A robust generalized Bayes estimator improving on the James-Stein estimator for spherically symmetric distributions

Yuzo Maruyama

Summary: The problem of estimating a mean vector for spherically symmetric distributions with the quadratic loss function is considered. A robust generalized Bayes estimator improving on the James-Stein estimator is given.

1 Introduction

Consider the linear regression model

\[ Y = A\beta + e \]  

(1.1)

where \( y \) is an \( N \times 1 \) response vector, \( A \) is an \( N \times p \) matrix of rank \( p \leq N \) of known constants, \( \beta \) is a \( p \times 1 \) vector of unknown parameters, and \( e \) is an \( N \times 1 \) vector of unobservable random errors. We assume that the error \( e \) has a spherically symmetric density \( \sigma^{-N} f(\|e\|^2/\sigma^2) \), where \( \sigma^2 \) is an unknown parameter and \( f(\cdot) \) is a nonnegative function on the nonnegative real line. We can easily derive the canonical form of (1.1). Let \( P \) be an \( N \times N \) orthogonal matrix such that

\[ PA = \begin{pmatrix} (A'A)^{1/2} \\ 0 \end{pmatrix} \]

and let \( \theta = (A'A)^{1/2}\beta \). Hence two random vectors \( X = (X_1, \ldots, X_p)' \) and \( Z = (Z_1, \ldots, Z_n)' \) where

\[ (X, Z) = PY \quad \text{and} \quad n = N - p \]

have the joint density of the form of

\[ \sigma^{-p-n} f((\|x - \theta\|^2 + \|z\|^2)/\sigma^2), \]  

(1.2)
where \( \theta \) is a \( p \times 1 \) vector of unknown parameters. References on these distributions which generalize the multivariate normal distribution in linear regression model are given by Kelker [8], Eaton [5], Fang and Anderson [6] and Kubokawa and Srivastava [11]. Then we consider the problem of estimating the mean vector \( \theta \) by \( \delta(X, Z) \) relative to the quadratic loss function \( L(\theta, \sigma^2, d) = ||d - \theta||^2/\sigma^2 \).

For the normal model, it is well-known that the usual minimax estimator \( X \) is inadmissible for \( p \geq 3 \) as shown in Stein [14]. James and Stein [7] succeeded in giving an explicit form of an estimator dominating \( X \) as

\[
\delta^JS = \left(1 - \frac{p-2}{n+2} \frac{||Z||^2}{||X||^2} \right) X,
\]

which is called the James-Stein estimator. For the spherically symmetric model, Cellier et al. [4] and Kubokawa and Srivastava [11] showed that the James-Stein estimator dominates the usual estimator \( X \) independent of \( f \) in (1.2), which does not need to be known. However it turns out that the James-Stein estimator is inadmissible since its positive-part estimator is superior to it as shown in Baranchik [2]. Moreover its positive-part estimator is not analytic and thus inadmissible. Therefore it has been of interest to derive analytic (generalized Bayes, if possible) estimators improving on \( X \) and the James-Stein estimator.

For the normal model, Lin and Tsai [12] derived a class of minimax generalized Bayes estimators. Kubokawa [9] showed that the estimator \( \delta_K = (1 - \phi_K(W)/W)X \), where \( W = ||X||^2/||Z||^2 \) and

\[
\phi_K(w) = w \int_0^1 \frac{\lambda^{p/2-1}(1 + \lambda w)^{-p/2-1} d\lambda}{\int_0^1 \lambda^{p/2-2}(1 + \lambda w)^{-p/2-1} d\lambda},
\]

which is included in Lin and Tsai’s class, dominates the James-Stein estimator. Moreover Kubokawa [10] derived a sufficient condition for domination over the James-Stein estimator and showed that \( \delta_K \) satisfies it. Maruyama [13] showed that some generalized Bayes estimators besides \( \delta_K \) satisfy Kubokawa’s [10] condition. However such estimators have not been derived for the spherically symmetric model yet.

In this paper, we show that Lin and Tsai [12] and Kubokawa’s [9, 10] results remain robust under a broad subclass of spherically symmetric distributions although these seem to depend upon the normality. In particular we recommend the use of \( \delta_K \) for any spherical symmetric distribution since it is minimax for any such \( f \), dominates the James-Stein estimator for those \( f \) which are unimodal, and is also generalized Bayes under the condition of a finite fourth moment.

2 Generalized Bayes estimators for spherically symmetric distributions

Letting \( \eta = \sigma^{-2} \), we consider the prior distribution whose joint density of \( \theta \) and \( \eta \) is proportional to \( \eta^a||\theta||^{-b} \). From the fact

\[
||\theta||^{-b} = \frac{\eta^{b/2}}{\Gamma(b/2)2^{b/2}} \int_0^1 \lambda^{b/2-1}(1 - \lambda)^{-b/2-1} \exp \left( -\frac{\eta \lambda}{2(1 - \lambda)||\theta||^2} \right) d\lambda,
\]

(2.1)
for $b > 0$, this prior is interpreted as the hierarchical prior

$$\theta | \lambda, \eta \sim N_p \left(0, \eta^{-1} \frac{1-\lambda}{\lambda} I_p\right), \quad \lambda \propto \lambda^{b/2-p/2-1}(1-\lambda)^{-b/2+p/2-1}, \quad \eta \propto \eta^{b/2-p/2+a},$$

which is a special case of ones considered in Lin and Tsai [12] and Alam [1] in the normal case of our problem. In the estimation of a multivariate normal mean with known variance, Baranchik [2] investigated the generalized Bayes estimator with respect to $\eta\|d-\theta\|^2$.

Under the quadratic loss function $\eta\|d-\theta\|^2$, the generalized Bayes estimator is given by $E(\eta|\theta,X,Z)/E(\eta|X,Z)$ and we have the generalized Bayes estimator with respect to our prior,

$$\begin{align*}
\int_{R^p} \int_0^{\infty} \theta \eta^{(n+p)/2+a+1} f(\eta(\|X-\theta\|^2 + \|Z\|^2)) \|\theta\|^{-b} d\theta d\eta \\
\int_{R^p} \int_0^{\infty} \eta^{(n+p)/2+a+1} f(\eta(\|X-\theta\|^2 + \|Z\|^2)) \|\theta\|^{-b} d\theta d\eta
\end{align*}$$

if $\int \eta^{(n+p)/2+a+1} f(\eta) d\eta < \infty$, which is equivalent to the finiteness of the $2(a+2)$-th moment of the distribution of $X$ and $Z$. That is to say, the generalized Bayes estimator under the spherically symmetric case does not depend on $f$ and hence coincides with one under the normal case. From (2.1) we have

$$\begin{align*}
\int_{R^p} \theta \exp \left( -\frac{\tau}{2} \|x-\theta\|^2 - \frac{\tau\lambda\|\theta\|^2}{2(1-\lambda)} \right) d\theta &= \left(\frac{2\pi(1-\lambda)}{\tau}\right)^{p/2} \exp(-\lambda\tau\|x\|^2/2)(1-\lambda)x \\
\int_{R^p} \exp \left( -\frac{\tau}{2} \|x-\theta\|^2 - \frac{\tau\lambda\|\theta\|^2}{2(1-\lambda)} \right) d\theta &= \left(\frac{2\pi(1-\lambda)}{\tau}\right)^{p/2} \exp(-\lambda\tau\|x\|^2/2).
\end{align*}$$

Moreover from the relation

$$\int_0^{\infty} \tau^{n/2+a+1+b/2} \exp \left( -\frac{\tau\lambda\|x\|^2 + \|z\|^2}{2} \right) = \frac{\Gamma(n/2 + a + b/2 + 2) 2^{n/2+a+b/2} 2^{n/2+a+b/2+2}}{\Gamma(1/2 + w\lambda)}$$

where $w = \|x\|^2/\|z\|^2$, we have the generalized Bayes estimator $\delta_{a,b}(X,Z) = (1 - \phi_{a,b}(W)/W)\mathcal{X}$ where

$$\phi_{a,b}(w) = w \int_0^1 \lambda^{b/2}(1-\lambda)^{p/2-b/2-1}(1 + w\lambda)^{-n/2-a-b/2-2} d\lambda$$

which is well-defined if $0 < b < p$ and $n/2 + a + b/2 + 2 > 0$. Note that $\delta_K = \delta_{a,b}$ for $a = 0$ and $b = p - 2$.

Here we summarize the result on the property of $\delta_{a,b}$. 
Theorem 2.1 For $0 < b < p$, $n/2 + a + b/2 + 2 > 0$ and any spherically symmetric distribution, the $2(a+2)$-th moment of which is finite, $\delta_{a,b}$ is generalized Bayes with respect to the density $n^a \| \theta \|^{-b}$.

The properties of the behavior of $\phi_{a,b}(w)$ is as follows.

Theorem 2.2

1. $\phi_{a,b}(w)$ is monotone increasing in $w$ if $0 < p \leq p - 2$.

2. $\lim_{w \to \infty} \phi_{a,b}(w) = b/(n + 2a + 2)$ if $a > -n/2 - 1$ and

$$|\phi_{a,b}(w) - b/(n + 2a + 2)| = \begin{cases} O\{(w + 1)^{-n/2-a-1}\} & \text{for } b = p - 2 \\ O\{(w + 1)^{-1}\} & \text{for } 0 < b < p - 2. \end{cases}$$

Proof: By the change of variables, we have

$$\phi_{a,b}(w) = \frac{\int_0^w t^{b/2}(1 - t/w)^{p/2-b/2-1}(1 + t)^{-n/2-a-b/2-2} dt}{\int_0^w t^{b/2-1}(1 - t/w)^{p/2-b/2-1}(1 + t)^{-n/2-a-b/2-2} dt}.$$ 

For $w_1 > w_2$ and $0 < b \leq p - 2$,

$$\frac{\int_0^{w_1} t^{b/2}(w_1 - t)^{p/2-b/2-1}(1 + t)^{-n/2-a-b/2-2} dt}{\int_0^{w_1} t^{b/2-1}(w_1 - t)^{p/2-b/2-1}(1 + t)^{-n/2-a-b/2-2} dt} \geq \frac{\int_0^{w_2} t^{b/2}(w_1 - t)^{p/2-b/2-1}(1 + t)^{-n/2-a-b/2-2} dt}{\int_0^{w_2} t^{b/2-1}(w_1 - t)^{p/2-b/2-1}(1 + t)^{-n/2-a-b/2-2} dt} \geq \frac{\int_0^{w_2} t^{b/2}(w_2 - t)^{p/2-b/2-1}(1 + t)^{-n/2-a-b/2-2} dt}{\int_0^{w_2} t^{b/2-1}(w_2 - t)^{p/2-b/2-1}(1 + t)^{-n/2-a-b/2-2} dt}.$$ 

The first inequality is from the fact that the ratio of integrands of the numerator and the denominator is increasing, the second inequality from the fact that $\{(w_2 - t)/(w_1 - t)\}^{p/2-b/2-1}$ is increasing. This completes the proof of part 1.

From an identity

$$\int_0^1 \lambda^a (1 - \lambda)^\beta (1 + w\lambda)^{-\gamma} d\lambda = (w + 1)^{-\alpha-1} \int_0^1 t^a (1 - t)^\beta (1 - tw/(w + 1))^{-\alpha-\beta+\gamma-2} dt,$$

we have

$$\phi_{a,b}(w) = v \frac{\int_0^1 t^{b/2}(1 - t)^{p/2-b/2-1}(1 - vt)^{-p/2+n/2+a+b/2+1} dt}{\int_0^1 t^{b/2-1}(1 - t)^{p/2-b/2-1}(1 - vt)^{-p/2+n/2+a+b/2+2} dt}$$

for $v = w/(w + 1)$, which implies that

$$\lim_{w \to \infty} \phi_{a,b}(w) = \frac{\int_0^1 t^{b/2}(1 - t)^{n/2+a} dt}{\int_0^1 t^{b/2-1}(1 - t)^{n/2+a+1} dt} = \frac{b}{n + 2a + 2}.$$
for $0 < b < p$ and $a > -n/2 - 1$. When $0 < b \leq p - 2$, applying an integration by parts gives

$$
\phi_{a,b}(w) = - \frac{(n/2 + a + 1)^{-1}(1 - v)^{n/2+a+1}}{\int_0^1 t^{p/2-2}(1 - vt)^{n/2+a+1}dt} + \frac{b}{n + 2a + 2} \tag{2.3}
$$

for $b = p - 2$ and

$$
\phi_{a,b}(w) = \frac{b}{n + 2a + 2} - \frac{p - b - 2}{n + 2a + 2}(1 - v)
\times \frac{\int_0^1 t^{b/2}(1 - t)^{p/2-b/2-2}(1 - vt)^{-p/2+n/2+a+b/2+1}dt}{\int_0^1 t^{b/2-1}(1 - t)^{p/2-b/2-1}(1 - vt)^{-p/2+n/2+a+b/2+2}dt}, \tag{2.4}
$$

for $0 < b < p - 2$. From (2.3) and (2.4), we can easily see that $|b/(n + 2a + 2) - \phi_{a,b}| = O\{(w + 1)^{-n/2-a-1}\}$ for $b = p - 2$ and $= O\{(w + 1)^{-1}\}$ for $0 < b < p - 2$, which completes the proof of part 2.

3 Improving on $X$ and the James-Stein estimator


$$
E[(X_i - \theta_i)h(X)] = \sigma^2 \int \frac{\partial}{\partial x_i} h(x) F \left( \frac{\|x - \theta\|^2 + \|z\|^2}{\sigma^2} \right) dx dz,
$$

$$
E[\|z\|^2 g(\|z\|^2)] = \sigma^2 \int \left( ng(\|z\|^2) + 2\|z\|^2 g'(\|z\|^2) \right) F \left( \frac{\|x - \theta\|^2 + \|z\|^2}{\sigma^2} \right) dx dz,
$$

for suitable $f$ and $g$. By using these identities, the risk of an estimator of the form

$$
\delta_\phi(X, Z) = (1 - \phi(\|X\|^2/\|Z\|^2)\|Z\|^2/\|X\|^2)X
$$

is expressed as

$$
R(\theta, \sigma^2, \delta_\phi) = E \left[ \frac{\|X - \theta\|^2}{\sigma^2} \right] + \sigma^{-2} E \left[ \phi^2 \left( \frac{\|X\|^2}{\|Z\|^2} \right) \frac{\|Z\|^4}{\|X\|^4} \right] + 2E \left[ \frac{(X - \theta)'X}{\sigma^2} \phi \left( \frac{\|X\|^2}{\|Z\|^2} \right) \frac{\|Z\|^2}{\|X\|^2} \right]
$$

$$
= p + \int_{\Re^p} \int_{\Re^p} \left( \frac{\phi(w)}{w} \right) \{ (n + 2)\phi(w) - 2(p - 2) \}
- 4\phi'(w)(1 + \phi(w)) \sigma^{-p-n} F((\|x - \theta\|^2 + \|z\|^2)/\sigma^2) dx dz, \tag{3.1}
$$

where $w = \|x\|^2/\|z\|^2$ and $F(u) = 2^{-1} \int_0^\infty f(t)dt$. Hence a sufficient condition for dominance over the natural estimator $X$ is derived as follows.

**Theorem 3.1 (Kubokawa and Srivastava)** Assume that $\phi(w)$ is monotone nondecreasing and $0 \leq \phi(w) \leq 2(p - 2)/(n + 2)$ for every $w \geq 0$. Then $\delta_\phi$ dominates $X$. 

This sufficient condition under the normal case was derived by Baranchik [3].

Among a simple class of shrinkage estimators \( \delta_{\phi} \) with \( \phi(w) = c \), the optimal \( c \), denoted by \( c^* \), is \( (p-2)/(n+2) \) and \( \delta_{c^*} \) is of course the James-Stein estimator. It is well-known that the James-Stein estimator is inadmissible since its positive-part estimator dominates it. In the following, we present a sufficient condition for dominance over the James-Stein estimator which is the generalization of Kubokawa’s [10] result.

**Theorem 3.2** Assume that \( \phi(w) \) is monotone nondecreasing, \( \lim_{w \to \infty} \phi(w) = (p-2)/(n+2) \) and \( \phi(w) \geq \phi_K(w) \) for every \( w \geq 0 \). Then \( \delta_{\phi} \) dominates the James-Stein estimator for unimodal spherically symmetric distributions.

**Proof:** By letting \( \Phi(w) = w^{-1}\phi\{(n+2)\phi - 2(p-2)\} - 4\phi'(1 + \phi) \) and by using the transformation to the polar coordinates, the second term of the right-hand side of (3.1) is written as

\[
\int_{R^p} \int_{R^n} \Phi(w) \sigma^{p-2} F((\|x - \theta\|^2 + \|z\|^2)/\sigma^2) dx dz
\]

\[
\begin{align*}
&= \int_{R^p} \int_{R^n} \Phi \left( \frac{\|x\|^2}{\|z\|^2} \right) F((\|x - \theta\|/\sigma)^2 + \|z\|^2) dx dz \\
&= C \int_0^\infty \int_0^\infty \int_0^{\pi} \Phi \left( \frac{s^2}{t^2} \right) F(s^2 - 2s\lambda^{1/2}\cos \varphi + \lambda + t^2) \\
&\quad \cdot \sin^{p-2} \varphi s^{p-1} t^{n-1} ds dt d\varphi \\
&= C \int_0^\infty \int_0^\infty \int_0^{\pi} \Phi(w) F(s^2 - 2s\lambda^{1/2}\cos \varphi + \lambda + s^2/w) \\
&\quad \cdot w^{-(n+1)/2} \sin^{p-2} \varphi w^{p+n-1} ds dw d\varphi,
\end{align*}
\]

where \( C = 4\pi^{(p+n-1)/2}/\{\Gamma((p-1)/2)\Gamma(n/2)\} \) and \( \lambda = \|\theta\|^2/\sigma^2 \). Letting

\[
g_\lambda(w) = \frac{w^{p/2-1}}{(1 + w)^{(p+n)/2}} \int_0^\pi \int_0^\infty F(u^2 - 2\lambda^{1/2}(1 + 1/w)^{-1/2} u \cos \varphi + \lambda) \\
\quad \cdot u^{p+n-1} \sin^{p-2} \varphi du d\varphi,
\]

we have

\[
R(\theta, \sigma^2, \delta_{JS}) - R(\theta, \sigma^2, \delta_{\phi}) = C \int_0^\infty \left( -\frac{(p-2)^2}{(n+2)w} - \Phi(w) \right) g_\lambda(w) dw.
\]

By a definite integral

\[
\left[ v(w) \int_0^w s^{-1} g_\lambda(s) ds \right]_0^\infty = \int_0^\infty v'(w) \int_0^w s^{-1} g_\lambda(s) ds dw + \int_0^\infty \frac{v(w)}{w} g_\lambda(w) dw
\]
for a differentiable function $v(w)$. Letting $(w) = \phi\{(n + 2)\phi - 2(p - 2)\}$ in the above equality and noting that $\lim_{w \to \infty} \phi(w) = (p - 2)/(n + 2)$, we have

$$R(\theta, \sigma^2, \delta_{JS}) - R(\theta, \sigma^2, \delta_\phi) = C \int_0^\infty \phi'(w) \left\{ (n + 2)\phi(w) - p + 2 \right\} \int_w^\infty s^{-1}g_\lambda(s)ds + 2(1 + \phi(w))g_\lambda(w) \right\} dw.$$ (3.2)

Since $g_\lambda(w)/g_0(w)$ is nondecreasing in $w$ by Lemma 3.3 in the below, we have

$$g_\lambda(w)/\int_0^w s^{-1}g_\lambda(s)ds \geq g_0(w)/\int_0^w s^{-1}g_0(s)ds,$$

and hence we have

$$R(\theta, \sigma^2, \delta_{JS}) - R(\theta, \sigma^2, \delta_\phi) \geq C \int_0^\infty \phi'(w) \left\{ (n + 2)\phi(w) - p + 2 + \frac{2(1 + \phi(w))g_0(w)}{\int_0^w s^{-1}g_0(s)ds} \right\} \int_0^w s^{-1}g_\lambda(s)ds dw.$$ (3.3)

Since $\phi_K(w)$ can be expressed as

$$\phi_K(w) = \frac{(p - 2)\int_0^w s^{-1}g_0(s)ds - 2g_0(w)}{(n + 2)\int_0^w s^{-1}g_0(s)ds + 2g_0(w)}$$ (3.3)

we have the theorem.

Combining (3.2) and (3.3), we have $R(0, \sigma^2, \delta_{JS}) = R(0, \sigma^2, \delta_K)$, which implies that $\delta_K$ can never dominate the James-Stein positive-part estimator unfortunately.

**Lemma 3.3** $g_\lambda(w)/g_0(w)$ is nondecreasing in $w$ if $f(\cdot)$ is monotone nonincreasing.

We note that the unimodality assumption corresponds to the fact the function $f$ is nonincreasing.

**Proof:** We have only to show that

$$h(w) = \int_0^\pi F(u^2 - 2\lambda^{1/2}(1 + 1/w)^{-1/2}u \cos \varphi + \lambda \sin^{p-2} \varphi \cos \varphi d\varphi,$$

for fixed $u$ and $\lambda$ is nondecreasing. The derivative of $h(w)$ is

$$\frac{d}{dw}h(w) = d\int_0^\pi f(u^2 - 2\lambda^{1/2}(1 + 1/w)^{-1/2}u \cos \varphi + \lambda \cos \varphi \sin^{p-2} \varphi \cos \varphi d\varphi$$

$$= d\int_0^{\pi/2} f(u^2 - 2\lambda^{1/2}(1 + 1/w)^{-1/2}u \cos \varphi + \lambda \cos \varphi \sin^{p-2} \varphi \cos \varphi d\varphi$$

$$- d\int_0^{\pi/2} f(u^2 + 2\lambda^{1/2}(1 + 1/w)^{-1/2}u \cos \varphi + \lambda \cos \varphi \sin^{p-2} \varphi \cos \varphi d\varphi$$
where \( d = 2^{-1}w^{-2}(1 + 1/w)^{-3/2}\lambda^{1/2}u \). The assumption of the lemma on \( f \) guarantees that \( h(w) \) is nondecreasing in \( w \).

Combining Theorem 3.1 and 3.2, we have the results on the decision-theoretic properties of \( \delta_{a,b} \).

**Theorem 3.4**

1. \( \delta_{a,b} \) is minimax for any spherically symmetric distributions if \( a > -n/2 - 1, \ 0 < b \leq p - 2 \) and \( b/(n + 2a + 2) < 2(p - 2)/(n + 2) \).

2. \( \delta_K \), which equals to \( \delta_{0,p-2} \), dominates the James-Stein estimator under the unimodal spherically symmetric distributions.

Note that among the minimax generalized Bayes estimators \( \delta_{a,b} \), only \( \delta_{0,p-2} \) satisfies the sufficient condition of Theorem 3.2. We thus recommend the use of \( \delta_K \) for any spherical symmetric distribution since it is minimax for any such \( f \), dominates the James-Stein estimator for those \( f \) which are unimodal, and is also generalized Bayes under the condition of a finite fourth moment.

**Remark 3.5**
The problem addressed in the paper is expected to have a relationship with some other estimation problems, in particular, the estimation of a scale parameter \( \sigma^2 \) with an unknown \( \theta \) and the estimation of \( \theta \) with known \( \sigma^2 \). But unfortunately we cannot establish a similar robustness property.

**Acknowledgements**

I would like to thank three referees for many valuable comments and helpful suggestions that led to an improved version of the paper.

**References**


Yuzo Maruyama  
Center for Spatial Information Science  
Faculty of Economics  
The University of Tokyo  
7–3–1 Hongo, Bunkyo–ku  
Tokyo, 113–0033, Japan  
maruyama@csis.u-tokyo.ac.jp