REVSTAT – Statistical Journal
Volume 8, Number 2, November 2010, 125–138

IMPROVING ON MINIMUM RISK EQUIVARIANT AND LINEAR MINIMAX ESTIMATORS OF BOUNDED MULTIVARIATE LOCATION PARAMETERS

Authors: Éric Marchand
– Département de Mathématiques, Université de Sherbrooke, Sherbrooke, Canada
eric.marchand@usherbrooke.ca

Amir T. Payandeh Najafabadi
– Department of Mathematical Sciences, Shahid Beheshti University, G.C. Evin, 1983963113, Tehran, Iran
amirtpayandeh@sbu.ac.ir

Received: February 2010 Revised: June 2010 Accepted: June 2010

Abstract:

- We propose improvements under squared error loss of the minimum risk equivariant and the linear minimax estimators for estimating the location parameter θ of a p-variate spherically symmetric distribution, with θ restricted to a ball of radius m centered at the origin. Our construction of explicit improvements relies on a multivariate version of Kubokawa's Integral Expression of Risk Difference (IERD) method. Applications are given for univariate distributions, for the multivariate normal, and for scale mixture of multivariate normal distributions.

Key-Words:

- *decision theory; spherical symmetric distribution; restricted parameter; minimum risk equivariant estimator; linear minimax estimator; dominating estimators; squared error loss.*

AMS Subject Classification:

- 62F10, 62F30, 62H12.

1. INTRODUCTION

We consider the problem of estimating, under squared error loss, the location parameter θ of a p-variate spherically symmetric distribution under the constraint $\|\theta\| \leq m$, with $m > 0$ known. With several authors having obtained interesting results relative to this problem, and more generally for restricted parameter space problems (see Marchand and Strawderman, 2004; van Eeden, 2006 for useful reviews), we focus on the determination of benchmark estimators such as the maximum likelihood estimator (MLE), the minimum risk equivariant estimator (MRE), and the linear minimax estimator (LMX). In this regard, Marchand and Perron (2001) provide for the multivariate normal case improvements on the (always) inadmissible MLE for all (m, p). These include Bayesian improvements, but conditions are then required on (m, p). Complementary findings for the multivariate normal and parallel findings for other spherically symmetric distributions, including in particular multivariate student distributions, were obtained respectively by Fourdrinier and Marchand (2010) and Marchand and Perron (2005); but again conditions for the studied priors π (typically boundary uniform, uniform on spheres, and fully uniform) of the form $m \leq c_\pi(p)$ for the Bayes estimator δ_π to dominate the MLE are necessitated. Hence, the problem of finding a Bayesian or an admissible improvement for any (m, p), for any given spherically symmetric distribution remains unsolved (even for $p = 1$ or the multivariate normal distribution).

Alternatively, for the objective of passing the minimum test of improving upon the minimum risk equivariant estimator, positive findings for the univariate case ($p = 1$) were obtained by Marchand and Strawderman (2005), as well as by Kubokawa (2005). The former establish a general dominance result for the fully uniform prior Bayes estimator, which actually applies more generally for a wider not necessarily symmetric class of location model densities and location invariant losses. The latter provides on the other hand a large class of priors which lead to Bayesian improvements for the univariate version of our problem of symmetric densities and squared error loss. A key feature of these dominance results is the use of Kubokawa's (1994) Integral Expression of Risk Difference (IERD) technique.

For multivariate settings, a lovely result by Hartigan (2004) tells us that for multivariate normal distributions, the fully uniform Bayes procedure improves upon the minimum risk equivariant estimator. The result is actually more general and applies for convex restricted parameter spaces with non-empty interiors. However, Hartigan's result does require normality and hence a spherically symmetric analog remains an open question. Moreover, Hartigan's result does not apply to the benchmark linear minimax estimator, which represents itself a simple improvement on the minimum risk equivariant estimator.

With the above background, our motivation here resides in extending the univariate dominance results to the multivariate case, extending Hartigan's result for balls to spherically symmetric distributions, and considering improvements upon the linear minimax procedure as well. We provide preliminary results in this direction in terms of sufficient conditions for dominating either the minimum risk equivariant estimator, the linear minimax estimator, or both. Our treatment possesses the interesting feature of being unified with respect to the dimension p and the given spherically symmetric distribution. Moreover, we arrive at our dominance results through a novel multivariate variant of Kubokawa's IERD technique. The main dominance results are presented in Section 2, and various examples or illustrations are pursued in Section 3. These include univariate distributions, the multivariate normal distribution, and scale mixture of multivariate normal distributions.

2. MAIN RESULTS

Let X be a p-variate random vector with spherically symmetric density

$$f\left(\|x-\theta\|^2\right), \tag{2.1}$$

where the location parameter θ is constrained to a ball centered at the origin and of radius m, say Θ_m. We seek improvements on the minimum risk equivariant (MRE) estimator $\delta_0(X)=X$, and the linear minimax estimator $\delta_{\mathrm{LMX}}(X)=\frac{m^2}{m^2+p\sigma^2}X$ under squared error loss $L(\theta,d)=\|d-\theta\|^2$, where $E_\theta(\|X-\theta\|^2)=p\,\sigma^2<\infty$. Hereafter, we denote the norms of X, x, and θ by R, r, and λ respectively. Our results bring into play the orthogonally invariant in θ and nonnegative quantities $H(t,\lambda)=\frac{E_\theta\left(\theta^T X|\ \|X\|\geq t\right)}{E_\theta\left(X^T X|\ \|X\|\geq t\right)}$ and $H^*(t,\lambda)=\frac{\lambda\, E_\theta\left(\|X\|\,\|X\|\geq t\right)}{E_\theta\left(|X^T X\|X\|\geq t\right)}$, $t\geq 0$, $\lambda\geq 0$. We will make use of the inequality $H(t,\lambda)\leq H^*(t,\lambda)$ for all $t\geq 0$, $\lambda\geq 0$, which follows as a simple application of the Cauchy–Schwartz inequality. Now, we present the main dominance results of this paper.

Theorem 2.1. *For a model as in (2.1), $\delta_g(X)=g(\|X\|)X$ dominates $g(0)X$, whenever:*

- **(i)** *g is absolutely continuous, nonconstant, and nonincreasing;*
- **(ii)** *and $g(r)\geq\sup_{\lambda\in[0,m]}H(r,\lambda)$ for all $r\geq 0$.*

Moreover, if conditions (i) and (ii) are satisfied, and

- **(iii)** *$g(0)\in\left[\frac{m^2-p\sigma^2}{m^2+p\sigma^2}\,,\,1\right)$,*

then $\delta_g(X)=g(\|X\|)X$ also dominates $\delta_0(X)=X$.

Remark 2.1. By virtue of the inequality $H(t,\lambda) \le H^*(t,\lambda)$ for all $t \ge 0$, $\lambda \ge 0$, condition (ii) of Theorem 2.1 can be replaced by the weaker, but nevertheless useful, condition

(ii′) and $g(r) \ge \sup_{\lambda\in[0,m]} H^*(r,\lambda)$ for all $r \ge 0$.

Proof of the Theorem: It is straightforward to verify that $g(0)X$ dominates X under condition (iii), so that conditions for which $\delta_g(X)$ dominates $g(0)X$, such as (i) and (ii), are necessarily conditions for which $\delta_g(X) = g(\|X\|)X$ also dominates $\delta_0(X) = X$. Now, using Kubokawa's IERD technique, the risk difference between the estimators $\delta_g(X)$ and $g(0)X$ can be written as

$$\begin{aligned}
\frac{1}{2}\,\Delta(\theta) &= \frac{1}{2}\Big[R\big(\theta, g(\|X\|)X\big) - R\big(\theta, g(0)X\big)\Big] \\
&= \frac{1}{2}\Big[E_\theta\big\|g(\|X\|)X-\theta\big\|^2 - \big\|g(0)X-\theta\big\|^2\Big] \\
&= \frac{1}{2}\,E_\theta\bigg(\int_0^{\|X\|}\frac{\partial}{\partial t}\big\|g(t)X-\theta\big\|^2\,dt\bigg) \\
&= \int_{\mathbb{R}^p}\int_0^{\|x\|} g'(t)\,\big[g(t)\,x-\theta\big]^T x\,f\big(\|x-\theta\|^2\big)\,dt\,dx \\
&= \int_0^\infty g'(t)\int_{\{x\in\mathbb{R}^p:\,\|x\|\ge t\}}\big[g(t)\,x^Tx-\theta^Tx\big]\,f\big(\|x-\theta\|^2\big)\,dx\,dt\,.
\end{aligned}$$

Now, observe that conditions (i) and (ii) imply that $\Delta(\theta) \le 0$ for all $\theta \in \Theta_m$, establishing the result. $\square$

Here are some further remarks and observations in relationship to Theorem 2.1.

The nonincreasing property of condition (i) is not necessarily restrictive. Indeed, for the multivariate normal case, Marchand and Perron (2001, theorem 5) establish that the nonincreasing property holds for all Bayesian estimators associated with symmetric, logconcave prior densities on $[-m,m]$. The conditions of Theorem 2.1 suggest the bounds (ii) and (ii′) themselves $\sup_{\lambda\in[0,m]} H(r,\lambda)$ and $\sup_{\lambda\in[0,m]} H^*(r,\lambda)$ as candidate g functions. These functions are of the form $H(r,\lambda(r))$ and $H^*(r,\lambda(r))$, where $\lambda(\cdot)$ is some function taking values on $[0,m]$. All such functions lead to range preserving estimators δ_g; i.e., $\|\delta_g(x)\| \le m$ for all $x \in \mathbb{R}^p$; since for all $r \ge 0$ and $\|\theta\| = \lambda(r)$:

$$0 \le H(r,\lambda(r)) \le H^*(r,\lambda(r)) = \frac{\lambda(r)\,E_\theta\big(\|X\|\,\|X\|\ge r\big)}{E_\theta\big(|X^TX\,\|X\|\ge r\big)} \le \frac{\lambda(r)}{r} \le \frac{m}{r}\,,$$

and since $\|\delta_g(x)\| \le m$ for all $x \in \mathbb{R}^p$ whenever $0 \le g(r) \le \frac{m}{r}$ for all $r > 0$. Finally, as a consequence of the above, observe that the projection of $\delta_0(X)$ onto Θ_m, given by δ_{g_p} with $g_p(r) = \frac{m}{r}\wedge 1$, satisfies the conditions of Theorem 2.1.

We now focus on related implications for the estimators $\delta_H(X) = H(\|X\|, m)X$ and $\delta_{H^*}(X) = H^*(\|X\|, m)X$, which will turn out in several cases to be the smallest possible g's satisfying respectively conditions (ii) and (ii′) of Theorem 2.1.

Corollary 2.1.

(**a**) *If $H(r,\lambda)$ increases in $\lambda \in [0,m]$ for all $r \geq 0$, and decreases in $r \in [0,\infty]$ for all $\lambda \in [0,m]$, then $\delta_H(X) = H(\|X\|,m)X$ dominates both the linear minimax estimator $\delta_{\mathrm{LMX}}(X)$ and the MRE estimator $\delta_0(X)$;*

(**b**) *If $H^*(r,\lambda)$ increases in $\lambda \in [0,m]$ for all $r \geq 0$, then $\delta_{H^*}(X) = H^*(\|X\|,m)X$ dominates the MRE estimator $\delta_0(X)$.*

Proof: Part (a) follows as a direct application of Theorem 2.1 as $H(0,m) = \frac{E_\theta(\theta^T X)}{E_\theta(X^T X)} = \frac{m^2}{m^2+p\sigma^2} \in \left[\frac{m^2-p\sigma^2}{m^2+p\sigma^2}, 1\right)$, for $\|\theta\| = m$. Part (b) follows for two reasons. First, for any positive random variable Y with density g_Y, and its biased version W with density proportional to $w g_Y(w)$, the ratio $\frac{E(Y^2|\,Y>t)}{E(Y|\,Y>t)} = E(W|\,W>t)$ is increasing in t, which implies that $H^*(\cdot,m)$ is a decreasing function on $[0,\infty)$. Secondly, for $\|\theta\| = m$, $H^*(0,m) = m\frac{E_\theta(\|X\|)}{E_\theta(\|X\|^2)} = \frac{E_\theta(\|X/m\|)}{E_\theta(\|X/m\|^2)} < \frac{E_\theta(\|X/m\|)^2}{E_\theta(\|X/m\|^2)} < 1$, since $E_\theta(\|X\|) > \|E_\theta(X)\| = m$. □

3. EXAMPLES

The following subsections are devoted to applications of Corollary 2.1, with the key difficulty arising in checking the monotonicity conditions relative to H and H^*. We focus on general univariate cases (subsection 3.1), the multivariate normal distribution (subsection 3.2.), and scale mixtures of multivariate normal distributions (subsection 3.3).

3.1. Univariate spherically symmetric distributions

We express the symmetric univariate densities in (2.1) as

$$f_\theta(x) = e^{-q(x-\theta)}\,, \tag{3.1}$$

and restrict ourselves to cases where

$$q \in Q^* = \Big\{q\colon\ q(\cdot) \text{ is increasing and convex on } (0,\infty), \\ \text{and } q'(\cdot) \text{ is concave on } (0,\infty)\Big\}.$$

Examples of such distributions include normal, Laplace, exponential power densities with $q(y) = \alpha\, y^\beta + c$, $\alpha > 0$, $1 \leq \beta \leq 2$; Hyperbolic Secant, Logistic, Generalized logistic densities with $q(y) = -y + \frac{2}{\alpha}\log(1+e^{\alpha y}) + c$, $\alpha > 0$; and Champernowne densities with $q(y) = \log(\cosh(y)+\beta)$, $\beta \in [0,2]$, (also see Marchand

and Perron, 2009; Marchand *et al.*, 2008). The next theorem establishes for such densities the applicability of part (a) of Corollary 2.1 and dominance of $\delta_H(X)$ over both the linear minimax estimator, $\delta_{\mathrm{LMX}}(X)$, and the MRE estimator $\delta_0(X)$.

Theorem 3.1. *For model densities as in (3.1) with $q \in Q^*$, the estimator $\delta_H(X) = H(\|X\|, m)X$ dominates both the linear minimax estimator $\delta_{\mathrm{LMX}}(X)$ and the MRE estimator $\delta_0(X)$.*

Proof: By virtue of Corollary 2.1, it suffices to show that $H(r,\lambda)$ decreases in $r \in [0,\infty)$ for all $\lambda \in [0,m]$, and increases in $\lambda \in [0,m]$ for all $r \geq 0$. First, $H(r,\lambda)$ can be written as

$$\begin{aligned} H(r,\lambda) &= \lambda \, \frac{\int_r^\infty x\big(f_0(x-\lambda) - f_0(x+\lambda)\big)\,dx}{\int_r^\infty x^2\big(f_0(x-\lambda) + f_0(x+\lambda)\big)\,dx} \\ &= \lambda^2 E_\lambda\left(\frac{\tanh\Big(\big(q(Y+\lambda) - q(Y-\lambda)\big)/2\Big)}{\lambda Y}\right), \end{aligned}$$

where Y is a random variable with density proportional to $y^2(f_0(y-\lambda)+f_0(y+\lambda))\cdot$ $\cdot 1_{[r,\infty)}(y)$. Such a family of densities with parameter r has increasing monotone likelihood ratio in Y. Furthermore, since $q \in Q^*$, a result of Marchand *et al.* (2008) (Lemma 1, part e) tells us that the inner function of the above expectation in Y is nonincreasing. Hence, we conclude that, for all $\lambda \in [0,m]$, $H(\lambda,\cdot)$ decreases on $[0,\infty)$. Turning to the monotonicity of $H(\cdot,r)$, begin by writing

$$\begin{aligned} H(r,\lambda) &= \lambda \, \frac{\int_r^\infty x\big(f_0(x-\lambda) - f_0(x+\lambda)\big)\,dx}{\int_r^\infty x^2\big(f_0(x-\lambda) + f_0(x+\lambda)\big)\,dx} \\ &= \lambda \, \frac{\int_{r-\lambda}^\infty (y+\lambda)\,f_0(y)\,dy \;-\; \int_{r+\lambda}^\infty (y-\lambda)\,f_0(y)\,dy}{\int_{r-\lambda}^\infty (y+\lambda)^2\,f_0(y)\,dy \;+\; \int_{r+\lambda}^\infty (y-\lambda)^2\,f_0(y)\,dy} \\ &= \lambda \, \frac{A(r,\lambda)}{B(r,\lambda)}\,, \end{aligned}$$

where $A(r,\lambda)$ and $B(r,\lambda)$ are the numerator and denominator of the above fraction, respectively. Manipulations yield:

$$\begin{aligned} B^2(r,\lambda)\,\frac{\partial H(r,\lambda)}{\partial\lambda} &= A(r,\lambda)\,B(r,\lambda) + \lambda\,A'(r,\lambda)\,B(r,\lambda) - \lambda\,A(r,\lambda)\,B'(r,\lambda) \\ &= \Big[l(r,\lambda) + A_1(r,\lambda)\Big] \\ &\quad \cdot \Big[B_1(r,\lambda) + r\lambda\big(\lambda f_0(r-\lambda) + \lambda f_0(r+\lambda) - r f_0(r-\lambda) + r f_0(r+\lambda)\big)\Big] \\ &\quad + \Big[r\lambda\big(f_0(r-\lambda) + f_0(r+\lambda)\big) + A_1(r,\lambda)\Big]\,\Big[B_1(r,\lambda) + \lambda\,l(r,\lambda)\Big] \\ &= r\lambda\,G(r,\lambda)\,f_0(r-\lambda) + 2\,A_1(r,\lambda)\,B_1(r,\lambda) + r\lambda^2 f_0(r+\lambda)\,l(r,\lambda) \\ &\quad + r^2\lambda\,f_0(r+\lambda)\,l(r,\lambda) + r\lambda\,f_0(r+\lambda)\,B_1(r,\lambda) + r\lambda^2 f_0(r+\lambda)\,l(r,\lambda)\,, \end{aligned}$$

where

$$\begin{aligned}
G(r,\lambda) &= 2\,\lambda \int_{r-\lambda}^{r+\lambda} y f_0(y)\,dy - r\int_{r-\lambda}^{r+\lambda} y f_0(y)\,dy + \int_{r-\lambda}^{\infty} y^2 f_0(y)\,dy + \int_{r+\lambda}^{\infty} y^2 f_0(y)\,dy \ , \\
l(r,\lambda) &= \int_{r-\lambda}^{r+\lambda} y f_0(y)\,dy \ , \\
A_1(r,\lambda) &= \lambda\bigg(\int_{r-\lambda}^{\infty} f_0(y)\,dy + \int_{r+\lambda}^{\infty} f_0(y)\,dy\bigg) \ , \\
B_1(r,\lambda) &= \int_{r-\lambda}^{\infty} y^2 f_0(y)\,dy + \int_{r+\lambda}^{\infty} y^2 f_0(y)\,dy \ .
\end{aligned}$$

Now, observe that for all $r \geq 0$, $\lambda \in [0,m]$, the quantities $B_1(r,\lambda)$, $A_1(r,\lambda)$, and (r,λ) are nonnegative. Hence, to show the positivity of $\frac{\partial H(r,\lambda)}{\partial \lambda}$, it will suffice to show the positivity of $G(r,\lambda)$. But, we have

$$\begin{aligned}
G(r,\lambda) &\geq \int_{r-\lambda}^{r+\lambda} y\, f_0(y)\,(2\,\lambda - r + y)\;dy \\
&\geq \int_0^{r+\lambda} \lambda\, y\, f_0(y)\,dy\;\, 1_{[\lambda,\infty)}(r) \;+ \int_{r-\lambda}^{\lambda-r} y\, f_0(y)\,(2\,\lambda - r + y)\;dy\;\, 1_{[0,\lambda)}(r) \\
&\geq \int_0^{\lambda-r} 2\,y^2 f_0(y)\,dy\;\, 1_{[0,\lambda)}(r) \;\geq\; 0 \ ,
\end{aligned}$$

which completes the proof. □

3.2. Multivariate normal distributions

We consider here multivariate normal models in (2.1) $X \sim N_p(\theta,\sigma^2)$ with $\|\theta\| \leq m$. We take $\sigma^2 = 1$ without loss of generality (since $\frac{X}{\sigma} \sim N_p(\theta' = \frac{\theta}{\sigma}, I_p)$ with $\|\theta'\| \leq m' = \frac{m}{\sigma}$). We require the following key properties relative to $\rho(\lambda,r) = E_\theta\Big(\frac{\theta^T X}{\|X\|} \mid \|X\| = r\Big)$, where $\lambda = \|\theta\|$. These properties involve modified Bessel functions $\mathbb{I}_v$ of order v, and more specifically ratios of the form $\rho_v(t) = \mathbb{I}_{v+1}(t)/\mathbb{I}_v(t)$, $t > 0$.

Lemma 3.1 (Watson, 1983; Marchand and Perron, 2001)**.**

(i) *We have* $\rho(\lambda,r) = \lambda\rho_{p/2-1}(\lambda\, r)$;

(ii) $\rho_{p/2-1}(\cdot)$ *is increasing and concave on* $[0,\infty)$, *with* $\rho_{p/2-1}(0) = 0$ *and* $\lim_{t\to\infty} \rho_{p/2-1}(t) = 1$;

(iii) $\rho_{p/2-1}(t)/t$ *is decreasing in* t *with* $\lim_{t\to 0^+} \rho_{p/2-1}(t)/t = 1/p$;

(iv) $\rho_{p/2}(t) = \rho^{-1}_{p/2-1}(t) - p/t$.

Denoting $f_p(\cdot,\lambda)$ and $\bar{F}_p(\cdot,\lambda)$ as the probability density and survival functions of $R = \|X\| \sim \sqrt{\chi^2_p(\lambda^2)}$, we will also require the following useful properties.

Lemma 3.2.

(i) *We have* $f_p(r,\lambda) = r\left(\frac{r}{\lambda}\right)^{p/2-1} \mathbb{I}_{p/2-1}(r\lambda)\exp\left\{-\frac{r^2+\lambda^2}{2}\right\}$;

(ii) $r^2 f_p(r,\lambda) = \lambda^2 f_{p+4}(r,\lambda) + p\, f_{p+2}(r,\lambda)$;

(iii) $r\, f_p(r,\lambda)\, \rho_{p/2-1}(\lambda r) = \lambda\, f_{p+2}(r,\lambda)$;

(iv) *the ratio* $\frac{\bar{F}_{p+2}(r,\lambda)}{\bar{F}_p(r,\lambda)}$ *decreases in* $\lambda \in [0,\infty)$, *for all* $p \geq 1$ *and* $r > 0$.

Proof: Parts (ii) and (iii) follow directly from (i), while (i) consists of a well known Bessel function representation of the noncentral chi-square distribution. Part (iv) follows from the identity $2\,\frac{\partial}{\partial\lambda}\bar{F}_p(r,\lambda) = \bar{F}_{p+2}(r,\lambda) - \bar{F}_p(r,\lambda)$, and the logconcavity of $\bar{F}_p(r,\cdot)$ on $[0,\infty)$ (see Das Gupta and Sarkar, 1984; Finner and Roters, 1997). $\square$

We now seek to apply part (a) of Corollary 2.1.

Theorem 3.2. *For multivariate normal densities, the estimator* $\delta_H(X) = H(\|X\|,m)X$ *dominates both the linear minimax estimator* $\delta_{\mathrm{LMX}}(X)$ *and the MRE estimator* $\delta_0(X)$.

Proof: By virtue of Corollary 2.1, it suffices to show that $H(r,\lambda)$ decreases in $r \in [0,\infty)$ for all $\lambda \in [0,m]$, and increases in $\lambda \in [0,m]$ for all $r \geq 0$. Making use of Lemmas 3.1 and 3.2, we obtain

$$
\begin{aligned}
H(r,\lambda) &= \frac{E_\theta\Big(\|X\|\, E_\theta\Big(\frac{\theta^T X}{\|X\|} \mid \|X\| \geq r\Big)\Big)}{E_\theta\big(\|X\|^2 \mid \|X\| \geq r\big)} \\
&= \frac{\int_\infty^r y\, E_\theta\Big(\frac{\theta^T X}{\|X\|} \mid \|X\| = y\Big) f_p(y,\lambda)\,dy}{\int_r^\infty y^2 f_p(y,\lambda)\,dy} \\
&= \frac{\int_\infty^r y\,\lambda\rho_{p/2-1}(\lambda y)\, f_p(y,\lambda)\,dy}{\int_r^\infty y^2 f_p(y,\lambda)\,dy} \\
&= \frac{\int_\infty^r \lambda^2 f_{p+2}(y,\lambda)\,dy}{\int_r^\infty y^2 f_p(y,\lambda)\,dy} \\
&= \left\{\frac{p}{\lambda^2} + \frac{\int_r^\infty f_{p+4}(y,\lambda)\,dy}{\int_r^\infty f_{p+2}(y,\lambda)\,dy}\right\}^{-1} \\
&= \left\{\frac{p}{\lambda^2} + \frac{\bar{F}_{p+4}(r,\lambda)}{\bar{F}_{p+2}(r,\lambda)}\right\}^{-1}.
\end{aligned}
\tag{3.2}
$$

The monotonicity property of $H(r,\cdot)$ on $[0,m]$ for all $r \geq 0$ now follows from the above expression and part (iv) of Lemma 3.2.

Now, to show that $H(r,\lambda)$ decreases in r, make use of (3.2) to write

$$\begin{aligned} H(r,\lambda) &= \lambda\, E_r\left(\frac{E_Y\left(\frac{\theta^T X}{\|X\|} \mid \|X\| = Y\right)}{Y}\right) \\ &= \lambda\, E_r\left(\frac{\rho_{p/2-1}(\lambda Y)}{Y}\right), \end{aligned}$$

where Y has density proportional to $y\, f_p(y,\lambda)\, 1_{[r,\infty)}(y)$. Since this family of densities with parameter r has increasing monotone likelihood ratio in Y, we conclude indeed that $H(r,\lambda)$ decreases for $r \geq 0$ for all $\lambda \in [0,m]$ by making use of part (iii) of Lemma 3.1. $\square$

3.3. Scale mixtures of multivariate normal distributions

We consider here in this subsection scale mixtures of multivariate normal distributions where X admits the representation: $X|Z=z \sim N_p(\theta, zI_p)$, Z having Lebesgue density g on $\mathbb{R}^+$. The corresponding density in (2.1) is of the form

$$\int_0^\infty (2\pi z)^{-p/2} \exp\left\{-\frac{\|x-\theta\|^2}{2z}\right\} g(z)\, dz\,; \tag{3.3}$$

and we further assume that g is logconcave on either $\mathbb{R}^+$ or some open interval (a,b) of $\mathbb{R}^+$. Uniform densities on (a,b) are included. With such a representation, since $X/\sqrt{Z}|Z=z \sim N_p(\theta/\sqrt{z}, I_p)$, we infer from part (i) of Lemma 3.2 that the density function of $R = \|X\|$ is given by

$$\int_0^\infty \frac{y}{z}\left(\frac{y}{\lambda}\right)^{p/2-1} \mathbb{I}_{p/2-1}\left(\frac{\lambda y}{z}\right) \exp\left\{-\frac{y^2+\lambda^2}{2z}\right\} g(z)\, dz\,. \tag{3.4}$$

We now seek to apply part (a) of Corollary 2.1.

Theorem 3.3. *For scale mixtures of multivariate normal densities as in (3.3) with g logconcave, the estimator $\delta_{H^*}(X) = H^*(\|X\|, m)X$ dominates the MRE estimator $\delta_0(X)$.*

Proof: By virtue of Corollary 2.1, it suffices to show that $H^*(r,\cdot)$ is nondecreasing on $[0,m]$ for all $r \geq 0$ under the given logconcave assumption on g.

Starting from the definition of H^* and making use of 3.4, we obtain

$$
\begin{aligned}
H^*(r,\lambda) &= \frac{\lambda\, E_\theta\big(R \,|\, R \ge r\big)}{E_\theta\big(R^2 \,|\, R \ge r\big)} \\
&= \frac{\lambda \int_r^\infty \int_0^\infty y^{\frac{p}{2}+1}\, \frac{g(z)}{z}\, \mathbb{I}_{\frac{p}{2}-1}\big(\frac{y\lambda}{z}\big)\, e^{-\frac{y^2+\lambda^2}{2z}}\, dz\, dy}{\int_r^\infty \int_0^\infty y^{\frac{p}{2}+2}\, \frac{g(z)}{z}\, \mathbb{I}_{\frac{p}{2}-1}\big(\frac{y\lambda}{z}\big)\, e^{-\frac{y^2+\lambda^2}{2z}}\, dz\, dy} \\
&= \frac{\int_{r/\lambda}^\infty \int_0^\infty x^{\frac{p}{2}+1}\, \frac{g(\lambda^2 t)}{t}\, \mathbb{I}_{\frac{p}{2}-1}\big(\frac{x}{t}\big)\, e^{-\frac{1+x^2}{2t}}\, dt\, dx}{\int_{r/\lambda}^\infty \int_0^\infty x^{\frac{p}{2}+2}\, \frac{g(\lambda^2 t)}{t}\, \mathbb{I}_{\frac{p}{2}-1}\big(\frac{x}{t}\big)\, e^{-\frac{1+x^2}{2t}}\, dt\, dx}\,,
\end{aligned}
$$

with the change of variables $(y,z) = (\lambda x, \lambda^2 t)$. Simple differentiation leads to $\frac{\partial}{\partial\lambda} H^*(r,\lambda) = \frac{1}{B^2}\{A_1 - A_2 + A_3 - A_4\}$, where B is the above denominator of H^*,

$$
\begin{aligned}
A_1 &= 2\lambda \int_{r/\lambda}^\infty \int_0^\infty x\, M(x,t)\, dt\, dx \int_{r/\lambda}^\infty \int_0^\infty x^{\frac{p}{2}+1}\, g'(\lambda^2 t)\, \mathbb{I}_{\frac{p}{2}-1}\left(\frac{x}{t}\right) e^{-\frac{1+x^2}{2t}}\, dt\, dx\,, \\
A_2 &= 2\lambda \int_{r/\lambda}^\infty \int_0^\infty M(x,t)\, dt\, dx \int_{r/\lambda}^\infty \int_0^\infty x^{\frac{p}{2}+2}\, g'(\lambda^2 t)\, \mathbb{I}_{\frac{p}{2}-1}\left(\frac{x}{t}\right) e^{-\frac{1+x^2}{2t}}\, dt\, dx\,, \\
A_3 &= \frac{r}{\lambda^2} \int_{r/\lambda}^\infty \int_0^\infty x\, M(x,t)\, dt\, dx \int_0^\infty \frac{g(\lambda^2 t)}{t} \left(\frac{r}{\lambda}\right)^{\frac{p}{2}+1} \mathbb{I}_{\frac{p}{2}-1}\left(\frac{r}{\lambda t}\right) e^{-\frac{\lambda^2+r^2}{2\lambda^2 t}}\, dt\,, \\
A_4 &= \frac{r}{\lambda^2} \int_{r/\lambda}^\infty \int_0^\infty M(x,t)\, dt\, dx \int_0^\infty \frac{g(\lambda^2 t)}{t} \left(\frac{r}{\lambda}\right)^{\frac{p}{2}+2} \mathbb{I}_{\frac{p}{2}-1}\left(\frac{r}{\lambda t}\right) e^{-\frac{\lambda^2+r^2}{2\lambda^2 t}}\, dt\,,
\end{aligned}
$$

with $M(x,t) = \frac{g(\lambda^2 t)}{t}\, x^{\frac{p}{2}+1}\, \mathbb{I}_{\frac{p}{2}-1}\big(\frac{x}{t}\big)\, e^{-\frac{1+x^2}{2t}}$. Obviously, $A_3 - A_4 \ge 0$, because $x \ge \frac{r}{\lambda}$ on the domain of integration. Furthermore, by setting $h(z) = \big(-g'(z)/g(z)\big)\cdot$ $\cdot 1_{\{z: g(z)>0\}}(z)$, we have

$$
\begin{aligned}
A_1 - A_2 &= 2\lambda \int_{r/\lambda}^\infty \int_0^\infty M(x,t)\, dt\, dx \int_{r/\lambda}^\infty \int_0^\infty h(\lambda^2 t)\, xt\, M(x,t)\, dt\, dx \\
&\quad - 2\lambda \int_{r/\lambda}^\infty \int_0^\infty x\, M(x,t)\, dt\, dx \int_{r/\lambda}^\infty \int_0^\infty h(\lambda^2 t)\, t\, M(x,t)\, dt\, dx\,.
\end{aligned}
$$

Now, since h is increasing with the logconcavity of g, the FKG's inequality (see Lemma A.1 in the Appendix) implies that $A_1 - A_2$ is nonnegative whenever $M(x_1,t_2)\, M(x_2,t_1) - M(x_1,t_1)\, M(x_2,t_2) \le 0$, for $0 \le x_1 \le x_2$ and $0 \le t_1 \le t_2$. From the definition of M, manipulations yield for non-zero values of $M(x_1,t_2)\cdot$

$\cdot M(x_2,t_1) - M(x_1,t_1)\,M(x_2,t_2)$:

$$\frac{t_1\,t_2\,e^{(1/t_1+1/t_2)}}{(x_1x_2)^{p/2+1}\,g(\lambda^2 t_1)\,g(\lambda^2 t_2)}\,\Big\{M(x_1,t_2)\,M(x_2,t_1) - M(x_1,t_1)\,M(x_2,t_2)\Big\} =$$

$$= \mathbb{I}_{\frac{p}{2}-1}\!\left(\frac{x_1}{t_2}\right)\mathbb{I}_{\frac{p}{2}-1}\!\left(\frac{x_2}{t_1}\right) - \mathbb{I}_{\frac{p}{2}-1}\!\left(\frac{x_1}{t_1}\right)\mathbb{I}_{\frac{p}{2}-1}\!\left(\frac{x_2}{t_2}\right)\,\exp\Big\{-(x_1^2-x_2^2)\,(1/t_1-1/t_2)\Big\}$$

$$= \mathbb{I}_{\frac{p}{2}-1}\!\left(\frac{x_1}{t_2}\right)\mathbb{I}_{\frac{p}{2}-1}\!\left(\frac{x_2}{t_2}\right)\left[\frac{\mathbb{I}_{\frac{p}{2}-1}\big(\frac{x_2}{t_1}\big)}{\mathbb{I}_{\frac{p}{2}-1}\big(\frac{x_2}{t_2}\big)} - \frac{\mathbb{I}_{\frac{p}{2}-1}\big(\frac{x_1}{t_1}\big)}{\mathbb{I}_{\frac{p}{2}-1}\big(\frac{x_1}{t_2}\big)}\,\exp\Big\{-(x_1^2-x_2^2)\,(1/t_1-1/t_2)\Big\}\right]$$

$$\le \mathbb{I}_{\frac{p}{2}-1}\!\left(\frac{x_1}{t_2}\right)\mathbb{I}_{\frac{p}{2}-1}\!\left(\frac{x_2}{t_1}\right)\left(\frac{t_2}{t_1}\right)^{p/2-1}\left[1-\exp\Big\{(x_2^2-x_1^2+x_1)\,(1/t_1-1/t_2)\Big\}\right]$$

$$\le\ 0\ ,$$

where the former inequality follows from the Ross inequality applications (see Lemma A.2 in Appendix): $\frac{\mathbb{I}_{p/2-1}(x_2/t_1)}{\mathbb{I}_{p/2-1}(x_2/t_2)} \le (t_2/t_1)^{p/2-1}$ and $\frac{\mathbb{I}_{p/2-1}(x_1/t_1)}{\mathbb{I}_{p/2-1}(x_1/t_2)} \ge (t_2/t_1)^{p/2-1}\exp\big\{x_1/t_1 - x_1/t_2\big\}$, and where the latter inequality follows from the fact that $(x_2^2 - x_1^2 + x_1)\,(1/t_1 - 1/t_2) \ge 0$, for $0 \le x_1 \le x_2$ and $0 \le t_1 \le t_2$. $\square$

APPENDIX

The FKG inequality due to Fortuin, Kasteleyn, and Ginibre (1971) is useful for Theorem 3.3.

Lemma A.1 (**FKG inequality**). *Suppose a p-variate random variable X is distributed with probability density function ξ and with positive measure ν. For two points $y=(y_1,...,y_p)$ and $z=(z_1,...,z_p)$, in the sample space of X, we define $y\wedge z=(y_1\wedge z_1,\,...,\,y_p\wedge z_p)$ and $y\vee z=(y_1\vee z_1,\,...,\,y_p\vee z_p)$, where $a\wedge b=\min(a,b)$, $a\vee b=\max(a,b)$. Suppose that ξ satisfies $\xi(y)\,\xi(z)\le\xi(y\vee z)\,\xi(y\wedge z)$ and that $\alpha(y)$, $\beta(y)$ are nondecreasing in each argument and α, β and $\alpha\,\beta$ are integrable with respect to ξ. Then $\int\alpha\,\beta\,\xi\,d\nu\ge\int\alpha\,\xi\,d\nu\int\beta\,\xi\,d\nu$.*

The following lemma, referred to as the Ross inequality is due to Joshi and Bissu (1991) and establishes useful bounds for a ratio of modified Bessel functions.

Lemma A.2. *Suppose $\mathbb{I}_v(x)$ and $\mathbb{I}_v(y)$ are two modified Bessel functions. Moreover, suppose that $y\ge x$ and $v\ge-\frac{1}{2}$. Then*

$$e^{x-y}\left(\frac{x}{y}\right)^{v}\le\frac{\mathbb{I}_v(x)}{\mathbb{I}_v(y)}\le\left(\frac{x}{y}\right)^{v}.$$

ACKNOWLEDGMENTS

Éric Marchand acknowledges the partial research support from NSERC of Canada. We are grateful to Dan Kucerovsky and Reinaldo Arellano-Valle for useful suggestions on the Ross inequality and properties of the H function in Theorem 2.1. Thanks to an anonymous reviewer for constructive comments.

REFERENCES

[1] Das Gupta, S. and Sarkar, S.K. (1984). On TP_2 and log-concavity, *Inequalities in Statistics and Probability, IMS Lecture Notes Monogr. Ser.*, **5**, 54–58.

[2] Fortuin, C.M.; Kasteleyn, P.W. and Ginibre, J. (1971). Correlation inequalities on some partially ordered sets, *Comm. Math. Phys.*, **22**, 89–103.

[3] Fourdrinier, D. and Marchand, É. (2010). On Bayes estimators with uniform priors on spheres and their comparative performance with maximum likelihood estimators for estimating bounded multivariate normal means, *Journal of Multivariate Analysis*, **101**, 1390–1399.

[4] Hartigan, J. (2004). Uniform priors on convex improve risk, *Statistics and Probability Letters*, **67**, 285–288.

[5] Finner, H. and Roters, M. (1997). Log-concavity and inequalities for Chi-square, F and Beta distributions with applications in multiple comparisons, *Statistica Sinica*, **7**, 771–787.

[6] Joshi, C.M. and Bissu, S.K. (1991). Some inequalities of Bessel and modified Bessel functions, *Australian Mathematical Society Journal, Series A, Pure Mathematics and Statistics*, **50**, 333–342.

[7] Kubokawa, T. (1994). A unified approach to improving equivariant estimators, *Annals of Statistics*, **22**, 290–299.

[8] Kubokawa, T. (2005). Estimation of bounded location and scale parameters, *Journal of the Japan Statistical Society*, **35**, 221–249.

[9] Marchand, 'E.; Ouassou, I.; Payandeh, A.T. and Perron, F. (2008). On the estimation of a restricted location parameter for symmetric distributions, *Journal of the Japan Statistical Society*, **38**, 1–17.

[10] Marchand, É. and Perron, F. (2001). Improving on the MLE of a bounded normal mean, *The Annals of Statistics*, **29**, 1078–1093.

[11] Marchand, É. and Perron, F. (2005). Improving on the MLE of a bounded mean for spherical distributions, *Journal of Multivariate Analysis*, **92**, 227–238.

[12] Marchand, É. and Strawderman, W.E. (2004). Estimation in restricted parameter spaces: a review, *A Festschrift for Herman Rubin, IMS Lecture Notes Monogr. Ser.*, **45**, 21–44.

[13] Marchand, É. and Strawderman, W.E. (2005). Improving on the minimum risk equivariant estimator for a location parameter which is constrained to an interval or a half-interval, *Annals of the Institute of Statistical Mathematics*, **57**, 129–143.

[14] van Eeden, C. (2006). *Restricted parameter space estimation problems. Admissibility and minimaxity properties*, Lecture Notes in Statistics, **188**, Springer, New York.

[15] Watson, G.S. (1983). *Statistics on Spheres*, John Wiley and Sons, Inc., New York.