then, for any $p\geqslant 7$ and any $m\geqslant 5$, there exists k_0 such that, for any $k\geqslant k_0$, the robust estimator $\delta^{\text{opt}}_{\text{RJS}}$ uniformly and simultaneously dominates all estimators δ^a_{JS}.*

Proof. As the density of the radius of the t_m distribution is proportional to $(1+R^2/m)^{-(p+k+m)/2}\,R^{p+k-1}$, the left-hand side of condition (7) equals, by letting

$$A=\frac{(p+k)(p-6)}{(p-2)^2},\qquad B=\frac{p+k}{p-4},\qquad \text{and}\qquad C=\frac{p+k-4}{p-4},$$

$$r=\frac{\left(\int_0^\infty \frac{R^{p+k+1}}{R^2+A\,\|\theta\|^2}\left(1+\frac{R^2}{m}\right)^{-(p+k+m)/2} dr\right)^2}{\left[\int_0^\infty \frac{R^{p+k-1}}{R^2+C\,\|\theta\|^2}\left(1+\frac{R^2}{m}\right)^{-(p+k+m)/2} dr \times\int_0^\infty \frac{R^{p+k+3}}{R^2+B\,\|\theta\|^2}\left(1+\frac{R^2}{m}\right)^{-(p+k+m)/2} dr\right]}.$$

Through the change of variable $R=(m(1-t)/t)^{1/2}$ the quantity r can be written, after obvious simplifications, as

$$r=\frac{\left(\int_0^1 \frac{((1-t)/t)^{(p+k+1)/2}}{m((1-t)/t)+A\,\|\theta\|^2}\,t^{(p+k+m)/2}\,\frac{t^{-3/2}}{(1-t)^{1/2}}\,dt\right)^2}{\left[\int_0^1 \frac{((1-t)/t)^{(p+k-1)/2}}{m((1-t)/t)+C\,\|\theta\|^2}\,t^{(p+k+m)/2}\,\frac{t^{-3/2}}{(1-t)^{1/2}}\,dt \times \int_0^1 \frac{((1-t)/t)^{(p+k+3)/2}}{m((1-t)/t)+B\,\|\theta\|^2}\,t^{(p+k+m)/2}\,\frac{t^{-3/2}}{(1-t)^{1/2}}\,dt\right]}$$

$$=\frac{\left(\int_0^1 \frac{1}{1-t(1-A(\|\theta\|^2/m))}\,t^{(m-2)/2}(1-t)^{(p+k)/2}\,dt\right)^2}{\left[\int_0^1 \frac{1}{1-t(1-C(\|\theta\|^2/m))}\,t^{m/2}(1-t)^{(p+k-2)/2}\,dt \times \int_0^1 \frac{1}{1-t(1-B(\|\theta\|^2/m))}\,t^{(m-4)/2}(1-t)^{(p+k+2)/2}\,dt\right]}$$

$$=\frac{\frac{m-2}{m}\,\frac{p+k}{p+k+2}\left(\int_0^1 \frac{1}{1-t(1-A(\|\theta\|^2/m))}\,B\left(\frac{m}{2},\frac{p+k+2}{2},dt\right)\right)^2}{\left[\int_0^1 \frac{1}{1-t(1-C(\|\theta\|^2/m))}\,B\left(\frac{m+2}{2},\frac{p+k}{2},dt\right) \times \int_0^1 \frac{1}{1-t(1-B(\|\theta\|^2/m))}\,B\left(\frac{m-2}{2},\frac{p+k+4}{2},dt\right)\right]}.$$

Convexity, for fixed z, of $(1-zt)^{-1}$ and Jensen's inequality for the integrals of the denominator give

$$r\leqslant\frac{m-2}{m}\,\frac{p+k}{p+k+2}$$

$$\times\frac{\left[\int_0^1 \frac{1}{1-t(1-(A/m)\,\|\theta\|^2)}\,B\left(\frac{m}{2},\frac{p+k+2}{2},dt\right)\right]^2}{\left[\frac{1}{1-((m-2)/(p+k+m+2))(1-(B/m)\,\|\theta\|^2)} \times\frac{1}{1-((m+2)/(p+k+m+2))(1-(C/m)\,\|\theta\|^2)}\right]}$$

$$= \frac{m-2}{m} \frac{p+k}{p+k+2}$$

$$\times \int_0^1 \frac{\frac{m-2}{p+k+m+2} B\left[\frac{p+k+4}{(m-2)\,B} + \frac{\|\theta\|^2}{m}\right]}{tA\left[\frac{1-t}{tA} + \frac{\|\theta\|^2}{m}\right]} B\left(\frac{m}{2}, \frac{p+k+2}{2}, dt\right)$$

$$\times \int_0^1 \frac{\frac{m+2}{p+k+m+2} C\left[\frac{p+k}{(m+2)\,C} + \frac{\|\theta\|^2}{m}\right]}{tA\left[\frac{1-t}{tA} + \frac{\|\theta\|^2}{m}\right]} B\left(\frac{m}{2}, \frac{p+k+2}{2}, dt\right)$$

$$= \frac{m-2}{m} \frac{m-2}{p+k+m+2} \frac{m+2}{p+k+m+2}$$

$$\times \frac{p+k-4}{p+k+2} \frac{(p-2)^4}{(p+k)^2\,(p-6)^2}$$

$$\times \int_0^1 \frac{1}{t} \frac{\frac{(p+k-4)(p-4)}{(m-2)(p+k)} + \frac{\|\theta\|^2}{m}}{\frac{1-t}{t} \frac{(p+k)(p-6)}{(p-2)^2} + \frac{\|\theta\|^2}{m}} B\left(\frac{m}{2}, \frac{p+k+2}{2}, dt\right)$$

$$\times \int_0^1 \frac{1}{t} \frac{\frac{(p+k)(p-4)}{(m+2)(p+k-4)} + \frac{\|\theta\|^2}{m}}{\frac{1-t}{t} \frac{(p+k)(p-6)}{(p-2)^2} + \frac{\|\theta\|^2}{m}} B\left(\frac{m}{2}, \frac{p+k+2}{2}, dt\right)$$

where we have replaced A, B, and C by their values in the last expression.

Applying Lemma 3 to each integral in the last expression yields

$$r \leqslant \frac{m-2}{m} \frac{m-2}{p+k+m+2} \frac{m+2}{p+k+m+2} \frac{p+k-4}{p+k+2} \frac{(p-2)^4}{(p+k)^2\,(p-6)^2}$$

$$\times \left[\frac{B\left(\frac{m}{2}-1, \frac{p+k+2}{2}\right)}{B\left(\frac{m}{2}, \frac{p+k+2}{2}\right)} \right.$$

$$+\frac{(p+k-4)(p-4)(p-2)^2}{(m-2)(p+k)^2\,(p-6)}\frac{B\left(\frac{m}{2},\frac{p+k}{2}\right)}{B\left(\frac{m}{2},\frac{p+k+2}{2}\right)}\Bigg]$$

$$\times\left[\frac{B\left(\frac{m}{2}-1,\frac{p+k+2}{2}\right)}{B\left(\frac{m}{2},\frac{p+k+2}{2}\right)}\right.$$

$$\left.+\frac{(p-4)(p-2)^2}{(p+k-4)(p-6)(m+2)}\frac{B\left(\frac{m}{2},\frac{p+k}{2}\right)}{B\left(\frac{m}{2},\frac{p+k+2}{2}\right)}\right]. \tag{15}$$

Noting that $B((m/2)-1, (p+k+2)/2)/B(m/2, (p+k+2)/2)=(p+k+m)/(m-2)$ and $B(m/2, (p+k+2)/2)/B(m/2, (p+k)/2)=(p+k+m)/(p+k)$, inequality (15) can be rewritten as

$$r\leqslant\frac{m-2}{m}\left(\frac{p+k+m}{p+k+m+2}\right)^2\frac{p+k-4}{p+k+2}\frac{(p-2)^4}{(p-6)^2}\frac{1}{(p+k)^2}$$
$$\times\left[1+\frac{1}{(p+k^2}\frac{p+k-4}{p+k}\frac{(p-4)(p-2)^2}{p-6}\right]$$
$$\times\left[\frac{m+2}{m-2}+\frac{1}{(p+k-4)(p+k)}\frac{(p-4)(p-2)^2}{p-6}\right]. \tag{16}$$

Note that the right-hand side of (16) goes to zero when p and m are fixed and k goes to infinity. Now the right-hand side of (7) goes to 1 under these conditions. Hence the theorem follows. ∎

Remark 6. Proposition 8 does not give an explicit value of k_0 such that, for $k\geqslant k_0$, we have uniform domination of $\delta_{\mathrm{RJS}}^{\mathrm{opt}}$ over all δ_{JS}^{a}. Comparing with Maple the right-hand side of (16) with the right-hand side of (7) yields Table I which gives values of k_0 corresponding to fixed values of m ($\geqslant 5$) and p ($\geqslant 7$).

7. CONCLUSION

Since Stein's fundamental 1956 article [10] many papers have studied properties of Stein like shrinkage estimators. In 1961 James and Stein [8]

TABLE I

Minimum Values of k for Which $\delta^{\text{opt}}_{\text{RJS}}$ Dominates All the δ^{a}_{JS}

$m \backslash p$	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5	22	13	11	10	9	9	10	10	10	11	11	11	12	12
6	21	13	10	10	9	9	9	10	10	10	11	11	12	12
7	21	13	10	10	9	9	9	10	10	10	11	11	11	12
8	21	13	10	9	9	9	9	9	10	10	10	11	11	12
9	20	12	10	9	9	9	9	9	10	10	10	11	11	12
10	10	12	10	9	9	9	9	9	10	10	10	11	11	12
11	10	12	10	9	9	9	9	9	10	10	10	11	11	11
12	10	12	10	9	9	9	9	9	10	10	10	11	11	11
13	10	12	10	9	9	9	9	9	10	10	10	11	11	11
14	10	12	10	9	9	9	9	9	10	10	10	11	11	11
15	10	12	10	9	9	9	9	9	10	10	10	11	11	11
16	10	12	10	9	9	9	9	9	10	10	10	11	11	11
17	10	12	10	9	9	9	9	9	10	10	10	11	11	11
18	10	12	10	9	9	9	9	9	10	10	10	11	11	11

showed that δ^{a}_{JS} dominates the usual estimator, X, in the $\mathcal{N}(\theta, \sigma^2 I)$ case for $0 < a < 2(p-2)\,\sigma^2$ when σ^2 is known and also that $\delta^{\text{opt}}_{\text{RJS}}$ improves X when σ^2 is unknown. Several authors have studied improvements in the nonnormal case with the broadest developments occurring for spherically symmetric distributions. Typically, δ^{a}_{JS} has been shown to dominate X or a range of values depending on the distribution (see Brandwein and Strawderman [1] for an extensive bibliography). Cellier *et al.* [7] showed hat $\delta^{\text{opt}}_{\text{RJS}}$ dominates X for all spherically symmetric distributions simultaneously.

A key feature of this last result is that it holds whether the variance of the underlying distribution is known or unknown (provided a residual vector is present). This extremely strong distributional robustness property of $\delta^{\text{opt}}_{\text{RJS}}$ makes it a very attractive alternative to δ^{a}_{JS} even in the known variance case.

It was this distributional robustness property that led us to the comparison of the two estimators and to the paradox studied in this paper —namely that substituting an estimated value $U'U/(k+2)$ for the known value σ^2 can lead to an improved procedure. In particular, it was most interesting that the paradox occurs for the multivariate-t model, perhaps the most important alternative to the normal model, as soon as the dimension of the residual vector is sufficiently large.

That the paradox can occur at all is of course due to the fact that, as soon as one departs from the normal model, even in the known variance case, X is no longer sufficient and $(X, U'U)$ is a minimal sufficient statistic with $U'U$ being ancillary. Of course this does not explain why the paradox occurs, only that it is not impossible.

The James–Stein estimator is not the only, and in practice is usually not the best, estimator to dominate X. Other alternatives such as the Lindley–Smith estimator $X-(a/\|X-\bar{X}1\|^2)(X-\bar{X}1)$ (see [9]) offer substantial improvements in certain portions of the parameter space and effectively utilize prior information about θ. A very general class of such estimators, $X+a\sigma^2 g(X)$ where $a\,\|g(X)\|^2+2\,\mathrm{div}\,g(X)<0$, was shown in Stein [11] to dominate X in the $\mathcal{N}(\theta, \sigma^2 I)$ case with σ^2 known.

Again, this result has been extended to the general spherically symmetric case by Brandwein and Strawderman [2]. Cellier and Fourdrinier [6] showed that the distributionnaly robust version $X+(U'U/(k+2))\,g(X)$ dominates X for all spherically symmetric distribution simultaneously.

We have not studied the paradox in any detail for this broader class but it seems likely that it persists. In particular, for estimators of the Lindley–Smith type or others which shrink to a subspace, our calculations do apply and show that the robust version will dominate for the t-distribution if the dimension of the residual vector is large.

We believe that the results add substantially to the attractiveness of the robust forms of general shrinkage estimators and have the potential to be more than just an interesting paradox.

APPENDIX

The first two lemmas deal with expectations conditioned on the radius of a spherically symmetric distribution in $\mathbb{R}^p \times \mathbb{R}^k$ centered at $(\theta, 0)$ where $\theta \in \mathbb{R}^p$. These expectations reduce to integrals with respect to the uniform distribution $U_{R,\theta}$ on the sphere $S_{R,\theta}=\{y=(x,u)\in\mathbb{R}^p\times\mathbb{R}^k/(\|x-\theta\|^2+\|u\|^2)^{1/2}=R\}$. If $E_{R,\theta}[\psi]$ is the expectation of some function ψ with respect to $U_{R,\theta}$, the expectation with respect to the entire distribution is given by $E_\theta[\psi]=E[E_{R,\theta}[\psi]]$ where E is the expectation with respect to the distribution of the radius. For $\theta=0$, we denote by S_R, U_R and E_R the respective expressions $S_{R,0}$, $U_{R,0}$, and $E_{R,0}$.

When the spherical distribution has a density with respect to the Lebesgue measure, it is necessarily of the form $f(\|x-\theta\|^2+\|u\|^2)$ for some function f. Then the radius has density $R \to \sigma_{p+k} f(R^2)\,R^{p+k-1}$ where $\sigma_{p+k}=2\pi^{p+k}/\Gamma((p+k)/2)$. Therefore the expectation of any function ψ can be written as

$$E_\theta[\psi]=\int_0^\infty \left[\int_{S_{R,\theta}} \psi(y)\, U_{R,\gamma}(dy)\right] f(R)\, dR.$$

Note that for a vector $y=(x, u)\in S_{R,\theta}$, we have $x=\pi(y)$ and $\|u\|^2=R^2-\|\pi(y)-\theta\|^2$ where π is the orthogonal projector from $\mathbb{R}^p\times\mathbb{R}^k$ onto $\mathbb{R}^p$. Under $U_{R,\theta}$, the distribution $\pi(U_{R,\theta})$ of this projector has a density with respect to the Lebesgue measure on $\mathbb{R}^p$ given by $x\to C_R^{p,k}(R^2-\|x-\theta\|^2)^{k/2-1}\mathbf{1}_{B_{R,\theta}}(x)$ where $C_R^{p,k}=(\Gamma((p+k)/2)\,R^{2-p-k}/(\Gamma(k/2)\,\pi^{p/2})$ and $\mathbf{1}_{B_{R,\theta}}$ is the indicator function of the ball $B_{R,\theta}=\{x\in\mathbb{R}^p/\|x-\theta\|\leqslant R\}$ of radius R centered at θ in $\mathbb{R}^p$.

According to the above, as a spherically symmetric distribution on $\mathbb{R}^p$ around θ, the radius of $\pi(U_{R,\theta})$ has density

$$\begin{aligned} r &\to \sigma_p C_R^{p,k}(R^2-r^2)^{k/2-1}\,\mathbf{1}_{]0,R[}(r)\,r^{p-1} \\ &= \frac{2R^{2-p-k}}{B\left(\frac{p}{2},\frac{k}{2}\right)}\,r^{p-1}(R^2-r^2)^{k/2-1}\mathbf{1}_{]0,R[}(r). \end{aligned}$$

We repeatedly use the fact that any such projection onto a space of dimension greater than 0 and less than $p+k$ is spherically symmetric with a density. Then we also often make use of its radial density.

LEMMA 1. *Assume $p\geqslant 2$. If g and h are two measurable real valued functions then, for any $R>0$ and any $\theta\in\mathbb{R}^p$,*

$$\begin{aligned} &E_{R,\theta}[g(X'X)\cdot h(U'U)] \\ &\quad=\int_0^1 h(R^2(1-u))\int_0^1\frac{1}{2}\,[g(R^2u+\|\theta\|^2-2\,\|\theta\|\,Ru^{1/2}v^{1/2}) \\ &\quad\quad+g(R^2u+\|\theta\|^2+2\,\|\theta\|\,Ru^{1/2}v^{1/2})] \\ &\quad\quad\times B\left(\frac{1}{2},\frac{p-1}{2},dv\right)B\left(\frac{p}{2},\frac{k}{2},du\right), \end{aligned}$$

provided these expectations exist.

Proof. Let $R\geqslant 0$ and $\theta\in\mathbb{R}^p$ be fixed. Then

$$\begin{aligned} &E_{R,\theta}[g(X'X)\,h(U'U)] \\ &\quad=\int_{S_{R,\theta}} h(\|y-\pi(y)\|^2)\;g(\|\pi(y)\|^2)\;U_{R,\theta}(dy) \\ &\quad=\int_{S_{R,\theta}} h(R^2-\|\pi(y)-\theta\|^2)\;g(\|\pi(y)\|^2)\;U_{R,\theta}(dy) \end{aligned}$$

$$= \int_{B_{R,\theta}} h(R^2 - \|x-\theta\|^2)\, g(\|x\|^2)\, \pi(U_{R,\theta})(dx)$$

$$= \frac{2R^{2-p-k}}{B\left(\frac{p}{2}, \frac{k}{2}\right)} \int_0^R h(R^2 - r^2) \left[\int_{S_{r,\theta}} g(\|x\|^2)\, U_{r,\theta}(dx) \right] r^{p-1} (R^2 - r^2)^{k/2-1}\, dr$$

using the density of the radius of the distribution $\pi(U_{R,\theta})$. Then, with the change of variable $r = R\sqrt{u}$, we obtain

$$E_{R,\theta}[g(X'X)\, h(U'U)]$$
$$= \int_0^1 h(R^2(1-u)) \left[\int_{S_{R\sqrt{u},\theta}} g(\|x\|^2)\, U_{R\sqrt{u},\theta}(dx) \right] B\left(\frac{p}{2}, \frac{k}{2}, du\right).$$

The rest of the calculation depends on the evaluation of the innermost integral of the last expression. Indeed, for $r \geqslant 0$ fixed, we have

$$\int_{S_{r,\theta}} g(\|x\|^2)\, u_{r,\gamma}(dx)$$
$$= \int_{\Delta_\theta} \left[\int_{\Delta_\theta^\perp} g(\|u\|^2 + \|t\|^2)\, U_{r,\theta}(du/\pi_\theta = t) \right] \pi_\theta(U_{r,\theta})(dt)$$

where Δ_θ is the one-dimensional linear subspace spanned by θ, $\Delta_\theta^\perp$ is its $(p-1)$-dimensional orthogonal subspace, π_θ is the orthogonal projector onto Δ_θ, and $U_{r,\theta}(\cdot\,, /\pi_\theta = t)$ is the conditional probability of $U_{r,\theta}$ given $\pi_\theta = t$.

In the right-hand side of the last equality, as $U_{r,\theta}(\cdot/\pi_\theta = t) = U_{\sqrt{r^2 - \|t-\theta\|^2}} \otimes \delta_t$ (where δ_t is the Dirac measure in Δ_θ at t), the innermost integrals equals

$$\int S_{\sqrt{r^2 - \|t-\theta\|^2}}\, g(\|u\|^2 + \|t\|^2)\, U_{\sqrt{r^2 - \|t-\theta\|^2}}(du) = g(r^2 - \|t-\theta\|^2 + \|t\|^2).$$

Hence using the density of $\pi_\theta(U_{r,\theta})$, we obtain

$$\int_{S_{r,\theta}} g(\|x\|^2)\, U_{r,\theta}(dx)$$
$$= \frac{2r^{2-p}}{B\left(\frac{1}{2}, \frac{p-1}{2}\right)} \int_0^r \left[\int_{S_{\tau,\theta}} g(r^2 - \|v-\theta\|^2 + \|v\|^2)\, U_{\tau,\theta}(dv) \right]$$
$$\times (r^2 - \tau^2)^{((p-1)/2)-1}\, d\tau.$$

Since the sphere $S_{\tau,\theta}$ is one-dimensional, then $U_{\tau,\theta} = \frac{1}{2}(\delta_{\theta - \tau v} + \delta_{\theta + \tau v})$ where $v = \theta / \|\theta\|$. Therefore

$$\int_{S_{r,\theta}} g(\|x\|^2)\, U_{r,\theta}(dx)$$

$$= \frac{r^{2-p}}{\Gamma\left(\frac{p-1}{2}, \frac{1}{2}\right)} \int_0^r [\,g(r^2 - \tau^2 + \|\theta - \tau v\|^2) + g(r^2 - \tau^2 + \|\theta + \tau v\|^2)]$$

$$\times (r^2 - \tau^2)^{((p-1)/2) - 1}\, d\tau$$

$$= \int_0^1 \frac{1}{2} [\,g(r^2 + \|\theta\|^2 - 2\,\|\theta\|\, r v^{1/2}) + g(r^2 + \|\theta\|^2 + 2\,\|\theta\|\, r v^{1/2})]$$

$$\times B\left(\frac{1}{2}, \frac{p-1}{2}, dv\right)$$

after expanding the variable in g and using the change of variable $\tau = r\sqrt{v}$.

Replacing this integral in the expression of $E_{R,\theta}[\,g(X'X)\, h(U'U)]$ found above gives the desired result. ∎

LEMMA 2. *If g is a vector-valued function and h is a real valued function then, for any $\theta \in \mathbb{R}^p$, provided these expectations exist, we have*

$$E_\theta[h(U'U)(X - \theta)'\, g(X)] = -E_\theta\left[\frac{H(U'U)}{(U'U)^{k/2-1}} \operatorname{div} g(X)\right],$$

where H is the indefinite integral, vanishing at 0, *of the function $t \to -\frac{1}{2} h(t)\, t^{k/2-1}$.*

Proof. Conditionally on the radius R, we have

$$E_{R,\theta}[h(U'U)(X - \theta)'\, g(X)]$$

$$= C_R^{p,k} \int_{B_{R,\theta}} h(R^2 - \|x - \theta\|^2)(x - \theta)'\, g(x)(R^2 - \|x - \theta\|^2)^{k/2-1}\, dx$$

$$= C_R^{p,k} \int_{B_{R,\theta}} (\nabla H(R^2 - \|x - \theta\|^2))'\, g(x)\, dx$$

since

$$\nabla H(R^2 - \|x - \theta\|^2) = -2H'(R^2 - \|x - \theta\|^2)(x - \theta)$$

$$= h(R^2 - \|x - \theta\|^2)(R^2 - \|x - \theta\|^2)^{k/2-1}\,(x - \theta).$$

Then, by the divergence formula,

$$E_{R,\theta}[h(U'U)(X-\theta)'\,g(X)]$$
$$= C_R^{p,k} \int_{B_{R,\theta}} \operatorname{div}\,(H(R^2-\|x-\theta\|^2)\,g(x))\,dx$$
$$- C_R^{p,k} \int_{B_{R,\theta}} H(R^2-\|x-\theta\|^2)\operatorname{div}\,g(x)\,dx.$$

Now, if $\sigma_{R,\theta}$ denotes the area measure on the sphere $S_{R,\theta}$, the divergence theorem insures that the first integral equals

$$C_R^{p,k} \int_{S_{R,\theta}} (H(R^2-\|x-\theta\|^2)\,g(x))'\,\frac{x-\theta}{\|x-\theta\|}\,\sigma_{R,\theta}(dx)$$

and is null since, for $x \in S_{R,\theta}$, $R^2-\|x-\theta\|^2=0$ and $H(0)=0$.

Hence, in terms of expectation, we have

$$E_{R,\theta}[h(U'U)(X-\theta)'\,g(X)]$$
$$= -C_R^{p,k} \int_{B_{R,\theta}} \frac{H(R^2-\|x-\theta\|^2)}{(R^2-\|x-\theta\|^2)^{k/2-1}} \operatorname{div}\,g(x)(R^2-\|x-\theta\|^2)^{k/2-1}\,dx$$
$$= -E_{R,\theta}\left[\frac{H(U'U)}{(U'U)^{k/2-1}}\operatorname{div}\,g(X)\right]$$

which is, when we uncondition, the desired result. ∎

Corollary. *For any $\alpha \in \mathbb{R}$ and any $\theta \in \mathbb{R}^p$, we have*

$$E_\theta[(U'U)^\alpha\,(X-\theta)'\,g(X)] = \frac{1}{k+2\alpha}\,E_\theta[(U'U)^{\alpha+1}\operatorname{div}\,g(X)]$$

provided these expectations exist.

Then, if $\alpha=0$,

$$E_\theta[(X-\theta)'\,g(X)] = \frac{1}{k}\,E_\theta[(U'U)\operatorname{div}\,g(X)]$$

and, if $\alpha=1$,

$$E_\theta[(U'U)(X-\theta)'\,g(X)] = \frac{1}{k+2}\,E_\theta[(U'U)^2\operatorname{div}\,g(X)].$$

Proof. The result comes from the fact that, if $h(t)=t^{\alpha}$, then the indefinite integral of $-\frac{1}{2}h(t)\,t^{k/2-1}=-\frac{1}{2}t^{k/2-1+\alpha}$ vanishing at 0 is $H(t)=-t^{k/2+\alpha}/(k+2\alpha)$. ∎

LEMMA 3. *Let F be a probability distribution on* $[0,1]$ *and a and b be two positive constants. Then, for any* $x\geqslant 0$, *we have*

$$\int_0^1 \frac{1}{t}\,\frac{a+x}{b\,\dfrac{1-t}{t}+x}\,dF(t)\leqslant \int_0^{t_1}\frac{1}{t}\,dF(t)+\frac{a}{b}\int_{t_1}^1\frac{1}{1-t}\,dF(t)$$

$$\leqslant E\left[\frac{1}{t}\right]+\frac{a}{b}\,E\left[\frac{1}{1-t}\right]$$

where t_1 *is such that* $(a/b)(t_1/(1-t_1))=1$ (*i.e.*, $t_1=(1+b/a)^{-1}$) *and E is the expectation with respect to F.*

Proof. Note that for $0\leqslant t\leqslant t_1$ (respectively, $t_1\leqslant t\leqslant 1$) $(a+x)/(b(1-t)/t)+x)$ is an increasing (respectively decreasing) function of x. The result follows by bounding this function by its value at infinity (respectively, at 0).

REFERENCES

[1] Brandwein, A. C., and Strawderman, W. E. (1990). Stein estimation: The spherically symmetric case. *Statist. Sci.* **5** (3), 356–369.
[2] Brandwein, A. C., and Strawderman, W. E. (1991). Generalizations of James–Stein estimators under spherical symmetry, *Annal. Statist.* **19** (3), 1639–1650.
[3] Brandwein, A. C., and Strawderman, W. E. (1991). Improved estimates of location in the presence of an unknown scale. *J. Multivar. Anal.* **39** 305–314.
[4] Brandwein, A. C., Ralescu, S., and Strawderman, W. E. (1991). Shrinkage estimators of the location parameter for certain spherically symmetric distributions. *Ann. Inst. Statist. Math.* **45** (3), 551–563.
[5] Casella, G., and Hwang, J. T. (1982). Limit expressions for the risk of James–Stein estimators. *Canad. J. Statist.* **10** 305–309.
[6] Cellier, D., and Fourdrinier, D. (1995). Shrinkage estimators under spherical symmetry for the general linear model. *J. Multivar. Anal.* **52** (2), 338–351.
[7] Cellier, D., Fourdrinier, D., and Robert, C. (1989). Robust shrinkage estimators of the location parameter for elliptically symmetric distributions. *J. Multivar. Anal.* **29** 39–52.
[8] James, W., and Stein, C. (1961). Estimation with quadratic loss. In *Proc. Fourth Berkeley Symposium Mathematics, Statistics, and Probability*, Vol. 1, pp. 361–380. University of California Press, Berkeley.

[9] Lindley, D. V., and Smith, A. F. M. (1972). Bayes estimates for the linear model (with discussion), *J. Roy. Statist. Soc.* **34** 1–41.

[10] Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate distribution. In *Proc. Third Berkeley Symposium Mathematics, Statistics, and Probability*, Vol. 1, pp. 197–206. Berkeley, University of California Press.

[11] Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.* **9** (6), 1135–1151.