IMS Collections
Borrowing Strength: Theory Powering Applications –
A Festschrift for Lawrence D. Brown
Vol. 6 (2010) 249–262

DOI: 10.1214/10-IMSCOLL617

Robust generalized Bayes minimax estimators of location vectors for spherically symmetric distributions with unknown scale

Dominique Fourdrinier[1,*], **William E. Strawderman**[2]

Université de Rouen and Rutgers University

Abstract: We consider estimation of the mean vector, θ, of a spherically symmetric distribution with unknown scale parameter σ under scaled quadratic loss. We show minimaxity of generalized Bayes estimators corresponding to priors of the form $\pi(\|\theta\|^2)\eta^b$ where $\eta = 1/\sigma^2$, for $\pi(\cdot)$ superharmonic with a non decreasing Laplacian under conditions on b and weak moment conditions. Furthermore, these generalized Bayes estimators are independent of the underlying density and thus have the strong robustness property of being simultaneously generalized Bayes and minimax for the entire class of spherically symmetric distributions.

1. Introduction

Let (X, U) be a random vector in $\mathbb{R}^p \times \mathbb{R}^k$ with density

$$\frac{1}{\sigma^{p+k}}\, f\left(\frac{\|x-\theta\|^2 + \|u\|^2}{\sigma^2}\right), \tag{1.1}$$

where $\theta \in \mathbb{R}^p$ and $\sigma \in \mathbb{R}_+^*$ are unknown. We assume throughout that $p \geq 3$.

We consider generalized Bayes estimators of θ for priors of the form

$$\pi(\|\theta\|^2)\,\eta^b, \tag{1.2}$$

where $\eta = 1/\sigma^2$, under the quadratic loss

$$\eta\,\|\delta - \theta\|^2. \tag{1.3}$$

We first show that, under weak moment conditions, such generalized Bayes estimators are robust in the sense that they do not depend on the underlying density f. Furthermore, we exhibit a large class of superharmonic priors π for which these generalized Bayes estimators dominate the usual minimax estimator X for the entire

*The support of the ANR grant 08-EMER-002 is gratefully acknowledged.
[1]Université de Rouen, LITIS EA 4108, Avenue de l'Université, BP 12, 76801 Saint-Étienne-du-Rouvray, France, e-mail: Dominique.Fourdrinier@univ-rouen.fr
[2]Rutgers University, Hill Center, Department of Statistics, 110 Frelinghuysen Rd., Piscataway NJ 08854, USA, e-mail: straw@Stat.rutgers.edu
Keywords and phrases: Bayes estimators, minimax estimators, spherically symmetric distributions, location parameter, scale parameter, quadratic loss, superharmonic priors.
AMS 2000 subject classifications: Primary 62C10, 62C20.

class of densities (1.1). Hence this subclass of estimators has the extended robustness property of being simultaneously generalized Bayes and minimax for the entire class of spherically symmetric distributions.

Note that the above model arises as the canonical form of the general linear model $Y = V\beta + \varepsilon$ where V is a $(p + k) \times p$ design matrix, β is a $p \times 1$ vector of unknown regression coefficients, and ε is an $(p+k) \times 1$ error vector with spherically symmetric density $f(\|\varepsilon\|^2/\sigma^2)/\sigma^{p+k}$.

In the following, for a real valued function $g(x, u)$, we denote by $\bigtriangledown_x g(x, u)$ and $\triangle_x g(x, u)$ the gradient and the Laplacian of $g(x, u)$ with respect to the variable x. Analogous notations hold with respect to the variable u. When $g(x, u)$ is a vector valued function, $\text{div}_x g(x, u)$ is the divergence with respect to x (here $\dim g(x, u) = \dim x$).

Stein [20] shows that, when the density in (1.1) is normal with known scale, the generalized Bayes estimator corresponding to a prior $\pi(\theta)$, for which the square root of the marginal density $m(x)$ is superharmonic, is minimax under the loss (1.3). Fundamental to this result is the development of an unbiased estimator of risk based on a differential expression involving $m(x)$ which has become a basic tool to prove minimaxity. This differential expression has been extended to non normal models such as (1.1) by several authors (see, for example, [5, 6, 2, 3, 4, 12, 13, 14, 16, 9] and [10]).

A notable aspect of many of the papers dealing with model (1.1), in particular in the presence of a residual vector U, is the development of robust estimators in the sense that they are minimax for a wide class of spherically symmetric distributions (see particularly, for example, [5, 4, 13]).

The general line of research pertinent to this paper is the development of Bayes and generalized Bayes minimax estimators. In the case of a normal distribution with known scale, see for example, [21, 7, 1, 11]. When the scale is unknown, see [22]. For variance mixture of normals with known scale and no residual vector, see [23, 16, 8]. For general spherically symmetric distributions (with no residual vector), see [9].

Maruyama [15] showed that, for spherically symmetric distributions with a residual vector U and unknown scale parameter, the generalized Bayes estimator with respect to a prior on θ and η proportioned to $\|\theta\|^{2-p}$ (i.e. the fundamental harmonic) is independent of the density f and is minimax under weak moment conditions (see also [18, 17] and [19]).

The goal of this paper is to extend the phenomenon in [15] to a broader class of priors of the form $\pi(\|\theta\|^2)\eta^b$ with $\pi(\|\theta\|^2)$ superharmonic. In particular, in Section 2.2, we show that the generalized Bayes estimators do not depend on the density f under weak moment conditions and, in Section 2.3, we prove that these generalized Bayes estimators are minimax provided the prior $\pi(\|\theta\|^2)$ is superharmonic and its Laplacian $\Delta\pi(\|\theta\|^2)$ is a non decreasing function of $\|\theta\|^2$, under conditions on b, p and k.

In the case of a known scale parameter in model (1), [9] studied the same class of priors $\pi(\|\theta\|^2)$ and proved minimaxity of generalized Bayes estimators for a large subclass of unimodal densities. We rely strongly on the techniques of that paper.

Our main result is given in Section 2 while, in Section 3, examples illustrate the theory. Section 4 contains some concluding remarks and the last section is an appendix.

2. Main result

2.1. Risk considerations

Any estimator $\delta = \delta(X,U)$ of θ is evaluated by its risk associated to the loss (1.3), that is by

$$R(\theta,\eta,\delta) = E_{\theta,\eta}\left[\eta\,\|\delta(X,U)-\theta\|^2\right], \tag{2.4}$$

where $E_{\theta,\eta}$ denotes the expectation with respect to the density (1.1) with $\eta = 1/\sigma^2$. For the rest of this paper, we assume

$$E_{\theta,\eta}\left[\|X-\theta\|^2\right] < \infty, \tag{2.5}$$

which guarantees that the standard estimator X has finite risk and is minimax. As $\delta(X,U)$ can be written as $\delta(X,U) = X + g(X,U)$, the finiteness of its risk is guaranteed by

$$E_{\theta,\eta}\left[\|g(X,U)\|^2\right] < \infty. \tag{2.6}$$

To express the risk difference between $\delta(X,U)$ and X, we introduce first the function F defined, for any $t > 0$, by

$$F(t) = \frac{1}{2}\int_t^\infty f(u)\,du.$$

Note that, according to (2.5), we have

$$c = \int_{\mathbb{R}^{p+k}} F(\|x\|^2 + \|u\|^2)\,dx\,du < \infty.$$

A version of the following lemma can be found in [12].

Lemma 2.1. *Assume that the function $g(x,u)$ is weakly differentiable from $\mathbb{R}^{p+k}$ into $\mathbb{R}^p$. Then*

$$E_{\theta,\eta}\left[(X-\theta)'g(X,U)\right] = cE^*_{\theta,\eta}\left[\mathrm{div}_X g(X,U)\right],$$

where $E^*_{\theta,\eta}$ is the expectation with respect to the density

$$\frac{\eta^{p+k}}{c}F\left(\eta\left(\|x-\theta\|^2 + \|u\|^2\right)\right),$$

provided either of the above expectations exists.

Similarly, for any weakly differentiable function h from $\mathbb{R}^{p+k}$ into $\mathbb{R}^p$,

$$E_{\theta,\eta}\left[U'h(X,U)\right] = c\,E^*_{\theta,\eta}\left[\mathrm{div}_U h(X,U)\right],$$

provided either of these expectations exists.

Thanks to Lemma 2.1, an expression of the risk difference between $\delta(X,U)$ and X is given in the following proposition.

Proposition 2.1. *Assume that $E_{\theta,\eta}\left[\|g(X,U)\|^2\right] < \infty$. The risk difference $\triangle_{\theta,\tau}$ between $\delta(X,U) = X + g(X,U)$ and X equals*

$$\triangle_{\theta,\eta} = \mathcal{R}(\theta,\eta,\delta) - \mathcal{R}(\theta,\eta,X) = c\,\eta E^*_{\theta,y}\left[\mathcal{O}g(X,U)\right],$$

where

$$\mathcal{O}g(X,U) = 2\,\mathrm{div}_X g(X,U) + \frac{k-2}{\|U\|^2}\|g(X,\|U\|^2)\|^2 + \frac{U'}{\|U\|^2}\nabla_U\|g(X,U)\|^2. \tag{2.7}$$

Proof. A straightforward calculation gives

$$\begin{aligned}\triangle_{\theta,\eta} &= \eta\, E_{\theta,\eta}\left[2\,(X-\theta)'g(X,U) + \|g(X,U)\|^2\right] \\ &= \eta\, E_{\theta,\eta}\left[2\,(X-\theta)'g(X,U) + U'\frac{U}{\|U\|^2}\|g(X,U)\|^2\right].\end{aligned}$$

Using Lemma 2.1 on each term in the brackets, we obtain

$$\begin{aligned}\triangle_{\theta,\eta} &= \eta\, c\, E^*_{\theta,\eta}\left[2\,\mathrm{div}_X g(X,U) + \mathrm{div}_U\left(\frac{U}{\|U\|^2}\,\|g(X,U)\|^2\right)\right] \\ &= \eta\, c\, E^*_{\theta,\eta}\left[2\,\mathrm{div}_X g(X,U) + \frac{k-2}{\|U\|^2}\,\|g(X,U)\|^2 + \frac{U'}{\|U\|^2}\,\nabla_U\|g(X,U)\|^2\right]\end{aligned}$$

by the divergence formula. □

2.2. Form of the Bayes estimators

We will see that, for priors of the form (1.2), the generalized Bayes estimators do not depend on the density (1.1); more precisely their expressions depend only on π and b provided that

$$\int_0^\infty f(\tau)\tau^{(p+k)/2+b+1}\,d\tau < \infty, \tag{2.8}$$

which is equivalent to

$$E_{0,1}\left[(\|X\|^2 + \|U\|^2)^{2(b+2)}\right] < \infty.$$

Proposition 2.2. *For a prior of the form (1.2), the generalized Bayes estimator $\delta(X,U) = X + g(X,U)$ is such that, for any $(x,u) \in \mathbb{R}^p \times \mathbb{R}^k$,*

$$g\,(x,u) = \frac{\int_{\mathbb{R}^p} \frac{\theta - x}{(\|x-\theta\|^2+\|u\|^2)^{(p+k)/2+b+2}}\,\pi(\|\theta\|^2)\,d\theta}{\int_{\mathbb{R}^p} \frac{1}{(\|x-\theta\|^2+\|u\|^2)^{(p+k)/2+b+2}}\,\pi(\|\theta\|^2)\,d\theta}, \tag{2.9}$$

provided (2.8) holds and (2.9) exists.

Remark 2.1. Note that $g(x,u)$ in (2.9) depends on u only through $\|u\|^2$ so that we write $g(x,u) = g(x,\|u\|^2)$. Note also that it arises as

$$\frac{\nabla_x M(x,\|u\|^2)}{m(x,\|u\|^2)},$$

where $m(x,\|u\|^2)$ is the marginal associated to π and the density

$$\varphi\left(\|x-\theta\|^2 + \|u\|^2\right) \;\propto\; \frac{1}{(\|x-\theta\|^2 + \|u\|^2)^{(p+k)/2+b+2}}, \tag{2.10}$$

and M is the marginal associated to ϕ with

$$\phi(t) = \frac{1}{2}\int_t^\infty \varphi(v)\,dv. \tag{2.11}$$

Therefore, for each fixed u, $\delta(X,u) = X + g(X,u)$ with $g(X,u)$ in (2.9) can be interpreted as the Bayes estimator of θ under the density φ and the prior π for fixed scale parameter ($\|u\|$). This observation will be important in the next section since it will allow us to use results in [9] which are developed for the case of known scale parameter.

Finally, note that existence of (2.9) will be guaranteed by the stronger finiteness risk condition developed in the proof of Theorem 2.1 in the appendix. More generally, it suffices that π be locally integrable and have tails that do not grow too fast at infinity. In particular, superharmonic priors are locally integrable and have bounded tails.

Proof of Proposition 2.2. The Bayes estimator under loss (1.3) is

$$\delta(X,U) = \frac{E[\eta\,\theta|X,U]}{E[\eta|X,U]} = X + g(X,\|U\|^2),$$

with, for any $(x,u) \in \mathbb{R}^p \times \mathbb{R}^k$,

$$\begin{aligned} g(x,\|u\|^2) &= \frac{E[\eta\,(\theta - x) \mid x,u]}{E[\eta|x,u]} \\ &= \frac{\int_0^\infty \int_{\mathbb{R}^p} \eta\,(\theta - x)\,\eta^{(p+k)/2}\, f(\eta\,(\|x-\theta\|^2 + \|u\|^2))\,\pi(\|\theta\|^2)\,\eta^b\, d\theta\, d\eta}{\int_0^\infty \int_{\mathbb{R}^p} \eta^{(p+k)/2+1}\, f(\eta\,(\|x-\theta\|^2 + \|u\|^2))\,\pi(\|\theta\|^2)\eta^b\, d\theta\, d\eta} \\ &= \frac{\int_{\mathbb{R}^p} \int_0^\infty \eta^{(p+k)/2+b+1}\, f(\eta\,(\|x-\theta\|^2 + \|u\|^2))\, d\eta\,(\theta - x)\,\pi(\|\theta\|^2)\, d\theta}{\int_{\mathbb{R}^p} \int_0^\infty \eta^{(p+k)/2+b+1}\, f(\eta\,(\|x-\theta\|^2 + \|u\|^2))\, d\,\eta\,\pi(\|\theta\|^2)\, d\theta}, \end{aligned}$$

by Fubini's theorem. Now, through the change of variable $\tau = \eta\,(\|x-\theta\|^2 + \|u\|^2)$ in the innermost integrals, we obtain

$$\begin{aligned} g(x,\|u\|^2) &= \frac{\int_{\mathbb{R}^p} \int_0^\infty \tau^{(p+k)/2+b+1}\, f(\tau)\, d\tau\, \frac{(\theta-x)\,\pi(\|\theta\|^2)}{(\|x-\theta\|^2+\|u\|^2)^{(p+k)/2+b+2}}\, d\theta}{\int_{\mathbb{R}^p} \int_0^\infty \tau^{(p+k)/2+b+1}\, f(\tau)\, d\tau \frac{\pi(\|\theta\|^2)}{(\|x-\theta\|^2+\|u\|^2)^{(p+k)/2+b+2}}\, d\theta} \\ &= \frac{\int_{\mathbb{R}^p} \frac{(\theta-x)\,\pi(\|\theta\|^2)}{(\|x-\theta\|^2+\|u\|^2)^{(p+k)/2+b+2}}\, d\theta}{\int_{\mathbb{R}^p} \frac{\pi(\|\theta\|^2)}{(\|x-\theta\|^2+\|u\|^2)^{(p+k)/2+b+2}}\, d\theta} \end{aligned}$$

thanks to (2.8). □

2.3. Minimaxity of generalized Bayes estimators

According to the expression of $g(X,U)$ in Remark 2.1, we give an expression of the differential operator $\mathcal{O}g(X,U)$ in (2.7).

Proposition 2.3. *For* $g(X,\|U\|^2) = \frac{\nabla_X M(X,\|U\|^2)}{m(X,\|U\|^2)}$*, (2.7) can be expressed as*

$$\begin{aligned} \mathcal{O}g(X,\|U\|^2) &= 2\,\frac{\Delta_X M(X,\|U\|^2)}{m(X,\|U\|^2)} - 2\,\frac{\nabla_X m(X,\|U\|^2)' \nabla_X M(X,\|U\|^2)}{m^2(X,\|U\|^2)} \\ &\quad + \frac{k-2}{\|U\|^2}\left\|\frac{\nabla_X M(X,\|U\|^2)}{m(X,\|U\|^2)}\right\|^2 + 2\,\frac{\partial}{\partial s}\left\|\frac{\nabla_X M(X,s)}{m(X,s)}\right\|^2\Bigg|_{s=\|U\|^2}, \end{aligned} \tag{2.12}$$

where, for any $(x,u) \in \mathbb{R}^p \times \mathbb{R}^k$*,*

$$m(x, \|u\|^2) = \int_{\mathbb{R}^p} \varphi(\|x-\theta\|^2 + \|u\|^2)\, \pi(\|\theta\|^2)\, d\theta, \tag{2.13}$$

and

$$M(x, \|u\|^2) = \int_{\mathbb{R}^p} \phi(\|x-\theta\|^2 + \|u\|^2)\, \pi(\|\theta\|^2)\, d\theta \tag{2.14}$$

with φ *and* ϕ *given by (2.10) and (2.11).*

Proof. The proof of Proposition 2.3 follows from straightforward calculations. □

[9] studied Bayes minimax estimation of a location vector in the case of spherically symmetric distributions with known scale parameter. For a subclass of spherically symmetric densities, they proved minimaxity of generalized Bayes estimators for spherically symmetric priors of the form $\pi(\|\theta\|^2)$ under the following assumptions.

Assumptions 2.1. (a) $\pi'(\|\theta\|^2) \;\le 0$ i.e. $\pi(\|\theta\|^2)$ is unimodal;
(b) $\Delta\pi(\|\theta\|^2) \;\le 0$ i.e. $\pi(\|\theta\|^2)$ is superharmonic;
(c) $\Delta\pi(\|\theta\|^2)$ is non decreasing in $\|\theta\|^2$.

Note that Condition (b) in fact implies Condition (a) by the mean value property of superharmonic functions.

Our main result below is that a generalized Bayes estimator of θ for a density (1.1), a prior (1.2) and the loss (1.3) is minimax under weak moment conditions and conditions on b, provided the prior satisfies Assumptions (2.1). We remind the reader that, according to Proposition 2.2, the generalized Bayes estimator is independent of the sampling density, f, provided the assumption (2.8) holds. Hence, each such estimator is simultaneously generalized Bayes and minimax for the entire class of spherically symmetric distributions.

Before developing our minimaxity result, we give a theorem which guarantees the risk finiteness of the generalized Bayes estimators.

Theorem 2.1. *Assume that* π *satisfies Assumption 2.1.b; assume also that* $b > -(\frac{k}{2}+1)$*. Then the generalized Bayes estimator associated to* π *has finite risk.*

Proof. The proof is postponed to the Appendix. □

We will need the following result which essentially gathers results in Lemmas 3.1–3.3 of [9].

Lemma 2.2. *Let* $m(x, \|u\|^2)$ *and* $M(x, \|u\|^2)$ *be as defined in (2.13) and (2.14) and let* $\cdot$ *be the inner product in* $\mathbb{R}^p$*. Then we have*

(1)

$$x \cdot \nabla_x m(x, \|u\|^2) = -2 \int_0^\infty H(v, \|x\|^2)\, v^{p/2}\, \varphi'(v + \|u\|^2)\, dv,$$

and

$$x \cdot \nabla_x M(x, \|u\|^2) = \int_0^\infty H(v, \|x\|^2) v^{p/2}\, \varphi(v + \|u\|^2)\, dv,$$

where, for $v > 0$*,*

$$H(v, \|x\|^2) = \lambda(B) \int_{B_{\sqrt{v},x}} x \cdot \theta\, \pi'(\|\theta\|^2)\, dV_{\sqrt{v},x}(\theta) \tag{2.15}$$

and $V_{\sqrt{v},x}$ is the uniform distribution on the ball $B_{\sqrt{v},x}$ of radius $\sqrt{v}$ centered at x and $\lambda(B)$ is the volume of the unit ball.

(2) for any $x \in \mathbb{R}^p$, the function $H(v, \|x\|^2)$ in (2.15) is non decreasing in v provided that $\Delta\pi(\|\theta\|^2)$ is non decreasing in $\|\theta\|^2$. (Assumption 2.1.c).

(3) for any $v > 0$ and any $x \in \mathbb{R}^p$, the function $H(v, \|x\|^2)$ in (2.15) is non positive provided $\pi'(\|\theta\|^2) \le 0$. (Assumption 2.1.a).

Given these preliminaries, we present our main result.

Theorem 2.2. *Suppose that π satisfies Assumptions 2.1. Then the generalized Bayes estimator associated to $\pi(\|\theta\|^2)\,\eta^b$ is minimax provided that $b \ge \frac{2p-k-2}{4}$. (Note that assumptions of Theorem 2.1 are satisfied.)*

Proof. It suffices to show that $\mathcal{O}g(X, U)$ in (2.12), with $m(X, \|U\|^2)$ and $M(X, \|U\|^2)$ given respectively by (2.13) and (2.14), is non positive since the assumptions guarantee that the generalized Bayes estimator δ is of the form $\delta(X, U) = X + \frac{\nabla_X M(X,\|U\|^2)}{m(X,\|U\|^2)}$ and has finite risk.

Due to the superharmonicity of $\pi(\|\theta\|^2)$, for any $(x, u) \in \mathbb{R}^p \times \mathbb{R}^k$, we have $\Delta_x M(x\|u\|^2) \le 0$ so that

$$\mathcal{O}g(x, \|u\|^2) \le -2\,\frac{\nabla_x m(x, \|u\|^2)'\nabla_x M(x, \|u\|^2)}{m^2(x, \|u\|^2)} + \frac{k-2}{\|u\|^2}\left\|\frac{\nabla_x M(x, \|u\|^2)}{m(x, \|u\|^2)}\right\|^2 + 2\frac{\partial}{\partial s}\left\|\frac{\nabla_x M(x, s)}{m(x, s)}\right\|^2\Bigg|_{s=\|u\|^2}.$$

Note that

$$\begin{aligned} m^2(x, s)\frac{\partial}{\partial s}\left\|\frac{\nabla_x M(x, s)}{m(x, s)}\right\|^2 &= \frac{\partial}{\partial s}\|\nabla_x M(x, s)\|^2 + \|\nabla_x M(x, s)\|^2\, m^2(x, s)\frac{\partial}{\partial s}\frac{1}{m^2(x, s)} \\ &\le \frac{\partial}{\partial s}\|\nabla_x M(x, s)\|^2 + (p + k + 2b + 4)\frac{1}{s}\|\nabla_x M(x, s)\|^2, \end{aligned}$$

since

$$\begin{aligned} \frac{\partial}{\partial s}\frac{1}{m^2(x, s)} &= \frac{-2}{m^3(x, s)}\int_{\mathbb{R}^p}\frac{-[(p+k)/2 + b + 2]}{(\|x-\theta\|^2 + s)^{(p+k)/2+b+3}}\,\pi(\|\theta\|^2)\,d\theta \\ &= \frac{p + k + 2b + 4}{m^3(x, s)}\,\frac{1}{s} \\ &\quad \times \int_{\mathbb{R}^p}\frac{s}{\|x-\theta\|^2 + s}\,\frac{1}{(\|x-\theta\|^2 + s)^{(p+k)/2+b+2}}\,\pi(\|\theta\|^2)\,d\theta \\ &\le \frac{p + k + 2b + 4}{m^2(x, s)}\,\frac{1}{s}. \end{aligned}$$

Therefore

$$\begin{aligned} m^2(x, s)\,\mathcal{O}g(x, s) \le &-2\,\nabla_x m(x, s)'\nabla_x M(x, s) \\ &+ \frac{k - 2 + 2(p + k + 2b + 4)}{s}\|\nabla_x M(x, s)\|^2 \\ &+ 2\frac{\partial}{\partial s}\|\nabla_x M(x, s)\|^2. \end{aligned} \tag{2.16}$$

As $m(x,s)$ and $M(x,s)$ depend on x only through $\|x\|^2$, it is easy to check that (as in [9])

$$\nabla_x m(x,s)' \nabla_x M(x,s) = \frac{x' \nabla_x m(x,s)\, x' \nabla_x M(x,s)}{\|x\|^2}$$

and

$$\|\nabla_x M(x,s)\|^2 = \frac{(x' \nabla_x M(x,s))^2}{\|x\|^2}.$$

Thus the right hand side of (2.16) will be non positive as soon as

$$-2x' \nabla_x m(x,s) + \frac{2p+3k+4b+6}{s}\, x' \nabla_x M(x,s) + 4 \frac{\partial}{\partial s} x' \nabla_x M(x,s) \geq 0, \tag{2.17}$$

since, according to Lemma 2.2, the common factor $x' \nabla_x M(x,s)$ is non positive. Using again Lemma 2.2, the left hand side of (2.17) equals

$$\begin{aligned} & 4 \int_0^\infty H(v, \|x\|^2)\, v^{p/2}\, \varphi'(v+s)\, dv \\ & \quad + \frac{2p+3k+4b+6}{s} \int_0^\infty H(v, \|x\|^2)\, v^{p/2}\, \varphi(v+s)\, dv \\ & \quad + 4 \int_o^\infty H(v, \|x\|^2) v^{p/2} \varphi'(v+s) dv \\ & \quad = \int_0^\infty v^{p/2}\, \varphi(v+s)\, dv \Big\{ 8E \left[H(v, \|x\|^2) \frac{\varphi'(v+s)}{\varphi(v+s)} \right] \\ & \qquad + \frac{2p+3k+4b+6}{s} E\left[H(v, \|x\|^2) \right] \Big\}, \end{aligned} \tag{2.18}$$

where E denotes the expectation with respect to the density proportional to $v \longmapsto v^{p/2} \varphi(v+s)$.

As

$$\frac{\varphi'(v+s)}{\varphi(v+s)} = \frac{-((p+k)/2+b+2)}{\mathrm{v}+s} \tag{2.19}$$

is non decreasing in v and, according to Lemma 2.2, $H(v, \|x\|^2)$ is also non decreasing in v, the first expectation in (2.18) satisfies

$$E\left[H(v, \|x\|^2) \frac{\varphi'(v+s)}{\varphi(v+s)} \right] \geq E\left[H(v, \|x\|^2) \right] E\left[\frac{\varphi'(v+s)}{\varphi(v+s)} \right]$$

by the covariance inequality. Therefore Inequality (2.17) will be satisfied as soon as

$$8\, E\left[\frac{\varphi'(v+s)}{\varphi(v+s)} \right] + \frac{2p+3k+4b+6}{s} \leq 0, \tag{2.20}$$

since $H(v, \|x\|^2) \leq 0$ by Lemma 2.2.

From (2.19) we have

$$(2.21)\quad E\left[\frac{\varphi'(v+s)}{\varphi(v+s)}\right] = -\big((p+k)/2+b+2\big)E\left[\frac{1}{v+s}\right]$$

$$= -\big((p+k)/2+b+2\big)\,\frac{\int_o^\infty \frac{1}{v+s}\, v^{p/2}\frac{1}{(v+s)^{(p+k)/2+b+2}}\,dv}{\int_0^\infty v^{p/2}\frac{1}{(v+s)^{(p+k)/2+b+2}}\,dv}$$

$$= -\big((p+k)/2+b+2\big)\,\frac{1}{s}\,\frac{\int_0^\infty \frac{z^{p/2}}{(z+1)^{(p+k)/2+b+3}}\,dz}{\int_0^\infty \frac{z^{p/2}}{(z+1)^{(p+k)/2+b+2}}\,dz}$$

$$= -\big((p+k)/2+b+2\big)\,\frac{1}{s}\,\frac{B(p/2+1,k/2+b+2)}{B(p/2+1,k/2+b+1)},$$

where $B(\alpha,\beta)$ is the beta function with parameters $\alpha > 0$ and $\beta > 0$. Then (2.21) becomes

$$(2.22)\quad \begin{aligned} E\left[\frac{\varphi'(v+s)}{\varphi(v+s)}\right] &= -\frac{((p+k)/2+b+2)}{s} \\ &\quad\times \frac{\Gamma((k/2+b+2))}{\Gamma((p+k)/2+b+3)}\,\frac{\Gamma((p+k)/2+b+2)}{\Gamma(k/2+b+1)} \\ &= \frac{-(k/2+b+1)}{s}. \end{aligned}$$

It follows from (2.22) that (2.20) reduces to

$$b \geq \frac{2p-k-2}{4},$$

which is the condition given in the theorem. □

The condition on b in Theorem 2.2 can be alternatively expressed as $k \geq 2p - 4b - 2$ which dictates that the dimension, k, of the residual vector, U, increases with the dimension, p, of θ. This dependence can be (essentially) eliminated provided the generalized Bayes estimator in Proposition 2.2 satisfies the following assumption:

Assumptions 2.2. The function $g(x,u)$ in (2.9) can be expressed as

$$g(x,u) = \frac{\nabla_x M(x,\|u\|^2)}{m(x,\|u\|^2)} = -\frac{r(\|x\|^2,\|u\|^2)\|u\|^2}{\|x\|^2}\,x,$$

where $r(\|x\|^2,\|u\|^2)$ is non-negative and non-increasing in $\|u\|^2$.

Assumption 2.2 is satisfied, for example, by the generalized Bayes estimator corresponding to the prior on (θ,η) proportional to $\pi(\|\theta\|^2) = \left(1/\|\theta\|^2\right)^{b/2}\eta^a$ for $0 < b \leq p-2$ and $a > -\frac{k}{2} - \frac{b}{2} - 2$, in which case the function $r(\|x\|^2,\|u\|^2) = \phi\left(\|x\|^2/\|u\|^2\right)$, where $\phi(t)$ is increasing in t, and hence $r(\|x\|^2,\|u\|^2)$ is decreasing in $\|u\|^2$ (see, e.g. 1 [15]).

We have the following corollary:

Corollary 2.1. *Suppose π satisfies Assumptions 2.1 and the assumptions of Theorem 2.2 and suppose also that the generalized Bayes estimator (which does not depend on the underlying density f) satisfies Assumption 2.2. Then the generalized Bayes estimator is minimax provided $b \geq -\frac{k+2}{4}$.*

Proof. Assumption 2.2 guarantees that

$$\frac{\partial}{\partial s}\left(\frac{1}{s^2}\left\|\frac{\nabla_x M(x,s)}{m(x,s)}\right\|^2\right) = \frac{\partial}{\partial s}\left(\frac{r^2(\|x\|^2,s)}{\|x\|^2}\right) \le 0.$$

Since

$$\begin{aligned}\frac{\partial}{\partial s}\left\|\frac{\nabla_x M(x,s)}{m(x,s)}\right\|^2 &= \frac{\partial}{\partial s}\left(\frac{s^2}{s^2}\left\|\frac{\nabla_x M(x,s)}{m(x,s)}\right\|^2\right)\\ &= \frac{2}{s}\left\|\frac{\nabla_x M(x,s)}{m(x,s)}\right\|^2 + s^2\frac{\partial}{\partial s}\left(\frac{1}{s^2}\left\|\frac{\nabla_x M(x,s)}{m(x,s)}\right\|^2\right),\end{aligned}$$

the inequality for $\mathcal{O}g(X,\|U\|^2)$ in the proof of Theorem 2.2 can be replaced by

$$\mathcal{O}g(x,\|u\|^2) \le -2\,\frac{\nabla_x m(x,\|u\|^2)'\nabla_x M(x,\|u\|^2)}{m^2(x,\|u\|^2)} + \frac{k+2}{\|u\|^2}\left\|\frac{\nabla_x M(x,\|u\|^2)}{m(x,\|u\|^2)}\right\|^2.$$

It follows that inequality condition (2.17) becomes

$$-2\,x'\nabla_x m(x,s) + \frac{k+2}{s}x'\nabla_x M(x,s) \ge 0,$$

and that inequality condition (2.20) becomes

$$4\,E\left[\frac{\varphi'(v+s)}{\varphi(v+s)}\right] + \frac{k+2}{s} \le 0,$$

which, by (2.22), becomes

$$4\left[-\left(\frac{k/2+b+1}{s}\right)\right] + \frac{k+2}{s} \le 0,$$

which is equivalent to $b \ge -(k+2)/4$. □

3. Examples

Several examples of priors which satisfies Assumptions 2.1 are given in [9]. We briefly summarize these.

Example 1 (priors related to the fundamental harmonic prior). Let $\pi(\|\theta\|^2) = (\frac{1}{A+\|\theta\|^2})^c$ with $A \ge 0$ and $0 \le c \le \frac{p}{2} - 1$.

Example 2 (mixtures of priors satisfying Assumption 2.2). Let $(\pi_\alpha)_{\alpha\epsilon A}$ be a family of priors such that Assumption 2.1 is satisfied for any $\alpha \in A$. Then any mixture of the form $\int_Y \pi_\alpha(\|\theta\|^2)\,dH(\alpha)$ where H is a probability on Y satisfies Assumption 2.1 as well. For instance, Example 1 with $c=1, p\ge 4, A=\alpha$ and the gamma density $\alpha \longmapsto \frac{\beta^{1-v}}{\Gamma(1-v)}\alpha^{-v}e^{-\beta\alpha}$ with $\beta > 0$ and $0 < v < 1$ leads to the prior

$$\|\theta\|^{-2-v}\,e^{\beta\|\theta\|^2}\,\Gamma(v,\beta\|\theta\|^2),$$

where

$$\Gamma(v,y) = \int_y^\infty e^{-x}x^{v-1}\,dx$$

is the complement of the incomplete gamma function.

Example 3 (variance mixtures of normals). Let

$$\pi(\|\theta\|^2) = \int_O^\infty \left(\frac{u}{2\pi}\right)^{p/2} \exp\left(\frac{-u\|\theta\|^2}{2}\right) h(u)\, du$$

a mixture of normals with respect to the inverse of the variance. As soon as, for any $u > 0$,

$$\frac{uh'(u)}{h(u)} \leq -2,$$

the prior $\pi(\|\theta\|^2)$ satisfies Assumptions 2.1. Note that the priors in Example 1 arise as such a mixture with $h(u) \propto \alpha u^{k-p/2-1} \exp(-\frac{A}{2}u)$.

Other examples can be given and a constructive approach is proposed in [9].

4. Concluding remarks

This paper shows that generalized Bayes estimators corresponding to priors of the form $\pi(\|\theta\|^2)\eta^b$ are robust in that they are independent of the underlying density $\eta^{(p+k)/2} f(\eta(\|x-\theta\|^2 + \|u\|^2))$. Furthermore, provided $\pi(\|\theta\|^2)$ is superharmonic with a non decreasing Laplacian, under weak moment conditions and conditions on b, these generalized Bayes estimator are minimax for the entire class of spherically symmetric distributions.

[15] developed similar results for priors of the form $\|\theta\|^{-a}\,\eta^b$ for $0 \leq a \leq p-2$ using techniques developed for Baranchik type estimators. This paper relies on techniques developed for the known scale case in [9] and its main contribution is to extend the minimaxity result to priors of the more general form $\pi(\|\theta\|^2)\,\eta^b$ where $\pi(\|\theta\|^2)$ is superharmonic, but not necessarily homogeneous.

A critical difference between the results of this paper and [9] (and other papers studying minimaxity in the known scale case) is that the minimax generalized Bayes estimators in the known scale case depend on the form of the underlying density $\eta^p\, f(\eta\,\|x-\theta\|^2)$, and of course, also on the scale. On the other hand, when the scale is unknown and residual vector is available, the form of the generalized Bayes estimator is independent of the form of the density $f(\cdot)$ provided the prior distribution has the form $\pi(\theta)\,\eta^b$. It is not necessary that $\pi(\theta)$ be spherically symmetric. It would be desirable to extend the minimaxity results to priors which are not necessarily spherically symmetric.

Appendix: Proof of Theorem 2.1

According to (2.9), the risk finiteness condition (2.6) is satisfied as soon as

$$\begin{aligned} &E_{\theta,\eta}\left[\left\|\frac{\displaystyle\int_{R^p}(\theta-X)\frac{\pi(\|\theta\|^2)}{(\|X-\theta\|^2+\|U\|^2)^{(p+k)/2+b+2}}\,d\theta}{\displaystyle\int_{R^p}\frac{\pi(\|\theta\|^2)}{(\|X-\theta\|^2+\|U\|^2)^{(p+k)/2+b+2}}\,d\theta}\right\|^2\right] \\ &\leq E_{\theta,\eta}\left[\frac{\displaystyle\int_{R^p}\|\theta-X\|^2\frac{\pi(\|\theta\|^2)}{(\|X-\theta\|^2+\|U\|^2)^{(p+k)/2+b+2}}\,d\theta}{\displaystyle\int_{R^p}\frac{\pi(\|\theta\|^2)}{(\|X-\theta\|^2+\|U\|^2)^{(p+k)/2+b+2}}\,d\theta}\right] \\ &< \infty. \end{aligned} \tag{A1}$$

Note that, for any $(x,u) \in \mathbb{R}^p \times \mathbb{R}^k$ and for any non negative function h on $\mathbb{R}_+ \times \mathbb{R}_+$,

$$\begin{aligned}\text{(A2)} \qquad & \int_{R^p} \pi(\|\theta\|^2)\, h(\|x-\theta\|^2, \|u\|^2)\, d\theta \\ & = \int_0^\infty \int_{S_{R,x}} \pi(\|\theta\|^2)\, d\mathcal{U}_{R,x}(\theta)\, \sigma(S)\, R^{p-1} h(R^2, \|u\|^2)\, dR,\end{aligned}$$

where $\mathcal{U}_{R,x}$ is the uniform distribution on the sphere $S_{R,x}$ of radius R and centered at x and $\sigma(S)$ is the area of the unit sphere. Through the change of variable $R = \sqrt{v}$, the right hand side of (A2) can be written as

$$\int_0^\infty \mathcal{S}_\pi(\sqrt{v}, x)\, v^{p/2-1}\, h(v, \|U\|^2)\, dv,$$

where

$$\mathcal{S}_\pi(\sqrt{v}, x) = \frac{\sigma(S)}{2} \int_{S_{\sqrt{v},x}} \pi(\|\theta\|^2)\, d\mathcal{U}_{\sqrt{v},x}(\theta)$$

is non increasing in v by the superharmonicity of $\pi(\|\theta\|^2)$.

Now we can express the last quantity in brackets in (A1) as

$$\begin{aligned}\text{(A3)} \qquad & \frac{\int_o^\infty \mathcal{S}_\pi(\sqrt{v}, x)\ \frac{v^{p/2}}{(v+\|u\|^2)^{(p+k)/2+b+2}}\, dv}{\int_o^\infty \mathcal{S}_\pi(\sqrt{v}, x)\ \frac{v^{p/2-1}}{(v+\|u\|^2)^{(p+k)/2+b+2}}\, dv} \\ & = E_1[v] \\ & \le E_2[v],\end{aligned}$$

where E_1 is the expectation with respect to the density $f_1(v)$ proportional to

$$\mathcal{S}_\pi(\sqrt{v}, x)\, \frac{v^{p/2-1}}{(v+\|u\|^2)^{(p+k)/2+b+2}},$$

and E_2 is the expectation with respect to the density $f_2(v)$ proportional to

$$\frac{v^{p/2-1}}{(v+\|u\|^2)^{(p+k)/2+b+2}}.$$

Indeed the ratio $\frac{f_2(v)}{f_1(v)}$ is non decreasing by the monotonicity of $\mathcal{S}_\pi(\sqrt{v}, x)$. In (A3), $E_2[v]$ is

$$\begin{aligned} E_2[v] &= \frac{\int_0^\infty \frac{v^{p/2}}{(v+\|u\|^2)^{(p+k)/2+b+2}}\, dv}{\int_0^\infty \frac{v^{p/2-1}}{(v+\|u\|^2)^{(p+k)/2+b+2}}\, dv} \\ &= \|u\|^2\ \frac{\int_0^\infty \frac{v^{p/2}}{(v+1)^{(p+k)/2+b+2}}\, dv}{\int_0^\infty \frac{v^{p/2-1}}{(v+1)^{(p+k)/2+b+2}}\, dv} \\ &= \|u\|^2\ \frac{B(p/2+1, k/2+b+1)}{B(p/2, k/2+b+2)},\end{aligned}$$

which is finite for $k/2+b+1>0$.

Finally the expectations in (A1) are bounded above by $K\, E_{\theta,\eta}[\|U\|^2]$ where K is a constant, and hence are finite. □

References

[1] Alam, K. (1973). A family of admissible minimax estimators of the mean of a multivariate normal distribution. *Ann. Statist.* **1** 517–525.

[2] 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.

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

[4] Cellier, D. and Fourdrinier, D. (1995). Shrinkage estimators under spherically symmetry for the general linear model. *J. Multivariate Anal.* **52** 338–351.

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

[6] Chou, J. P. and Strawderman, W. E. (1990). Minimax estimation of means of multivariate normal mixtures. *J. Multivariate Anal.* **35** 141–150.

[7] Faith, R. E. (1978). Minimax Bayes point estimators of a multivariate normal mean. *J. Multivariate Anal.* **8** 372–379.

[8] Fourdrinier, D., Kortbi, O. and Strawderman, W. E. (2008). Bayes minimax estimators of the mean of a scale mixture of multivariate normal distributions. *J. Multivariate Anal.* **99** 1 74–93.

[9] Fourdrinier, D. and Strawderman, W. E. (2008). Generalized Bayes minimax estimators of location vector for spherically symmetric distributions. *J. Multivariate Anal.* **9** 4 735–750.

[10] Fourdrinier, D. and Strawderman, W. E. (2008). A unified and generalized set of shrinkage bounds on minimax Stein estimates. *J. Multivariate Anal.* **99** 10 2221–2233.

[11] Fourdrinier, D., Strawderman, W. E. and Wells, M. T. (1998). On the construction of Bayes minimax estimators. *Ann. Statist.* **26** 2 660–671.

[12] Fourdrinier, D., Strawderman, W. E. and Wells, M. T. (2003). Robust shrinkage estimation for elliptically symmetric distributions with unknown covariance matrix. *J. Multivariate Anal.* **85** 24–39.

[13] Fourdrinier, D., Strawderman, W. E. and Wells, M. T. (2006). Estimation of a location parameter with restrictions or "vague information" for spherically symmetric distributions. *Ann. Inst. Statist. Math.* **58** 73–92.

[14] Kubokawa, T. (1991). An approach to improving the James-Stein estimator. *J. Multivariate Anal.* **36** 121–126.

[15] Maruyama, Y. (2003). A robust generalized Bayes estimator improving on the James-Stein estimator for spherically symmetric distributions. *Statist. Decisions* **21** 69–78.

[16] Maruyama, Y. (2003). Admissible minimax estimators of a mean vector of scale mixtures of multivariate normal distributions. *J. Multivariate Anal.* **84** 274–283.

[17] Maruyama, Y. and Strawderman, W. E. (2005). A new class of generalized Bayes minimax ridge regression estimators. *Ann. Statist.* **33** 1753–1770.

[18] Maruyama, Y. and Takemura, A. (2008). Admissibility and minimaxity of generalized Bayes estimators for spherically symmetric family. *J. Multivariate Anal.* **99** 50–73.

[19] Maruyama, Y. and Strawderman, W. E. (2009). An extended class of minimax generalized Bayes estimators of regression coefficients. *J. Multivariate*

Anal. To appear.

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

[21] Strawderman, W. E. (1971). Proper Bayes minimax estimators of the multivariate normal mean. *The Annals of Mathematical Statistics* **42** 385–388.

[22] Strawderman, W. E. (1973). Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Statist.* **1** 1189–1194.

[23] Strawderman, W. E. (1974). Minimax estimation of location parameters for certain spherically symmetric distributions. *J. Multivariate Anal.* **4** 255–264.