Journal of Multivariate Analysis 100 (2009) 2155–2166

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

An extended class of minimax generalized Bayes estimators of regression coefficients

Yuzo Maruyama [a,*], William E. Strawderman [b]

[a] *University of Tokyo, Japan*
[b] *Rutgers University, United States*

ARTICLE INFO

Article history:
Received 9 March 2008
Available online 6 June 2009

AMS subject classifications:
primary 62C20
secondary 62J07

Keywords:
Regression
Minimaxity
Shrinkage estimators
Generalized Bayes estimators
Invariant loss
Unknown variance
Hypergeometric function

ABSTRACT

We derive minimax generalized Bayes estimators of regression coefficients in the general linear model with spherically symmetric errors under invariant quadratic loss for the case of unknown scale. The class of estimators generalizes the class considered in Maruyama and Strawderman [Y. Maruyama, W.E. Strawderman, A new class of generalized Bayes minimax ridge regression estimators, Ann. Statist., 33 (2005) 1753–1770] to include non-monotone shrinkage functions.

1. Introduction

In this paper we consider minimax generalized Bayes estimators of the regression coefficients in the general linear model with homogeneous spherically symmetric errors. We start with the familiar linear regression model $Y = A\beta + \epsilon$ where Y is an $N \times 1$ vector of observations, A is the known $N \times p$ design matrix of rank p, β is the $p \times 1$ vector of unknown regression coefficients, and ϵ is an $N \times 1$ vector of experimental errors. We assume that ϵ has a spherically symmetric distribution with a density $\sigma^{-N} f(\epsilon'\epsilon/\sigma^2)$, where σ is an unknown scale parameter and $f(\cdot)$ is a nonnegative function on the nonnegative real line, which satisfies

$$\int_0^\infty t^{N/2-1} f(t) \mathrm{d}t < \infty. \tag{1.1}$$

The problem is to estimate β. The least squares estimator of β is $\hat{\beta} = (A'A)^{-1}A'y$. In order to treat the estimation problem from the decision-theoretic point of view, we measure the loss in estimating β by b with the so-called scale invariant "predictive loss" functions

$$L(b, \beta, \sigma^2) = \sigma^{-2}(b - \beta)'A'A(b - \beta). \tag{1.2}$$

* Corresponding address: University of Tokyo, Center for Spatial Information Science, Sougoukenkyuto 403, 113-0033 Tokyo, Japan.
E-mail addresses: maruyama@csis.u-tokyo.ac.jp (Y. Maruyama), straw@stat.rutgers.edu (W.E. Strawderman).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2009.06.001

Then the risk function of an estimator b is given by $R(b, \beta, \sigma^2) = E[L(b, \beta, \sigma^2)]$. The least squares estimator $\hat{\beta}$ is minimax with constant risk. Therefore, b is a minimax estimator of β if and only if $R(b, \beta, \sigma^2) \le R(\hat{\beta}, \beta, \sigma^2)$ for all β and σ^2, and the search for estimators better than $\hat{\beta}$ is a search for minimax estimators. It is, of course, well known that estimators dominating $\hat{\beta}$ exist when $p \ge 3$. In this paper, we study generalized Bayes minimax estimators.

To simplify expressions and to make matters a bit clearer it is helpful to rotate the problem via the following transformation, so that the matrix $(A'A)^{-1}$, which is proportional to the covariance matrix of $\hat{\beta}$ for the general situation, becomes diagonal. Let P be the $p \times p$ orthogonal matrix of eigenvectors of $(A'A)^{-1}$, with $d_1 \ge d_2 \ge \cdots \ge d_p$ as eigenvalues, it follows that

$$P'(A'A)^{-1}P = D, \qquad P'P = I_p$$

where $D = \mathrm{diag}(d_1, \ldots, d_p)$. Let also Q be an $N \times N$ orthogonal matrix such that

$$QA = \begin{pmatrix} D^{-1/2}P' \\ 0 \end{pmatrix}.$$

Next, define two random vectors $X = (X_1, \ldots, X_p)'$ and $Z = (Z_1, \ldots, Z_n)'$ where $n = N - p$ by

$$\begin{pmatrix} X \\ Z \end{pmatrix} = QY.$$

Then $(X', Z')'$ has the joint density given by

$$\sigma^{-p-n} f(\{(x-\theta)'(x-\theta) + z'z\}/\sigma^2), \tag{1.3}$$

where $\theta = D^{-1/2}P'\beta$. Notice also that X and $Z'Z$ can be expressed as $D^{-1/2}P'\hat{\beta}$ and $(y - A\hat{\beta})'(y - A\hat{\beta})$, respectively. Also, as is customary, denote $Z'Z$ by S. We assume throughout this paper that $p \ge 3$ and $n \ge 1$.

The original problem is thus equivalent to estimation of θ under the loss function $L(\delta, \theta, \sigma^2) = (\delta - \theta)'(\delta - \theta)/\sigma^2$. We will consider the problem in this equivalent canonical form.

This paper is best viewed as a companion paper to Maruyama and Strawderman [1]. In that paper a class of minimax generalized Bayes estimators was derived for the above canonical problem when the errors were normally distributed, and a subclass was shown to be generalized Bayes and minimax for the entire class of spherically symmetric error distributions.

In this paper we enlarge the class of generalized Bayes minimax estimators for each of above classes of distributions. To do so we must first enlarge the class of minimax estimators to accommodate non-monotone shrinkage functions. Section 2 is based on this generalization. Section 3 develops an extended class of generalized Bayes minimax estimators for spherical normal error distributions. To a large extend, the results therein represent extensions to the case of unknown variance, of the results of Maruyama [2,3] in the known variance case. Section 4 extends these results to the case of general spherically symmetric error distributions with unknown scale. A subclass of the estimators in Section 3 is shown to be generalized Bayes and minimax for the entire class of spherical error distributions simultaneously (subject to finiteness of moments). We give some numerical results in Section 4.1.

2. Minimaxity

In this section, we give a sufficient condition for minimaxity in the general spherically symmetric case.

Theorem 2.1. *Suppose $(X', Z')'$ has a distribution given by (1.3). Then δ_ϕ given by*

$$\delta_\phi = \left\{1 - \frac{S}{\|X\|^2}\phi\left(\frac{\|X\|^2}{S}\right)\right\} X \tag{2.1}$$

is minimax if

$$\frac{\phi(w)}{w}\{(n+2)\phi(w) - 2(p-2)\} - 4\phi'(w)\{1 + \phi(w)\} \le 0. \tag{2.2}$$

The interesting point of the theorem is that the sufficient condition for minimaxity does not depend on f. Such distributional robustness has already been noted in the literature, but only the following tractable subset of the above result is typically used:

Corollary 2.1. *If ϕ is monotone nondecreasing and $0 \le \phi(w) \le 2(p-2)/(n+2)$, δ_ϕ is minimax in the general spherically symmetric case.*

For completeness we give the proof. We use the version derived in Kubokawa and Srivastava [4] but earlier versions appear in Robert [5] and elsewhere.

Proof. Let

$$F(x)=\frac{1}{2}\int_x^{\infty} f(t)\mathrm{d}t$$

and define

$$E^f[h(X,Z)]=\int\int h(x,z)\sigma^{-N} f\left(\frac{(x-\theta)'(x-\theta)}{\sigma^2}+\frac{z'z}{\sigma^2}\right)\mathrm{d}x\mathrm{d}z$$
$$E^F[h(X,Z)]=\int\int h(x,z)\sigma^{-N} F\left(\frac{(x-\theta)'(x-\theta)}{\sigma^2}+\frac{z'z}{\sigma^2}\right)\mathrm{d}x\mathrm{d}z,$$

where $h(x,z)$ is an integrable function. Note that $F(x)$ is not necessarily a probability density, and hence E^F is not necessarily an expectation symbol in any strict sense.

The identities corresponding to the Stein and chi-square identities for the normal distribution [6,7],

$$E^f[(X_i-\theta_i)h(X,Z)]=\sigma^2E^F[(\partial/\partial X_i)h(X,Z)], \tag{2.3}$$
$$E^f[Sg(S)]=\sigma^2E^F[ng(S)+2Sg'(S)], \tag{2.4}$$

where $S=Z'Z$, are useful in our proof.

The risk of δ_ϕ is given by

$$\begin{aligned}R(\theta,\sigma^2,\delta_\phi)&=E^f\left[(\delta_\phi-\theta)'(\delta_\phi-\theta)/\sigma^2\right]\\&=R(\theta,\sigma^2,X)+E^f\left[\frac{S^2}{\sigma^2\|X\|^2}\phi^2\left(\frac{\|X\|^2}{S}\right)\right]-2E^f\left[\sum\frac{S}{\sigma^2\|X\|^2}X_i(X_i-\theta_i)\phi\left(\frac{\|X\|^2}{S}\right)\right].\end{aligned} \tag{2.5}$$

Let $W=\|X\|^2/S$. For the second term in (2.5), using (2.4), we have

$$E^f\left[\frac{S}{\sigma^2\|X\|^2}\left\{S\phi^2\left(\frac{\|X\|^2}{S}\right)\right\}\right]=E^F\left[(n+2)\frac{\phi^2(W)}{W}-4\phi(W)\phi'(W)\right].$$

For the third term in (2.5), using (2.3), we have

$$\sum E^f\left[\frac{1}{\sigma^2}(X_i-\theta_i)X_i\left(\frac{\|X\|^2}{S}\right)^{-1}\phi\left(\frac{\|X\|^2}{S}\right)\right]=E^F\left[p\frac{\phi(W)}{W}+2\frac{\|X\|^2}{S}\left\{\frac{\phi'(W)}{W}-\frac{\phi(W)}{W^2}\right\}\right].$$

Hence we have

$$R(\theta,\sigma^2,\delta_\phi)-R(\theta,\sigma^2,X)=E^F\left[\frac{\phi(W)}{W}\{(n+2)\phi(W)-2(p-2)\}-4\phi'(W)(1+\phi(W))\right],$$

which completes the proof. □

An immediate corollary, which we believe is often more tractable, is given by dividing the left-hand side of the inequality (2.2) by $\{1+\phi(w)\}\phi(w)/w$.

Corollary 2.2. (i) *δ_ϕ given by (2.1) is minimax if*

$$\frac{(n+2)\phi(w)-2(p-2)}{1+\phi(w)}-4w\frac{\phi'(w)}{\phi(w)}\le 0. \tag{2.6}$$

(ii) *Suppose $0\le\phi(w)\le M_1$ and $w\phi'(w)/\phi(w)\ge -M_2$. The estimator δ_ϕ given by (2.1) is minimax if*

$$\frac{(n+2)M_1-2(p-2)}{1+M_1}+4M_2\le 0.$$

It is clear that part (ii) of the corollary above allows ϕ to be non-monotonic. Actually we show in the next section that there exists a class of minimax generalized Bayes estimators with non-monotone ϕ.

3. An extended class of generalized Bayes minimax estimators in the normal case

In this section, we extend a class of generalized Bayes minimax estimators in Maruyama and Strawderman [1] in the normal case.

Suppose the sampling distribution of $(X', Z')'$ is normal with covariance matrix $\sigma^2 I_n$ and mean vector $(\theta', 0')'$. As in Maruyama and Strawderman [1], the class of hierarchical priors we consider is as follows;

$$\theta|\eta, \lambda \sim N_p(0, \eta^{-1}(1-\lambda)\lambda^{-1} I_p), \;\; \eta \sim \eta^e, \lambda \sim \lambda^a(1-\lambda)^b. \tag{3.1}$$

Then the marginal density of X, S, λ and η is proportional to

$$\begin{aligned}&\int_{R^p} \exp\left(-\frac{\eta}{2}\left\{\|x-\theta\|^2 + \frac{\lambda}{1-\lambda}\|\theta\|^2\right\} - \frac{\eta s}{2}\right) \frac{\eta^{p/2+n/2+e}\lambda^{p/2+a}}{(1-\lambda)^{p/2-b}} \mathrm{d}\theta \\ &\quad \propto \exp\left(-\frac{\eta s}{2}(1+\lambda w)\right)\eta^{p/2+n/2+e}\lambda^{p/2+a}(1-\lambda)^b,\end{aligned} \tag{3.2}$$

where $w = \|x\|^2/s$. Under quadratic loss, the generalized Bayes estimator for such a hierarchical priors can be expressed as (3.2),

$$\hat{\theta} = \frac{E(\eta\theta|X,S)}{E(\eta|X,S)} = \left(1 - \frac{E(\lambda\eta|X,S)}{E(\eta|X,S)}\right) X = \left(1 - \frac{\phi(W)}{W}\right) X. \tag{3.3}$$

When $p/2 + n/2 + e + 2 > 0$,

$$\int_0^\infty \eta^{p/2+n/2+e+1} \exp\left(-\frac{\eta s}{2}(1+\lambda w)\right) \mathrm{d}\eta \propto (1+\lambda w)^{-p/2-n/2-e-2}, \tag{3.4}$$

hence

$$\phi(w) = w\frac{\int_0^1 \lambda^{p/2+a+1}(1-\lambda)^b(1+w\lambda)^{-p/2-n/2-e-2}\mathrm{d}\lambda}{\int_0^1 \lambda^{p/2+a}(1-\lambda)^b(1+w\lambda)^{-p/2-n/2-e-2}\mathrm{d}\lambda}, \tag{3.5}$$

which is well defined for $a > -p/2 - 1$ and $b > -1$.

Here is a result of Maruyama and Strawderman [1].

Theorem 3.1 (*Maruyama and Strawderman [1]*). *Suppose $b \geq 0$, $e > -p/2 - n/2 - 1$ and $-p/2 - 1 < a < n/2 + e$. Then $\phi(w)$ is monotone increasing and approaches $(p/2 + a + 1)/(n/2 + e - a)$ as $w \to \infty$. Hence δ_ϕ with ϕ given by (3.5) is minimax if*

$$0 \leq \frac{p/2+a+1}{n/2+e-a} \leq 2\frac{p-2}{n+2}.$$

The key assumption of the theorem is $b \geq 0$ while the estimator with $\phi(w)$ itself is well defined when $b > -1$. In fact, there is a big difference between the two cases: $b \geq 0$ and $-1 < b < 0$. An immediately apparent difference is that $(1-\lambda)^b$ is unbounded for $-1 < b < 0$ while it is bounded for $b \geq 0$. Technically, because of this unboundedness, the standard integration by parts technique fails to work well when $-1 < b < 0$. Furthermore the shrinkage factor $\phi(w)$ for $-1 < b < 0$ is not often monotonic, which means the tractable sufficient condition for minimaxity given by Corollary 2.1 is not applicable.

We will see below that proving minimaxity with $b < 0$ requires a different approach from that in Maruyama and Strawderman [1]. Actually it will be done through the expression of ϕ by hypergeometric functions

$$F(\alpha, \beta; \gamma; z) = 1 + \sum_{i=1}^\infty \frac{(\alpha)_i(\beta)_i}{(\gamma)_i}\frac{z^i}{i!} \tag{3.6}$$

where $(\alpha)_i = \alpha \cdots (\alpha+i-1)$. The following lemma summarizes the relationships we use. All formulas are from Abramowitz and Stegun [8] and the number following AS in each expression below is the formula number in Abramowitz and Stegun [8].

Lemma 3.1. • *AS.15.3.1*

$$F(\alpha, \beta; \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)}\int_0^1 t^{\beta-1}(1-t)^{\gamma-\beta-1}(1-tz)^{-\alpha}\mathrm{d}t \tag{3.7}$$

when $\gamma > \beta > 0$.

- *AS.15.3.4*

$$F(\alpha, \beta; \gamma; z) = (1-z)^{-\alpha} F(\alpha, \gamma - \beta; \gamma; z/(z-1)). \tag{3.8}$$

- *AS.15.2.25*

$$\gamma(1-z)F(\alpha, \beta; \gamma; z) - \gamma F(\alpha, \beta - 1; \gamma; z) + z(\gamma - \alpha)F(\alpha, \beta; \gamma + 1; z) = 0. \tag{3.9}$$

- *AS.15.2.18*

$$(\gamma - \alpha - \beta)F(\alpha, \beta; \gamma; z) - (\gamma - \alpha)F(\alpha - 1, \beta; \gamma; z) + \beta(1-z)F(\alpha, \beta + 1; \gamma; z) = 0. \tag{3.10}$$

- *AS.15.1.20*

$$\lim_{z \to 1} F(\alpha, \beta; \gamma; z) = \frac{\Gamma(\gamma)\Gamma(\gamma - \alpha - \beta)}{\Gamma(\gamma - \alpha)\Gamma(\gamma - \beta)} \tag{3.11}$$

  when $\gamma \neq 0, -1, -2, \ldots$ *and* $\gamma - \alpha - \beta > 0$.
- *AS.15.3.10*

$$\lim_{z \to 1} \{-\log(1-z)\}^{-1} F(\alpha, \beta; \alpha + \beta; z) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}. \tag{3.12}$$

- *AS.15.3.3*

$$F(\alpha, \beta; \gamma; z) = (1-z)^{\gamma - \alpha - \beta} F(\gamma - \alpha, \gamma - \beta; \gamma; z). \tag{3.13}$$

- *AS.15.2.1*

$$(\mathrm{d}/\mathrm{d}z)F(\alpha, \beta; \gamma; z) = \{\alpha\beta/\gamma\}F(\alpha + 1, \beta + 1; \gamma + 1; z). \tag{3.14}$$

We first use Lemma 3.1 to re-express $\phi(w)$.

Lemma 3.2. *Let* ϕ *be given by* (3.5). *Then provided* $a > -p/2 - 1$, $b > -1$ *and* $p/2 + n/2 + e + 2 > 0$, $\phi(w)$ *is expressed as*

$$\phi(w) = \frac{1 - G(v)}{\frac{n/2+e-a}{p/2+a+1} + G(v)} \tag{3.15}$$

where

$$G(v) = \frac{F(b, p/2 + n/2 + e + 1; p/2 + a + b + 2; v)}{F(b + 1, p/2 + n/2 + e + 1; p/2 + a + b + 2; v)}.$$

Proof. Using (3.7) and (3.8), we have

$$\int_0^1 t^{\beta-1}(1-t)^{\gamma-\beta-1}(1+tz)^{-\alpha}\mathrm{d}t = \frac{\Gamma(\beta)\Gamma(\gamma - \beta)}{\Gamma(\gamma)}(1+z)^{-\alpha}F(\alpha, \gamma - \beta; \gamma; z/(1+z)) \tag{3.16}$$

when $\gamma > \beta > 0$ and hence by (3.5)

$$\phi(w) = \frac{v}{1-v}\,\frac{p/2 + a + 1}{d}\,\frac{F(b + 1, c; d + 1; v)}{F(b + 1, c; d; v)}, \tag{3.17}$$

where $v = w/(w+1)$, $c = p/2 + n/2 + e + 2$, and $d = p/2 + a + b + 2$. Using (3.9) and (3.10), we have

$$\begin{aligned} \phi(w) &= -1 + \frac{F(b + 1, c - 1; d; v)}{(1-v)F(b + 1, c; d; v)} \\ &= -1 + \frac{c - 1}{n/2 + e - a + (p/2 + a + 1)G(v)} \\ &= \frac{1 - G(v)}{\frac{n/2+e-a}{p/2+a+1} + G(v)}. \quad \square \end{aligned} \tag{3.18}$$

The next lemma gives properties of ratios of hypergeometric functions such as G, which we employ in demonstrating minimaxity.

Lemma 3.3. *Let*

$$K(v) = \frac{F(b, \beta; \gamma; v)}{F(b + 1, \beta; \gamma; v)} \tag{3.19}$$

for $\beta > 0$ *and* $0 < b + 1 < \gamma < b + 1 + \beta$. *Then,*

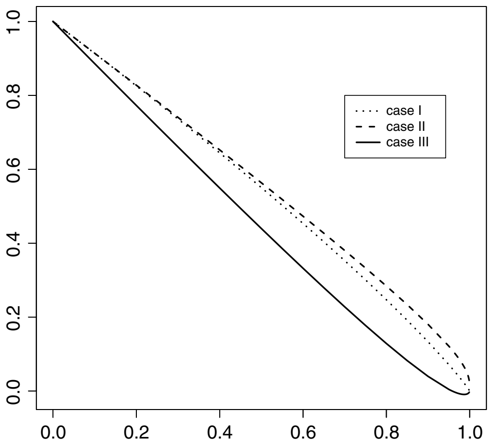

Fig. 1. The graph of $K(v)$.

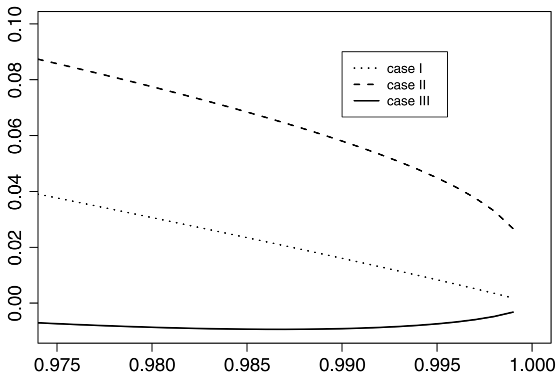

Fig. 2. The graph of $K(v)$ around the global minimum point.

(i) $K(0) = 1$
(ii) $\lim_{v\to 1} K(v) = 0$.
(iii) *If* $b \geq 0$, $K(v)$ *is monotone decreasing.*
(iv) *If* $-1 < b < 0$ *and* $\gamma \geq \beta$, $K(v)$ *is monotone decreasing.*
(v) *If* $-1 < b < 0$ *and* $\gamma < \beta$, *the minimum of* $K(v)$ *is a negative value and* $K(v)$ *approaches* 0 *from below as* v *approaches* 1. *Also* $\inf_{0\leq v\leq 1} K(v) \geq b/(b+1)$.

Note: In Fig. 1, we give a graph of behavior of the function $K(v)$ for the following cases:

(I) $b = 0.5, \beta = 1.5, \gamma = 1.75$,
(II) $b = -0.5, \beta = 1.5, \gamma = 1.75$,
(III) $b = -0.5, \beta = 2, \gamma = 1.75$,

where three cases (I), (II), and (III) correspond with (iii), (iv), and (v) of Lemma 3.3, respectively. Fig. 2 focuses attention on the behavior of $K(v)$ around the global minimum point in case (III). We clearly see that $K(v)$ for the case (III) is not a monotone function. Therefore small changes of β and γ around $\beta = \gamma$ affect the behavior of $K(v)$ dramatically.

Proof. First note that $K(0) = 1$ by (3.19). Next by (3.11) and (3.13),

$$\lim_{z\to 1}(1-z)^{\alpha+\beta-\gamma} F(\alpha, \beta; \gamma; z) = \frac{\Gamma(\gamma)\Gamma(\alpha+\beta-\gamma)}{\Gamma(\alpha)\Gamma(\beta)} \tag{3.20}$$

when $\gamma > 0$ and $\gamma - \alpha - \beta < 0$. Hence we have by (3.20), (3.12) and (3.11) respectively,

$$F(b, \beta; \gamma; v) \approx \begin{cases} (1-v)^{\gamma-b-\beta} \dfrac{\Gamma(\gamma)\Gamma(b+\beta-\gamma)}{\Gamma(b)\Gamma(\beta)} & \text{if } \gamma - b - \beta < 0 \\ \{-\log(1-v)\} \dfrac{\Gamma(\gamma)}{\Gamma(b)\gamma(\beta)} & \text{if } \gamma - b - \beta = 0 \\ \dfrac{\Gamma(\gamma)\Gamma(\gamma-b-\beta)}{\Gamma(\gamma-b)\Gamma(\gamma-\beta)} & \text{if } 0 < \gamma - b - \beta < 1 \end{cases} \tag{3.21}$$

and also by (3.20)

$$F(b+1, \beta; \gamma; v) \approx (1-v)^{\gamma-b-1-\beta} \frac{\Gamma(\gamma)\Gamma(b+1+\beta-\gamma)}{\Gamma(b+1)\Gamma(\beta)} \quad \text{if } \gamma - b - \beta < 1 \tag{3.22}$$

where $f(v) \approx g(v)$ means $\lim_{v\to 1} f(v)/g(v) = 1$. We easily see that the ratio $F(b, \beta; \gamma; v)/F(b+1, \beta; \gamma; v) = K(v)$ goes to zero as $v \to 1$ under the assumption $\gamma - b - \beta < 1$. Hence part (ii) follows.

When $b \geq 0$, $K(v)$ is decreasing from the monotone likelihood ratio property of the kernel of $k(v)$. Hence the first assertion of part (iii) follows.

When $-1 < b < 0$ and $\gamma - \beta \geq 0$, the numerator $F(b, \beta; \gamma; v)$ is always positive because it can be rewritten as $(1-v)^{\gamma-b-\beta}F(\gamma-b, \gamma-\beta; \gamma; v)$. Also the numerator $F(b, \beta; \gamma; v)$ with $-1 < b < 0$ is decreasing in v and the positive denominator $F(b+1, \beta; \gamma; v)$ is increasing in v. Hence the second assertion of part (iii) follows.

To show part (iv), note that $\Gamma(x) > 0$ if $x > 0$ and $\Gamma(y) < 0$ if $-1 < y < 0$. By assumption $-1 < b < 0$, $\gamma < \beta$, $\beta > 0$, $0 < b+1 < \gamma < b+1+\beta$. It then follows using the additional assumptions in the first and third lines of (3.21), that there is exactly 1 negative factor in each constant term and hence each constant term is negative. Since, also, the denominator of $K(v)$, $F(b+1, \beta; \gamma; v)$, is positive, it follows that $K(v)$ approaches 0 from below as v approaches 1. Let $K(v)$ take on its minimum value at v_0. Using the formula

$$(A/B)' = (B'/B)\{A'/B' - A/B\}$$

and (3.14), we have $K(v_0) = K_1(v_0)$ where

$$K_1(v) = \frac{b}{b+1} \frac{F(b+1, \beta+1; \gamma+1; v)}{F(b+2, \beta+1; \gamma+1; v)}.$$

Since $b/(b+1) < 0$ and $b+1 > 0$, $K_1(v)$ is increasing in v. Therefore

$$K(v) \geq K(v_0) = K_1(v_0) > K_1(0) = b/(b+1). \tag{3.23}$$

Hence part (iv) follows. □

The following corollary gives the behavior of $\phi(w)$ and follows immediately from Lemmas 3.2 and 3.3.

Corollary 3.1. *Assume that $e > -p/2 - n/2 - 1$, $-p/2 - 1 < a < n/2 + e$ and $b > -1$. Then*

(i) $\lim_{w\to\infty} \phi(w) = \alpha$ *where* $\alpha = (p/2 + a + 1)/(n/2 + e - a)$.
(ii) *When* $b \geq \min(n/2 + e - a - 1, 0)$, $\phi(w)$ *is monotone increasing.*
(iii) *When* $-1 < b < \min(n/2 + e - a - 1, 0)$, $\phi(w)$ *is not monotonic.*
(iv) *When*

$$-\frac{n/2 + e - a}{p/2 + n/2 + e + 1} < b < \min(n/2 + e - a - 1, 0), \tag{3.24}$$

$$0 \leq \phi(w) \leq \frac{1 - b/(b+1)}{1/\alpha + b/(b+1)} = \frac{(p/2 + a + 1)}{n/2 + e - a + b(p/2 + n/2 + e + 1)}.$$

Note: By analogy with the known variance case [9,10,3], it may be expected that the choice $-1 < b < 0$ would lead to a non-monotone ϕ. However part (ii) of the corollary shows that this need not be true in the unknown variance case, and in fact monotonicity depends on the relationship between n, a and e. The addition of the restriction (3.24) in part (iv) is necessitated by the fact that $\lim_{b\to -1} b/(b+1) = -\infty$. Hence a value of b close to -1 would cause the upper bound on ϕ, $\{1 - b/(b+1)\}/\{1/\alpha + b/(b+1)\}$ to be negative. Thanks to the restriction, the denominator is positive.

The next result gives a lower bound for $w\phi'(w)/\phi(w)$.

Lemma 3.4.

$$w\frac{\phi'(w)}{\phi(w)} \geq \frac{(p/2 + a + 2)b}{b+1} \tag{3.25}$$

provided that $e > -p/2 - n/2 - 1$, $-p/2 - 1 < a < n/2 + e - a$ and $-1 < b < \min(n/2 + e - a - 1, 0)$.

Proof. Note

$$\{wA(w)/B(w)\}'/\{A(w)/B(w)\} = 1 + wA'(w)/A(w) - wB'(w)/B(w).$$

Hence

$$\frac{w\phi'(w)}{\phi(w)} = 1 + cw\left\{-\frac{\int_0^1 \lambda^{p/2+a+2}(1-\lambda)^b(1+w\lambda)^{-c-1}\mathrm{d}\lambda}{\int_0^1 \lambda^{p/2+a+1}(1-\lambda)^b(1+w\lambda)^{-c}\mathrm{d}\lambda} + \frac{\int_0^1 \lambda^{p/2+a+1}(1-\lambda)^b(1+w\lambda)^{-c-1}\mathrm{d}\lambda}{\int_0^1 \lambda^{p/2+a}(1-\lambda)^b(1+w\lambda)^{-c}\mathrm{d}\lambda}\right\}$$

where $c = p/2 + n/2 + e + 2$. Using (3.7) and (3.8), we have

$$\frac{w\phi'(w)}{\phi(w)} = -cv\left\{\frac{\frac{p}{2}+a+2}{d+1}\frac{F(b+1,c+1;d+2;v)}{F(b+1,c;d+1;v)} - \frac{\frac{p}{2}+a+1}{d}\frac{F(b+1,c+1;d+1;v)}{F(b+1,c;d;v)}\right\} + 1,$$

where $v = w/(w+1)$ and $d = p/2 + a + b + 2$. Using (3.9) and (3.10),

$$\begin{aligned}\frac{w\phi'(w)}{\phi(w)} &= c(1-v)\left\{\frac{F(b+1,c+1;d+1;v)}{F(b+1,c;d+1;v)} - \frac{F(b+1,c+1;d;v)}{F(b+1,c;d;v)}\right\} + 1\\ &= \left(\frac{p}{2}+a+2\right)\frac{F(b,c;d+1;v)}{F(b+1,c;d+1;v)} - \left(\frac{p}{2}+a+1\right)\frac{F(b,c;d;v)}{F(b+1,c;d;v)}.\end{aligned}$$

Note $F(b,c;d+1;v) > F(b,c;d;v)$ since $-1 < b < 0$. For v which satisfies $F(b,c;d;v) > 0$, we have

$$\frac{w\phi'(w)}{\phi(w)} = F(b,c;d;v)\left(\frac{p/2+a+2}{F(b+1,c;d+1;v)} - \frac{p/2+a+1}{F(b+1,c;d;v)}\right) \ge 0,$$

because $F(b+1,c;d;v) > F(b+1,c;d+1;v)$. For v which satisfies $F(b,c;d+1;v) > 0 > F(b,c;d;v)$, $w\phi'(w)/\phi(w)$ is clearly nonnegative. For v which satisfies $F(b,c;d+1;v) < 0$ and $-1 < b < \min(n/2+e-a-1, 0)$, we have

$$\frac{w\phi'(w)}{\phi(w)} \ge \left(\frac{p}{2}+a+2\right)\frac{F(b,c;d+1;v)}{F(b+1,c;d+1;v)} \ge \left(\frac{p}{2}+a+2\right)\frac{b}{b+1}$$

by Lemma 3.3(iv). □

The main result is the following.

Theorem 3.2. *Suppose $e > -p/2 - n/2 - 1$.*

(i) *[monotone ϕ] When*

$$-p/2 - 1 < a \le \frac{c(p,n)(n/2+e) - p/2 - 1}{1 + c(p,n)} \tag{3.26}$$

where $c(p,n) = 2(p-2)/(n+2)$ and $b \ge \min(n/2+e-a-1, 0)$, the generalized Bayes estimator is minimax.

(ii) *[non-monotone ϕ] When*

$$-p/2 - 1 < a < \frac{c(p,n)(n/2+e) - p/2 - 1}{1 + c(p,n)} \tag{3.27}$$

$$-1 < -(n+2)\frac{c(p,n) - \alpha}{4(p+a+1)(\alpha+1)} \le b < \min(0, n/2+e-a-1) \tag{3.28}$$

where $\alpha = (p/2+a+1)/(n/2+e-a) = \lim_{w\to\infty}\phi(w)$, the generalized Bayes estimator is minimax.

Proof. First we prove part (i). Monotonicity of ϕ follows from Corollary 3.1(ii) since $b \ge \min(0, n/2+e-a-1)$. Also using Corollary 3.1(i) and (3.26)

$$0 \le \phi \le \alpha = \frac{p/2+a+1}{n/2+e-a} \le 2\frac{p-2}{n+2} = c(p,n).$$

Hence minimaxity follows from Corollary 2.1 and part (i) follows.

Next we consider part (ii), the non-monotonic ϕ case. The lower bound in (3.28)

$$-(n+1)\frac{c(p,n)-\alpha}{4(p+a+1)(\alpha+1)}$$

is negative because of (3.27). Also Corollary 3.1(i) and (3.27) implies

$$0 < \lim_{w\to\infty}\phi(w) = \alpha = \frac{p/2+a+1}{n/2+e-a} < 2\frac{p-2}{n+2} = c(p,n).$$

Also since the lower bound in (3.28) satisfies

$$\begin{aligned}-(n+2)\frac{c(p,n)-\alpha}{4(p+a+1)(\alpha+1)} - \left(-\frac{n/2+e-a}{p/2+n/2+e+1}\right) &\ge \frac{1}{\alpha+1}\left(1 - \frac{(n+2)c(p,n)}{4(p+a+1)}\right)\\ &= \frac{1}{\alpha+1}\left(1 - \frac{p-2}{p+2(p/2+a+1)}\right)\\ &\ge \frac{1}{\alpha+1}\frac{2}{p} > 0,\end{aligned}$$

it follows from Corollary 3.1(iv) that

$$\phi(w) \le \frac{\alpha}{b+1+\alpha b} = M_1.$$

Additionally, by Lemma 3.4(iv),

$$w\frac{\phi'(w)}{\phi(w)} \ge \frac{(p/2+a+2)b}{b+1} = -M_2.$$

By Corollary 2.2(ii), it follows that the generalized Bayes estimator is minimax provided

$$\frac{(n+2)\{M_1 - c(p,n)\}}{1+M_1} + 4M_2 \le 0 \tag{3.29}$$

but a straightforward calculation shows that this condition is equivalent to

$$b \ge -(n+2)\frac{c(p,n)-\alpha}{4(p+a+1)(\alpha+1)}. \tag{3.30}$$

Hence the generalized Bayes estimator is minimax since (3.28) guarantees that (3.29) is satisfied. This completes the proof. □

Note: A recent paper by Wells and Zhou [11] also derives generalized Bayes estimators in the normal case, some of which have non-monotone shrinkage functions ϕ (in our notation). The generalized Bayes minimax estimators of Theorem 3.2(ii) all have non-monotone shrinkage functions and their minimaxity cannot be shown by the methods in Wells and Zhou [11] which rely on integration by parts. As noted earlier, the assumption that $b < 0$ causes the usual integration by parts technique to fail.

The behavior of shrinkage functions of some generalized Bayes estimators and their risk performance will be given in Section 4.

4. Generalized Bayes estimators for spherically symmetric distributions

In this section, we consider generalized Bayes minimax estimators for spherically symmetric distributions. As shown in [10,1], the special choice $b = -a-2$ in the prior given by (3.1) leads to the separated joint density of θ and η, $\|\theta\|^{-2(p/2+a+1)}\eta^{-a-1+e}$. This follows since

$$\int_0^1 \exp\left(-\frac{\eta\lambda\|\theta\|^2}{2(1-\lambda)}\right)^{p/2}\left(\frac{\eta\lambda}{1-\lambda}\right)^{p/2}\lambda^a(1-\lambda)^b \mathrm{d}\lambda = \eta^{-a-1}\int_0^\infty t^{p/2+a}\exp\left(-\frac{t}{2}\|\theta\|^2\right)\mathrm{d}t$$
$$\propto \|\theta\|^{-2(p/2+a+1)}\eta^{-a-1}, \tag{4.1}$$

if $p/2+a+1>0$. Because of this simplification, we make the assumption $b=-a-2$ throughout the rest of this section. Under quadratic loss $\eta(d-\theta)'(d-\theta)$, even in the spherically symmetric situation, the generalized Bayes estimator is given by $E(\eta\theta|X,S)/E(\eta|X,S)$ and hence we have the generalized Bayes estimator with respect to our prior,

$$\frac{\int_{R^p}\int_0^\infty \theta\eta^{(n+p)/2-a+e}f(\eta\{\|X-\theta\|^2+S\})\|\theta\|^{-2(p/2+a+1)}\mathrm{d}\eta\mathrm{d}\theta}{\int_{R^p}\int_0^\infty \eta^{(n+p)/2-a+e}f(\eta\{\|X-\theta\|^2+S\})\|\theta\|^{-2(p/2+a+1)}\mathrm{d}\eta\mathrm{d}\theta}$$
$$= \frac{\int_{R^p}\theta(\{\|X-\theta\|^2+S\})^{-(n+p)/2+a-e-1}\|\theta\|^{-2(p/2+a+1)}\mathrm{d}\theta\int_0^\infty \eta^{(n+p)/2-a+e}f(\eta)\mathrm{d}\eta}{\int_{R^p}(\{\|X-\theta\|^2+S\})^{-(n+p)/2+a-e-1}\|\theta\|^{-2(p/2+a+1)}\mathrm{d}\theta\int_0^\infty \eta^{(n+p)/2-a+e}f(\eta)\mathrm{d}\eta}$$
$$= \frac{\int_{R^p}\theta(\{\|X-\theta\|^2+S\})^{-(n+p)/2+a-e-1}\|\theta\|^{-2(p/2+a+1)}\mathrm{d}\theta}{\int_{R^p}(\{\|X-\theta\|^2+S\})^{-(n+p)/2+a-e-1}\|\theta\|^{-2(p/2+a+1)}\mathrm{d}\theta} \tag{4.2}$$

if

$$\int_0^\infty \eta^{(n+p)/2-a+e}f(\eta)\mathrm{d}\eta < \infty. \tag{4.3}$$

Note that the estimator on the right-hand side of (4.2) does not depend on f and hence is equal to the generalized Bayes estimator in the normal case. In the normal case, as seen in Section 3, the estimator is well defined if $a > -p/2-1$, $b > -1$ and $e > -p/2-n/2-2$. Since $a=-b-2$, the inequality $-p/2-1 < a < -1$ is also satisfied. Hereafter the estimator on the right-hand side of (4.2) is denoted by $\delta^*_{a,e}$. In summary, we have the following result on $\delta^*_{a,e}$.

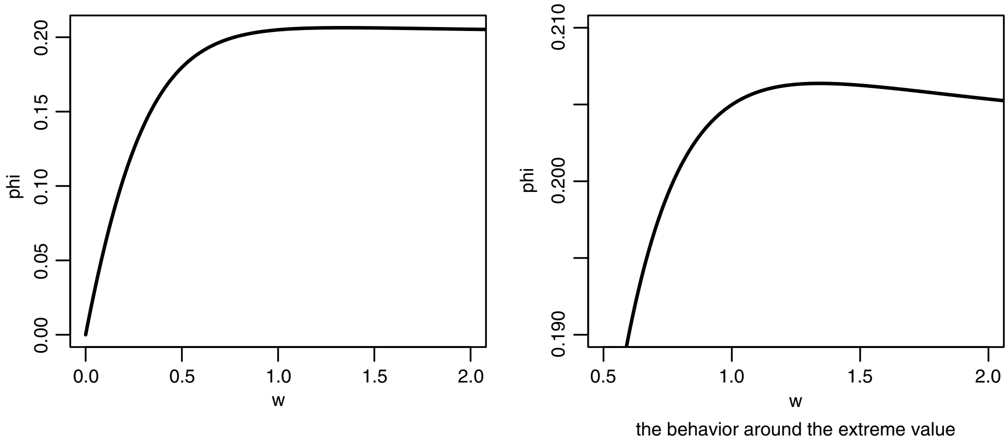

Fig. 3. The graph of $\phi(w)$ of $\delta^*_{a,e}$ when $p = 5$, $n = 10$, $a = a_*$, and $e = n/4 - 1$.

Theorem 4.1. *Suppose $-p/2 - 1 < a < -1$ and $e > -p/2 - n/2 - 2$. Then the estimator*

$$\delta^*_{a,e} = \frac{\int_{R^p} \theta(\{\|X-\theta\|^2 + S\})^{-(n+p)/2+a-e-1}\|\theta\|^{-2(p/2+a+1)}\mathrm{d}\theta}{\int_{R^p} (\{\|X-\theta\|^2 + S\})^{-(n+p)/2+a-e-1}\|\theta\|^{-2(p/2+a+1)}\mathrm{d}\theta}$$

is generalized Bayes with respect to $\|\theta\|^{-2(p/2+a+1)}\eta^{-a-1+e}$ for the class of spherically symmetric distributions which satisfy the moment condition (4.3).

We note that Theorem 3.2, together with the general results of Section 2, imply that the normal theory generalized Bayes estimators in Theorem 3.2 remain minimax (but not necessarily generalized Bayes) for the entire class of spherically symmetric distributions.

Adding the restriction $b = -a - 2$ in Theorem 3.2, we have the following generalized Bayes minimaxity result on the estimator $\delta^*_{a,e}$ for the entire class of spherically symmetric distributions.

Theorem 4.2. (i) *Suppose $-p/2 - n/2 - 1 < e \le -n/4 - 3/2$. When*

$$-p/2 - 1 < a \le \frac{c(p,n)(n/2+e) - p/2 - 1}{1 + c(p,n)} \tag{4.4}$$

where $c(p,n) = 2(p-2)/(n+2)$, the estimator $\delta^*_{a,e}$ is minimax for the entire class of spherically symmetric distributions.

(ii) *Suppose $e > n/4 - 3/2$ and $n \ge 2$. When*

$$-p/2 - 1 < a \le a_*(p,n,e) \tag{4.5}$$

where $a_*(p,n,e)$ is the larger solution of the equation $g(a) = 0$ where

$$\begin{aligned} g(a) = {} & (2p + 2n + 4e + 4)a^2 + a\{2p^2 + 2pn + 12p + 7n + 4(p+3)e + 10\} \\ & \times 4p^2 + \{7/2\}pn + 6(p+2)e + 7n + 13p + 10, \end{aligned}$$

the estimator $\delta^*_{a,e}$ is minimax for the entire class of spherically symmetric distributions.

(iii) *The minimax estimator given by* (i) *and* (ii) *is also generalized Bayes if* (4.3) *is satisfied.*

Note: Since $a_*(p,n,e)$ will be shown to be between $(-2,-1)$ in the proof, minimaxity of the estimator $\delta^*_{a,e}$ with $-2 < a \le a_*(p,n,e)$ $(-1 < -a_*(p,n,e) - 2 < b < 0)$, which has non-monotone shrinkage factor ϕ, is new compared to Maruyama and Strawderman [1]. See Fig. 3.

Proof. First consider that the upper bound of a for minimaxity in Theorem 3.2,

$$u_*(p,n,e) = \frac{c(p,n)(n/2+e) - p/2 - 1}{1 + c(p,n)}.$$

A simple calculation gives

$$u_*(p,n,e) + 2 = \frac{(p-2)(n+4e+6)}{2(n+2)\{1 + c(p,n)\}}.$$

Hence $u_*(p,n,e)$ is greater than -2 if and only if $e > -n/4 - 3/2$. To show (i), note that when $-p/2 - n/2 - 1 < e \le -n/4 - 3/2$, $u_*(p,n,e) < -2$ and (4.4) satisfies the sufficient condition on a for minimaxity of Theorem 3.2(i). Also since $-a - 2$ $(=b)$ is nonnegative, the condition on b is also satisfied. Hence part (i) follows.

Table 1
The relative risk.

df \ NC	0	1	2	3	4	5	10	$E[\|X-\theta\|^2]$
	$p=5, n=10$							
5	0.51	0.66	0.82	0.90	0.93	0.95	0.98	8.44
10	0.51	0.71	0.89	0.95	0.97	0.98	0.99	6.22
20	0.51	0.74	0.91	0.96	0.98	0.99	0.99	5.52
∞	0.51	0.75	0.93	0.97	0.98	0.99	0.99	5.00
	$p=10, n=10$							
5	0.38	0.58	0.77	0.86	0.91	0.94	0.98	16.5
10	0.38	0.63	0.84	0.92	0.95	0.97	0.99	12.6
20	0.38	0.66	0.87	0.94	0.96	0.98	0.99	11.1
∞	0.38	0.69	0.89	0.95	0.97	0.98	0.99	10.0

Next we show (ii). When $e > -n/4-3/2$, the estimator with $-p/2-1 < a \le -2$ is minimax because $-2 < u_*(p,n,e)$ and $-a-2\ (=b) \ge 0$ guarantees minimaxity by Theorem 3.2(i).

Finally we consider the case where $-2 < a < -1$ and $e > -n/4-3/2$. The condition $b = -a-2 < \min(n/2+e-a-1, 0)$ in Theorem 3.2(ii) is clearly satisfied under the assumption $n \ge 2$. Also a straightforward calculation shows that the inequality

$$-(n+2)\frac{c(p,n)-\alpha}{4(p+a+1)(\alpha+1)} \le -a-2 = b \tag{4.6}$$

is equivalent to $g(a) \le 0$. Note, (4.6), for $a=-1$, is not satisfied because

$$\begin{aligned}-a-2+(n+2)\frac{c(p,n)-\alpha}{4(p+a+1)(\alpha+1)} &= -1+(n+2)\frac{c(p,n)-\alpha}{4p(\alpha+1)}\\ &\le -1+\frac{2(p-2)}{4p(\alpha+1)}\\ &< 0.\end{aligned}$$

Also (4.6), for $a=-2$, is satisfied because

$$-a-2+(n+2)\frac{c(p,n)-\alpha}{4(p+a+1)(\alpha+1)} = (n+2)\frac{c(p,n)-\alpha}{4(p-1)(\alpha+1)} > 0.$$

Hence $g(-2) < 0$ and $g(-1) > 0$, which guarantees that the larger solution $a_*(p,n,e)$ of the quadratic equation $g(a)=0$ should be between -2 and -1. Additionally $g(u_*(p,n,e)) > 0$ because

$$-u_*(p,n,e)-2+(n+2)\frac{c(p,n)-\alpha}{4(p+u_*(p,n,e)+1)(\alpha+1)} = -u_*(p,n,e)-2 < 0,$$

which means that $a_*(p,n,e)$ is smaller than $u_*(p,n,e)$. Hence $-2 < a \le a_*(p,n,e)$ leads to minimaxity. This completes the proof of (ii).

Part (iii) follows from Part (i), (ii) and Theorem 4.1. □

4.1. Numerical results

In this subsection, we give some numerical results. We consider two cases $p=5, n=10$ and $p=10, n=10$. We also consider

$$f(t) \propto (1+t/\nu)^{-(p+n+\nu)/2}$$

for $\nu = 5, 10, 20, \infty$. These choices of ν correspond to multivariate-t with degrees of freedom ($\nu = 5, 10, 20$) and multivariate normal ($\nu=\infty$), respectively. As parameters of the estimator $\delta^*_{a,e}$, $e = n/4-1$ and $a = a_*(p,n,e)$ are taken. As pointed out in the note just below Theorem 4.2, the corresponding shrinkage function is not monotone (see Fig. 3). For $p = 5,\ 10$, $a_*(p,n,e)$ is about -1.82 in both cases. The moment condition (4.3) for being generalized Bayes is clearly satisfied when $-a+e-\nu/2 < -1$. Hence the estimator $\delta^*_{a,e}$ is justified as generalized Bayes only if $\nu \ge 9$. From Theorem 4.2, the estimator $\delta^*_{a,e}$ should be minimax for the entire class of spherically symmetric distributions, which include multivariate-t with $\nu \le 8$. Table 1 reports the minimax risk $E[\|X-\theta\|^2]$ and the relative risk which is defined by $E[\|\delta-\theta\|^2]/E[\|X-\theta\|^2]$. In the table, NC means (scaled) non-centrality $\|\theta\|/p^{1/2}$.

References

[1] Y. Maruyama, W.E. Strawderman, A new class of generalized Bayes minimax ridge regression estimators, Ann. Statist. 33 (2005) 1753–1770.
[2] Y. Maruyama, A unified and broadened class of admissible minimax estimators of a multivariate normal mean, J. Multivariate Anal. 64 (1998) 196–205.
[3] Y. Maruyama, Stein's idea and minimax admissible estimation of a multivariate normal mean, J. Multivariate Anal. 88 (2004) 320–334.
[4] T. Kubokawa, M.S. Srivastava, Robust improvement in estimation of a mean matrix in an elliptically contoured distribution, J. Multivariate Anal. 76 (2001) 138–152.
[5] C. Robert, The Bayesian Choice, Springer, New York, 1994.
[6] C. Stein, Estimation of the mean of a multivariate normal distribution, in: Proc. Prague Symp. Asymptotic Statist., 1973, pp. 345–381.
[7] B. Efron, C. Morris, Families of minimax estimators of the mean of a multivariate normal distribution, Ann. Statist. 4 (1976) 11–21.
[8] M Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, Dover Publications, New York, 1964.
[9] K. Alam, A family of admissible minimax estimators of the mean of a multivariate normal distribution, Ann. Statist. 1 (1973) 517–525.
[10] Y. Maruyama, A robust generalized Bayes estimator improving on the James–Stein estimator for spherically symmetric distributions, Statist. Decisions 21 (2003) 69–77.
[11] M.T. Wells, G. Zhou, Generalized Bayes minimax estimators of the mean of multivariate normal distribution with unknown variance, J. Multivariate Anal. 99 (10) (2008) 2208–2220.