CIRJE-F-473

The Stein Phenomenon in Simultaneous Estimation of Normal Precisions

Hisayuki Tsukuma
Toho University

Tatsuya Kubokawa
University of Tokyo

February 2007; Revised in November 2007

Minimax estimation of normal precisions via expansion estimators

Hisayuki Tsukuma* and Tatsuya Kubokawa†
Toho University and University of Tokyo

November 22, 2007

Abstract

In this paper, the simultaneous estimation of the precision parameters of k normal distributions is considered under the squared loss function in a decision-theoretic framework. Several classes of minimax estimators are derived by using the chi-square identity, and the generalized Bayes minimax estimators are developed out of the classes. It is also shown that the improvement on the unbiased estimators is characterized by the superharmonic function. This corresponds to Stein (1981)'s result in simultaneous estimation of normal means.

Key words and phrases: Bayes estimation, chi-square identity, decision theory, empirical Bayes estimation, James-Stein estimator, risk function, simultaneous estimation, superharmonic function.

1 Introduction

This paper is concerned with the so-called Stein problem in the simultaneous estimation of normal precision parameters, namely, the reciprocals of the variances. The Stein problem, discovered by Stein (1956), is one of the most surprising and interesting phenomenon in theoretical statistics, and has been studied extensively in the literature. Most studies have addressed the problem in the framework of the simultaneous estimation of means of normal distributions. Of these, Stein (1973, 81) developed the so-called Stein identity, a very powerful tool for studying the shrinkage estimation, and derived the wonderful theory that the inadmissibility of the maximum likelihood estimator is characterized by superharmonic functions, which suggests a deep connection between the Stein problem and the potential theory. The Stein phenomenon has been extended to other distributions since Hudson (1978), who extended the Stein identity to the exponential family and showed the inadmissibility of unbiased estimators. Berger (1980), Ghosh and Parsian

*Faculty of Medicine, Toho University, 5-21-16 Omori-nishi, Ota-ku, Tokyo 143-8540, JAPAN, E-Mail: tsukuma@med.toho-u.ac.jp

†Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, JAPAN, E-Mail: tatsuya@e.u-tokyo.ac.jp

(1980), DasGupta (1986, 89) and Bilodeau (1988) provided dominance results in the gamma distribution, and an extension to the continuous exponential family was studied by Haff and Johnson (1986). Although these extensions correspond to the estimation of variances of normal distributions, no dominance results have been established for the precision parameters.

In this paper, we consider the k sample model that statistics $s_1, \ldots, s_k$ are mutually independent and distributed as $\sigma_i^2 \chi_{n_i}^2$ for $i = 1, \ldots, k$, where $\chi_{n_i}^2$ denotes the chi-square distribution with n_i degrees of freedom. Assume that $n_i - 4 > 0$ and σ_i^2 is unknown for $i = 1, \ldots, k$. It is noted that this is a canonical model of k sample model from normal distributions. Suppose that we want to estimate simultaneously the precision parameters $\boldsymbol{\sigma}^{-2} = (\sigma_1^{-2}, \ldots, \sigma_k^{-2})^t$, the reciprocals of the σ_i^2's, under the squared loss function

$$L(\boldsymbol{\delta}, \boldsymbol{\sigma}^{-2}) = \sum_{i=1}^{k} (\delta_i - \sigma_i^{-2})^2 = \|\boldsymbol{\delta} - \boldsymbol{\sigma}^{-2}\|^2, \tag{1.1}$$

where $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_k)^t$ is an estimator of $\boldsymbol{\sigma}^{-2}$. Every estimator is evaluated by the risk function $R(\boldsymbol{\delta}, \boldsymbol{\sigma}^{-2}) = E[L(\boldsymbol{\delta}, \boldsymbol{\sigma}^{-2})]$.

Usual estimators of $\boldsymbol{\sigma}^{-2} = (\sigma_1^{-2}, \ldots, \sigma_k^{-2})^t$ are represented as $\boldsymbol{\delta}^c = (c_1/s_1, \ldots, c_k/s_k)^t$ for positive constants c_i's. For example, the unbiased estimator is

$$\boldsymbol{\delta}^{UB} = ((n_1 - 2)/s_1, \ldots, (n_k - 2)/s_k)^t, \tag{1.2}$$

and the maximum likelihood estimator (MLE) is

$$\boldsymbol{\delta}^{ML} = (n_1/s_1, \ldots, n_k/s_k)^t.$$

It can be also seen that the best constant c_i with respect to the loss (1.1) is $c_i = n_i - 4$ for $i = 1, \ldots, k$, namely the best usual estimator is given by

$$\boldsymbol{\delta}^{BU} = ((n_1 - 4)/s_1, \ldots, (n_k - 4)/s_k)^t. \tag{1.3}$$

Since $\boldsymbol{\delta}^{BU}$ is minimax relative to the loss $\|\boldsymbol{\delta} - \boldsymbol{\sigma}^{-2}\|^2 / \sum_{i=1}^{k} \{(n_i - 2)\sigma_i^4\}^{-1}$, we here call $\boldsymbol{\delta}^{BU}$ the minimax estimator. The purpose of this paper is (1) to derive classes of estimators improving on $\boldsymbol{\delta}^{BU}$ and/or $\boldsymbol{\delta}^{UB}$, and (2) to obtain generalized Bayes estimators out of these classes.

The merit of considering the simultaneous estimation of the precision parameters $\boldsymbol{\sigma}^{-2}$ under the loss (1.1) is that unbiased estimators of the risk functions can be easily derived by using the chi-square identity, given by

$$E\left[(\boldsymbol{\delta}^{UB} - \boldsymbol{\sigma}^{-2})^t \boldsymbol{g}(\boldsymbol{s})\right] = -2E\left[\nabla^t \boldsymbol{g}(\boldsymbol{s})\right],$$

where $\nabla^t = (\partial/\partial s_1, \ldots, \partial/\partial s_k)$, $\boldsymbol{s} = (s_1, \ldots, s_k)^t$ and $\boldsymbol{g}(\boldsymbol{s}) = (g_1(\boldsymbol{s}), \ldots, g_k(\boldsymbol{s}))^t$ for absolutely continuous functions $g_i(\boldsymbol{s})$'s. In Section 4, we utilize this argument to get the superharmonic condition $\nabla^t \nabla \sqrt{f(\boldsymbol{s})} \leq 0$ for a positive function $f(\boldsymbol{s})$ under which the estimator $\boldsymbol{\delta}_f^{GB*} = \boldsymbol{\delta}^{UB} - 2\nabla \log f(\boldsymbol{s})$ improves on the unbiased estimator $\boldsymbol{\delta}^{UB}$. It is also shown that the generalized Bayes estimator can be written in the form $\boldsymbol{\delta}_f^{GB*}$ and that the improvement of the generalized Bayes estimators can be characterized by the superharmonic condition of the prior density. It is interesting to note that these results just correspond to the nice story developed by Stein (1973, 81) in the estimation of normal means, where the chi-square identity in our problem corresponds to the Stein identity in the normal distribution.

A different point between the two estimation problems of precision parameters and means is that the unbiased, the minimax and the maximum likelihood estimators are different in our problem as stated above, while they are identical to the sample means in estimation of means. In this sense, the nice story of Stein (1981) can be obtained for the improvement on the unbiased estimator. Since the unbiased estimator is dominated by the minimax estimator $\boldsymbol{\delta}^{BU}$, we want to investigate the dominance properties for $\boldsymbol{\delta}^{BU}$. Section 2 derives the general condition for the improvement on the minimax estimator $\boldsymbol{\delta}^{BU}$ through the unbiased estimator of the risk, and constructs several classes of improved estimators. In Section 3, we introduce a hierarchical prior distribution based on a truncated prior and show that the resulting generalized Bayes estimators belong to the class of minimax estimators.

Another interesting observation is that the inadmissibility of $\boldsymbol{\delta}^{BU}$ can be established for $k \geq 2$, while the sample means are inadmissible for $k \geq 3$ in the case of normal means. This phenomenon corresponds to Berger (1980)'s result in estimation of normal variances. Through the results in Sections 2 and 4, it is also observed that the minimax estimator $\boldsymbol{\delta}^{BU}$ can be dominated by expansion estimators while the unbiased estimator $\boldsymbol{\delta}^{UB}$ is dominated by both expansion and shrinkage estimators. In Section 5, through a geometric interpretation, we explain how these estimators can be improved on by expansion or shrinkage procedures. The performances of the risk functions of the estimators are investigated in Section 6 through the simulation studies.

The organization of the paper is as follows: Several classes of minimax estimators are constructed in Section 2, and out of the classes, the generalized Bayes minimax estimators against the hierarchical prior are obtained in Section 3. Section 4 derives the same dominance results as in estimation of normal means and provides the generalized Bayes estimators improving on the unbiased one. The geometric interpretation is given in Section 5, and numerical investigations are studied in Section 6.

2 Minimax estimation via expansion estimators

In this section, we characterize estimators improving on

$$\boldsymbol{\delta}^{BU} = ((n_1 - 4)/s_1, \dots, (n_k - 4)/s_k)^t,$$

which is the best among estimators $(a_1/s_1, \dots, a_k/s_k)^t$ for constants $a_1, \dots, a_k$ relative to the loss (1.1). We call $\boldsymbol{\delta}^{BU}$ the best usual estimator or minimax estimator, because it can be shown to be minimax relative to the loss $\|\boldsymbol{\delta} - \boldsymbol{\sigma}^{-2}\|^2 / \sum_{i=1}^k \{(n_i - 2)\sigma_i^4\}^{-1}$.

2.1 General condition for minimaxity

We begin with deriving the general condition for the improvement. Consider estimators of the general form

$$\boldsymbol{\delta}_g = \boldsymbol{\delta}^{BU} + \boldsymbol{g}(\boldsymbol{s}),$$

where $\boldsymbol{g}(\boldsymbol{s}) = (g_1(\boldsymbol{s}), \ldots, g_k(\boldsymbol{s}))^t$. Assume that $g_i(\boldsymbol{s})$ is absolutely continuous for $i = 1, \ldots, k$. To derive an unbiased estimator of the risk function of the estimator $\boldsymbol{\delta}_g$, we use the chi-square identity

$$E\left[\left(\frac{n_i - 2}{s_i} - \sigma_i^{-2}\right) g_i(\boldsymbol{s})\right] = E\left[-2\frac{\partial g_i(\boldsymbol{s})}{\partial s_i}\right], \tag{2.1}$$

which leads to the identity

$$E\left[\left(\frac{n_i - 4}{s_i} - \sigma_i^{-2}\right) g_i(\boldsymbol{s})\right] = E\left[-2\frac{g_i(\boldsymbol{s})}{s_i} - 2\frac{\partial g_i(\boldsymbol{s})}{\partial s_i}\right],$$

namely,

$$E\left[(\boldsymbol{\delta}^{BU} - \boldsymbol{\sigma}^{-2})^t \boldsymbol{g}(\boldsymbol{s})\right] = -2E\left[(\boldsymbol{s}^{-1})^t \boldsymbol{g}(\boldsymbol{s}) + \nabla^t \boldsymbol{g}(\boldsymbol{s})\right],$$

where $\boldsymbol{s}^{-1} = (s_1^{-1}, \ldots, s_k^{-1})^t$. The risk function of the estimator $\boldsymbol{\delta}_g$ is written as

$$\begin{aligned} R(\boldsymbol{\delta}_g, \boldsymbol{\sigma}^{-2}) &= R(\boldsymbol{\delta}^{BU}, \boldsymbol{\sigma}^{-2}) + E\left[\|\boldsymbol{g}(\boldsymbol{s})\|^2 + 2(\boldsymbol{\delta}^{BU} - \boldsymbol{\sigma}^{-2})^t \boldsymbol{g}(\boldsymbol{s})\right] \\ &= R(\boldsymbol{\delta}^{BU}, \boldsymbol{\sigma}^{-2}) + E\left[\|\boldsymbol{g}(\boldsymbol{s})\|^2 - 4\left\{(\boldsymbol{s}^{-1})^t \boldsymbol{g}(\boldsymbol{s}) + \nabla^t \boldsymbol{g}(\boldsymbol{s})\right\}\right], \end{aligned}$$

which shows the following proposition.

Proposition 2.1 *The estimator $\boldsymbol{\delta}_g$ is minimax, namely, better than $\boldsymbol{\delta}^{BU}$ if $\boldsymbol{g}(\boldsymbol{s})$ satisfies the inequality*

$$\|\boldsymbol{g}(\boldsymbol{s})\|^2 - 4(\boldsymbol{s}^{-1})^t \boldsymbol{g}(\boldsymbol{s}) - 4\nabla^t \boldsymbol{g}(\boldsymbol{s}) \le 0. \tag{2.2}$$

The general condition (2.2) gives us various classes of improved estimators given below.

2.2 James-Stein type estimators

Let $\boldsymbol{g}(\boldsymbol{s}) = \{\phi(\|\boldsymbol{s}\|^2)/\|\boldsymbol{s}\|^2\}\boldsymbol{s}$, and consider the estimators

$$\boldsymbol{\delta}_\phi^{JS} = \boldsymbol{\delta}^{BU} + \frac{\phi(\|\boldsymbol{s}\|^2)}{\|\boldsymbol{s}\|^2}\boldsymbol{s}.$$

Since $\nabla^t \boldsymbol{g}(\boldsymbol{s}) = (k-2)\phi(\|\boldsymbol{s}\|^2)/\|\boldsymbol{s}\|^2 + 2\phi'(\|\boldsymbol{s}\|^2)$, the inequality (2.2) is written as $\phi(\|\boldsymbol{s}\|^2)\left\{\phi(\|\boldsymbol{s}\|^2) - 8(k-1)\right\}/\|\boldsymbol{s}\|^2 - 8\phi'(\|\boldsymbol{s}\|^2) \le 0$.

Proposition 2.2 *The estimator $\boldsymbol{\delta}_\phi^{JS}$ is minimax if the nonnegative function $\phi(\boldsymbol{s})$ satisfies the inequality*

$$\phi(\|\boldsymbol{s}\|^2)\left\{\phi(\|\boldsymbol{s}\|^2) - 8(k-1)\right\} - 8\|\boldsymbol{s}\|^2\phi'(\|\boldsymbol{s}\|^2) \le 0,$$

for $k \ge 2$. This condition is satisfied for a non-decreasing function $\phi(u)$ with $0 \le \phi(u) \le 8(k-1)$ for $k \ge 2$ and $u > 0$.

Proposition 2.2 provides various improved estimators. For example, let $\phi(u) = 4(k-1)$. The resulting improved estimator is $\boldsymbol{\delta}^{JS} = (\delta_1^{JS}, \ldots, \delta_k^{JS})^t$ with

$$\delta_i^{JS} = \frac{n_i - 4}{s_i} + \frac{4(k-1)}{\|\boldsymbol{s}\|^2} s_i, \quad i = 1, \ldots, k, \tag{2.3}$$

which expands $\boldsymbol{\delta}^{BU}$. Since the modification term $\{4(k-1)/\|\boldsymbol{s}\|^2\}\boldsymbol{s}$ corresponds to the James-Stein (1961) estimator in estimation of normal means, we call $\boldsymbol{\delta}_\phi^{JS}$ the James-Stein type estimator. Note that for even $k = 2$ it is possible to improve upon $\boldsymbol{\delta}^{BU}$ and see also Berger (1980).

To present another example, let $\phi(u) = 4\{a - 2u(\mathrm{d}/\mathrm{d}u)\log h(u)\}$. The resulting estimator is minimax if $u(\mathrm{d}/\mathrm{d}u)\log h(u)$ is non-increasing and

$$0 < a - 2\lim_{u\to\infty}\left\{u\frac{\mathrm{d}}{\mathrm{d}u}\log h(u)\right\} \le 2(k-1).$$

In the case of $h(u) = (1+u)^{-b}$, it observed that $u(\mathrm{d}/\mathrm{d}u)\log h(u) = -bu/(1+u)$ is non-increasing for $b \ge 0$ and that $\lim_{u\to\infty} bu/(1+u) = b$. The estimator is minimax if $0 < a + 2b \le 2(k-1)$ for $b \ge 0$ and $k \ge 2$.

In the case of $h(u) = \log(1+u)$, $u(\mathrm{d}/\mathrm{d}u)\log h(u) = u/\{(1+u)\log(1+u)\}$, which can be verified to non-increasing. Since $\lim_{u\to\infty} u/\{(1+u)\log(1+u)\} = 0$, the resulting estimator is minimax if $0 < a \le 2(k-1)$.

2.3 Empirical Bayes estimators

Let $\boldsymbol{g}(\boldsymbol{s}) = \{\phi(T)/T\}\mathbf{1}_k$ for $T = \sum_{i=1}^k s_i$ and the k-dimensional vector $\mathbf{1}_k = (1, \ldots, 1)^t$, and consider the estimators

$$\boldsymbol{\delta}_\phi^{EB} = \boldsymbol{\delta}^{BU} + \frac{\phi(T)}{T}\mathbf{1}_k. \tag{2.4}$$

In this case, $\nabla^t \boldsymbol{g}(\boldsymbol{s}) = k\{\phi'(T)/T - \phi(T)/T^2\}$, so that the inequality (2.2) is written as

$$k\frac{\phi^2(T)}{T^2} - 4k\left\{\frac{\phi'(T)}{T} - \frac{\phi(T)}{T^2}\right\} - 4\left(\sum_{i=1}^k s_i^{-1}\right)\frac{\phi(T)}{T} \le 0. \tag{2.5}$$

For the arithmetic and the harmonic means of the s_i's, it is noted that $(1/k)\sum_{i=1}^k s_i \ge k/(\sum_{i=1}^k s_i^{-1})$, namely, $(\sum_{i=1}^k s_i^{-1})(\sum_{i=1}^k s_i) \ge k^2$. If $\phi(T)$ is nonnegative, then the l.h.s. of (2.5) is less than or equal to

$$(k/T^2)\big[\phi(T)\{\phi(T) - 4(k-1)\} - 4T\phi'(T)\big].$$

Proposition 2.3 *The estimator $\boldsymbol{\delta}_\phi^{EB}$ is minimax if the nonnegative function $\phi(T)$ satisfies the inequality*

$$\phi(T)\{\phi(T) - 4(k-1)\} - 4T\phi'(T) \le 0,$$

for $k \ge 2$. This condition is satisfied for a non-decreasing function $\phi(T)$ with $0 \le \phi(T) \le 4(k-1)$ for $k \ge 2$.

Letting $\phi(T) = 2(k-1)$, we get the simple improved estimator $\boldsymbol{\delta}^{EB} = (\delta_1^{EB}, \ldots, \delta_k^{EB})^t$, where

$$\delta_i^{EB} = \frac{n_i - 4}{s_i} + \frac{2(k-1)}{\sum_{j=1}^k s_j}, \quad i = 1, \ldots, k. \tag{2.6}$$

It is noted that $\boldsymbol{\delta}^{EB}$ is an expansion estimator of $\boldsymbol{\delta}^{BU}$. Since it is rewritten as

$$\delta_i^{EB} = \frac{n_i - 4}{s_i} \left(1 + \frac{2(k-1)n_i}{n_i - 4} \times \frac{s_i/n_i}{\sum_{j=1}^k s_j} \right),$$

it is observed that δ_i^{EB} gives more expansion for larger sample variance s_i/n_i.

We shall show that $\boldsymbol{\delta}^{EB}$ can be characterized as an empirical Bayes estimator. Let $\eta_i = \sigma_i^{-2}$ for $i = 1, \ldots, k$ and let $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_k)^t$. The joint density function of $\boldsymbol{s}$ given $\boldsymbol{\eta}$ is

$$p(\boldsymbol{s}|\boldsymbol{\eta}) = C_0 \prod_{i=1}^k \left\{ \eta_i^{n_i/2} s_i^{n_i/2-1} e^{-s_i \eta_i/2} \right\},$$

for $C_0 = 1/\prod_{i=1}^k \left\{ \Gamma(n_i/2) 2^{n_i/2} \right\}$. Suppose that the prior density function of $\boldsymbol{\eta}$ for hyperparameter $\gamma > 0$ is given by

$$\pi(\boldsymbol{\eta}|\gamma) = C_1 \prod_{i=1}^k \left\{ \gamma^2 (\eta_i - \gamma)^{(n_i-6)/2} \eta_i^{-n_i/2} I(\eta_i > \gamma) \right\}, \tag{2.7}$$

where $C_1 = \prod_{i=1}^k \{(n_i - 2)(n_i - 4)/4\}$, and $I(A)$ is the indicator function such that $I(A) = 1$ if A is true, and $= 0$ otherwise. Then the posterior density of $\boldsymbol{\eta}$ given $\boldsymbol{s}$ and γ is

$$\pi(\boldsymbol{\eta}|\boldsymbol{s}, \gamma) = C_2 \prod_{i=1}^k \left\{ s_i^{(n_i-4)/2} (\eta_i - \gamma)^{(n_i-4)/2-1} e^{-s_i(\eta_i - \gamma)/2} I(\eta_i > \gamma) \right\},$$

for $C_2 = 1/\prod_{i=1}^k \{\Gamma((n_i - 4)/2) 2^{(n_i-4)/2}\}$, and the resulting Bayes estimator is

$$\delta_i^B(\gamma) = E\left[\eta_i|\boldsymbol{s}, \gamma\right] = E\left[\eta_i - \gamma|\boldsymbol{s}, \gamma\right] + \gamma = \frac{n_i - 4}{s_i} + \gamma. \tag{2.8}$$

Since the hyperparameter γ is unknown, it is estimated from the marginal distribution of $\boldsymbol{s}$ and γ. The marginal density of $\boldsymbol{s}$ given γ is

$$p_\pi(\boldsymbol{s}|\gamma) = \prod_{i=1}^k \left\{ \frac{1}{\Gamma(4/2)} (\gamma/2)^2 s_i e^{-\gamma s_i/2} \right\},$$

which means that $s_1, \ldots, s_k$ are marginally i.i.d. $\gamma^{-1}\chi_4^2$. Since $T = s_1 + \cdots + s_k$ has $\gamma^{-1}\chi_{4k}^2$, γ is estimated by $\hat{\gamma}_a = a/T$ for constant a. Minimizing the risk among estimators of the form $\boldsymbol{\delta}^{BU} + (a/T)\mathbf{1}_k$ with respect to a, we see that the best constant is given by $a = 2(k-1)$. Substituting the estimator $2(k-1)/T$ into (2.8) yields the empirical Bayes estimator which is identical to $\boldsymbol{\delta}^{EB}$.

2.4 Generalized Bayes estimators

Let $\boldsymbol{g}(\boldsymbol{s}) = -2\nabla \log q(\boldsymbol{s}) = -2\nabla q(\boldsymbol{s})/q(\boldsymbol{s})$ for a positive function $q(\boldsymbol{s})$, and consider the estimators

$$\boldsymbol{\delta}_q^{GB} = \boldsymbol{\delta}^{BU} - 2\nabla \log q(\boldsymbol{s}). \tag{2.9}$$

Note that

$$\nabla^t \boldsymbol{g}(\boldsymbol{s}) = -2\nabla^t \left\{ \frac{\nabla q(\boldsymbol{s})}{q(\boldsymbol{s})} \right\} = -2 \left\{ \frac{\nabla^t \nabla q(\boldsymbol{s})}{q(\boldsymbol{s})} - \frac{\|\nabla q(\boldsymbol{s})\|^2}{\{q(\boldsymbol{s})\}^2} \right\}. \tag{2.10}$$

Then the inequality (2.2) is written as

$$\{4/q(\boldsymbol{s})\} \left\{ 2(\boldsymbol{s}^{-1})^t \nabla q(\boldsymbol{s}) + 2\nabla^t \nabla q(\boldsymbol{s}) - \|\nabla q(\boldsymbol{s})\|^2/q(\boldsymbol{s}) \right\} \le 0.$$

Proposition 2.4 *The estimator $\boldsymbol{\delta}_q^{GB}$ is minimax if the positive function $q(\boldsymbol{s})$ satisfies the inequality*

$$2(\boldsymbol{s}^{-1})^t \nabla q(\boldsymbol{s}) + 2\nabla^t \nabla q(\boldsymbol{s}) - \|\nabla q(\boldsymbol{s})\|^2/q(\boldsymbol{s}) \le 0. \tag{2.11}$$

This inequality is satisfied by

$$(\boldsymbol{s}^{-1})^t \nabla q(\boldsymbol{s}) + \nabla^t \nabla q(\boldsymbol{s}) \le 0. \tag{2.12}$$

Since it can be seen that

$$2\nabla^t \nabla q(\boldsymbol{s}) - \|\nabla q(\boldsymbol{s})\|^2/q(\boldsymbol{s}) = 4\sqrt{q(\boldsymbol{s})}\nabla^t \nabla \sqrt{q(\boldsymbol{s})}, \tag{2.13}$$

the condition (2.11) is also expressed as

$$(\boldsymbol{s}^{-1})^t \nabla \sqrt{q(\boldsymbol{s})} + \nabla^t \nabla \sqrt{q(\boldsymbol{s})} \le 0.$$

The function $\sqrt{q(\boldsymbol{s})}$ being superharmonic is equivalent to the inequality $\nabla^t \nabla \sqrt{q(\boldsymbol{s})} \le 0$. Stein (1973, 81) showed that the superharmonicity of a function of the sample means implies the minimaxity in estimation of the normal mean vector. In our problem, however, this story does not hold, because the condition (2.12) depends on $(\boldsymbol{s}^{-1})^t \nabla q(\boldsymbol{s})$, the inner product of the vectors $\boldsymbol{s}^{-1}$ and $\nabla q(\boldsymbol{s})$.

We shall show that a generalized Bayes estimator can be expressed by (2.9). Assume that the prior density of $\boldsymbol{\eta}$ is given by $\pi(\boldsymbol{\eta})$, and let $g(\boldsymbol{\eta}) = \pi(\boldsymbol{\eta}) \prod_{i=1}^k \eta_i^3$. Then the generalized Bayes estimator of $\boldsymbol{\eta} = \boldsymbol{\sigma}^{-1}$ is

$$\boldsymbol{\delta}_\pi^{GB} = \frac{\int \boldsymbol{\eta} \{\prod_{i=1}^k \eta_i^{(n_i-4)/2-1} e^{-\eta_i s_i/2}\} g(\boldsymbol{\eta}) \mathrm{d}\boldsymbol{\eta}}{\int \{\prod_{i=1}^k \eta_i^{(n_i-4)/2-1} e^{-\eta_i s_i/2}\} g(\boldsymbol{\eta}) \mathrm{d}\boldsymbol{\eta}}.$$

Note that $\boldsymbol{\delta}_\pi^{GB} = \boldsymbol{\delta}^{BU}$ when $g(\boldsymbol{\eta}) \propto 1$. The generalized Bayes estimator can be written as

$$\begin{aligned} \boldsymbol{\delta}_\pi^{GB} &= \boldsymbol{\delta}^{BU} + (\boldsymbol{\delta}_\pi^{GB} - \boldsymbol{\delta}^{BU}) \\ &= \boldsymbol{\delta}^{BU} - 2\nabla \log \left(\left\{ \prod_{i=1}^k s_i^{(n_i-4)/2} \right\} \int \left\{ \prod_{i=1}^k \eta_i^{(n_i-4)/2-1} e^{-\eta_i s_i/2} \right\} g(\boldsymbol{\eta}) \mathrm{d}\boldsymbol{\eta} \right) \\ &= \boldsymbol{\delta}^{BU} - 2\nabla \log q_\pi(\boldsymbol{s}), \quad \text{say,} \end{aligned}$$

where the function $q_\pi(\boldsymbol{s})$ is given by

$$q_\pi(\boldsymbol{s}) = \int \Big\{ \prod_{i=1}^k \tau_i^{(n_i-4)/2-1} e^{-\tau_i/2} \Big\} g(\boldsymbol{\tau}/\boldsymbol{s}) \mathrm{d}\boldsymbol{\tau},$$

for $\boldsymbol{\tau}/\boldsymbol{s} = (\tau_1/s_1, \ldots, \tau_k/s_k)^t$. This shows that any generalized Bayes estimator is expressed in the form (2.9).

It is noted that $q_\pi(\boldsymbol{s})$ is not a marginal density. From Proposition 2.4, it is seen that the generalized Bayes estimator $\boldsymbol{\delta}_\pi^{GB}$ dominates $\boldsymbol{\delta}^{BU}$ if

$$(\boldsymbol{s}^{-1})^t \nabla q_\pi(\boldsymbol{s}) + \nabla^t \nabla q_\pi(\boldsymbol{s}) \le 0, \tag{2.14}$$

which provides the general characterization of the prior distribution.

Corollary 2.1 *If the prior distribution* $\pi(\boldsymbol{\eta}) = g(\boldsymbol{\eta}) \prod_{i=1}^k \eta_i^{-3}$ *satisfies the condition*

$$(\boldsymbol{s}^{-1})^t \nabla g(\boldsymbol{\tau}/\boldsymbol{s}) + \nabla^t \nabla g(\boldsymbol{\tau}/\boldsymbol{s}) < 0,$$

then $\boldsymbol{\delta}_\pi^{GB}$ *dominates* $\boldsymbol{\delta}^{BU}$ *relative to the loss* (1.1).

3 Bayesian minimax methods against a truncated prior

We now derive generalized Bayes estimators improving on the best usual one $\boldsymbol{\delta}^{BU}$. Assume that the following hierarchical prior distribution $\pi(\boldsymbol{\eta}) = \int \pi_1(\boldsymbol{\eta}|\gamma)\pi_2(\gamma)\mathrm{d}\gamma$:

$$\begin{aligned} \boldsymbol{\eta}|\gamma \sim & \pi_1(\boldsymbol{\eta}|\gamma) = C_1 \prod_{j=1}^k \Big\{ \gamma^2 (\eta_j - \gamma)^{(n_j-6)/2} \eta_j^{-n_j/2} I(\eta_j > \gamma) \Big\}, \\ \gamma \sim & \pi_2(\gamma), \end{aligned} \tag{3.1}$$

where C_1 is defined below (2.7), and the second stage prior $\pi_2(\gamma)$ will be specified later. The posterior density is written as

$$\pi(\boldsymbol{\eta}|\boldsymbol{s}) = C \int \prod_{j=1}^k \big\{ \gamma^2 (\eta_j - \gamma)^{(n_j-6)/2} e^{-s_j \eta_j/2} I(\eta_j > \gamma) \big\} \pi_2(\gamma) \mathrm{d}\gamma,$$

and the Bayes estimator is given by

$$\delta_i^{GB} = C \int \int (\eta_i - \gamma + \gamma) \prod_{j=1}^k \big\{ \gamma^2 (\eta_j - \gamma)^{(n_j-6)/2} e^{-s_j \eta_j/2} I(\eta_j > \gamma) \big\} \pi_2(\gamma) \mathrm{d}\gamma \mathrm{d}\boldsymbol{\eta}. \tag{3.2}$$

Making the transformation $\xi_j = \eta_j - \gamma$ with $|\partial \boldsymbol{\xi}/\partial \boldsymbol{\eta}| = 1$ for $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_k)^t$, we see that

$$\begin{aligned} & \int \int (\eta_i - \gamma) \prod_{j=1}^k \big\{ \gamma^2 (\eta_j - \gamma)^{(n_j-6)/2} e^{-s_j \eta_j/2} I(\eta_j > \gamma) \big\} \pi_2(\gamma) \mathrm{d}\boldsymbol{\eta} \mathrm{d}\gamma \\ & = \int \xi_i \prod_{j=1}^k \{ \xi_j^{(n_j-6)/2} e^{-s_j \xi_j/2} \} \mathrm{d}\boldsymbol{\xi} \times \int \gamma^{2k} e^{-(\sum_{j=1}^k s_j)\gamma/2} \pi_2(\gamma) \mathrm{d}\gamma, \end{aligned}$$

and that

$$\int\int \gamma \prod_{j=1}^{k} \left\{ \gamma^2(\eta_j-\gamma)^{(n_j-6)/2} e^{-s_j\eta_j/2} I(\eta_j>\gamma) \right\} \pi_2(\gamma) \mathrm{d}\boldsymbol{\eta} \mathrm{d}\gamma$$
$$= \int \prod_{j=1}^{k} \{\xi_j^{(n_j-6)/2} e^{-s_j\xi_j/2}\} \mathrm{d}\boldsymbol{\xi} \times \int \gamma^{2k+1} e^{-(\sum_{j=1}^{k} s_j)\gamma/2} \pi_2(\gamma)\mathrm{d}\gamma,$$

so that from (3.2), the generalized Bayes estimator δ_i^{GB} is expressed as

$$\delta_i^{GB} = \frac{n_i-4}{s_i} + \frac{\int \gamma^{2k+1} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}, \tag{3.3}$$

for $T = \sum_{j=1}^{k} s_j$. The following theorem provides a condition for $\boldsymbol{\delta}^{GB} = (\delta_1^{GB}, \dots, \delta_k^{GB})^t$ to dominate $\boldsymbol{\delta}^{BU}$.

Theorem 3.1 *Assume that $\pi_2(\gamma)$ is an absolutely continuous function on $(0,\infty)$ such that $\gamma\pi_2'(\gamma)/\pi_2(\gamma)$ is non-increasing in γ, and $\int \gamma^{2k} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma < \infty$. If $\pi_2(\gamma)$ satisfies the condition*

$$\lim_{T\to\infty} R(T) \geq 3, \tag{3.4}$$

where $R(T)$ is defined by

$$R(T) \equiv -\frac{\int \gamma^{2k+1} e^{-T\gamma/2}\pi_2'(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma},$$

then the generalized Bayes estimator $\boldsymbol{\delta}^{GB}$ given by (3.2) or (3.3) against the prior (3.1) is minimax.

Proof. To prove the theorem, we use Proposition 2.3, namely, we show that

$$\phi(T)\{\phi(T) - 4(k-1)\} - 4T\phi'(T) \leq 0. \tag{3.5}$$

The generalized Bayes estimator $\boldsymbol{\delta}^{GB}$ is expressed in the form (2.4) by putting

$$\phi(T) = T\frac{\int \gamma^{2k+1} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}.$$

Note that

$$\begin{aligned}\phi'(T) =& \frac{\int \gamma^{2k+1} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma} - \frac{T}{2}\frac{\int \gamma^{2k+2} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma} \\ &+ \frac{T}{2}\left(\frac{\int \gamma^{2k+1} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}\right)^2.\end{aligned}$$

Using the expressions $\phi(T)$ and $\phi'(T)$, we can rewrite the inequality (3.5) as

$$2T\frac{\int \gamma^{2k+2} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k+1} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma} - T\frac{\int \gamma^{2k+1} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k} e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma} \leq 4k. \tag{3.6}$$

By the integration by parts, it is noted that for constants c_1, c_2 and d,

$$\begin{aligned}\int_{c_1}^{c_2} \gamma^d e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma =& -\frac{2}{T}\left[\gamma^d e^{-T\gamma/2}\pi_2(\gamma)\right]_{c_1}^{c_2} \\ &+\frac{2}{T}\int_{c_1}^{c_2}\left\{d\gamma^{d-1}e^{-T\gamma/2}\pi_2(\gamma)+\gamma^d e^{-T\gamma/2}\pi_2'(\gamma)\right\}\mathrm{d}\gamma. \qquad (3.7)\end{aligned}$$

Note that $\pi_2(\gamma)$ is an absolutely continuous function such that $\lim_{\gamma\to 0}\gamma^d e^{-T\gamma/2}\pi_2(\gamma) = \lim_{\gamma\to\infty}\gamma^d e^{-T\gamma/2}\pi_2(\gamma) = 0$. Then,

$$\begin{aligned}T\frac{\int \gamma^{2k+2}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma} =& 2(2k+2)+2\frac{\int \gamma^{2k+2}e^{-T\gamma/2}\pi_2'(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}, \\ T\frac{\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma} =& 2(2k+1)+2\frac{\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2'(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}.\end{aligned}$$

Hence, the inequality (3.6) is equivalent to

$$\frac{\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2'(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma} - 2\frac{\int \gamma^{2k+2}e^{-T\gamma/2}\pi_2'(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma} \geq 3. \qquad (3.8)$$

Since $\gamma\pi_2'(\gamma)/\pi_2(\gamma)$ is non-increasing in γ, the following inequality holds:

$$\frac{\int \gamma^{2k+2}e^{-T\gamma/2}\pi_2'(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma} \leq \frac{\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2'(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma}. \qquad (3.9)$$

In fact, this inequality is equivalent to

$$E^*[\gamma^2\pi_2'(\gamma)/\pi_2(\gamma)] \leq E^*[\gamma]E^*[\gamma\pi_2'(\gamma)/\pi_2(\gamma)],$$

where $E^*[\cdot]$ is the expectation with respect to the probability

$$P^*(A) = \int_A \gamma^{2k}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma \Big/ \int \gamma^{2k}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma.$$

Since γ and $\gamma\pi_2'(\gamma)/\pi_2(\gamma)$ are monotone in opposite directions, the inequality (3.9) is guaranteed. Using (3.9), we see that the inequality (3.8) is satisfied if

$$-\frac{\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2'(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma} \equiv R(T) \geq 3. \qquad (3.10)$$

By differentiating $R(T)$ with respect to T, the derivative of $R(T)$ is proportional to

$$\begin{aligned}&\int \gamma^{2k+2}e^{-T\gamma/2}\pi_2'(\gamma)\mathrm{d}\gamma\int \gamma^{2k}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma \\ &-\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2'(\gamma)\mathrm{d}\gamma\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma\end{aligned}$$

which is not positive from the inequality (3.9). That is, $R(T)$ is non-increasing in T when $\gamma\pi_2'(\gamma)/\pi_2(\gamma)$ is non-increasing in γ. Hence form (3.10), we obtain the sufficient condition $\lim_{T\to\infty} R(T) \geq 3$, which is given by (3.4), and the proof is complete. ∎

When the second stage prior is given by $\pi_2(\gamma) = \gamma^{-3}g(\gamma)$ for an absolutely continuous function $g(\gamma)$, $R(T)$ given below (3.4) can be written as

$$R(T) = 3 - \frac{\int \gamma^{2k-2}e^{-T\gamma/2}g'(\gamma)\mathrm{d}\gamma}{\int \gamma^{2k-3}e^{-T\gamma/2}g(\gamma)\mathrm{d}\gamma},$$

so that we get the following corollary.

Corollary 3.1 *Assume that $g(\gamma)$ is an absolutely continuous function on $(0,\infty)$ such that $\gamma g'(\gamma)/g(\gamma)$ is non-increasing in γ, and $\int \gamma^{2k-3}e^{-T\gamma/2}g(\gamma)\mathrm{d}\gamma < \infty$. If $g'(\gamma) \leq 0$ for any $\gamma > 0$, then the generalized Bayes estimator $\boldsymbol{\delta}^{GB}$ given by (3.3) against the prior (3.1) with $\pi_2(\gamma) = \gamma^{-3}g(\gamma)$ is minimax.*

Theorem 3.1 and Corollary 3.1 provide various prior distributions which result in the generalized minimax estimators.

[1] Let $\pi_2(\gamma) = \gamma^{-a-3}$. Then, $g(\gamma) = \gamma^{-a}$, $\gamma g'(\gamma)/g(\gamma) = -a$, and the resulting estimator is

$$\boldsymbol{\delta}^{GB} = \boldsymbol{\delta}^{BU} + \frac{2\{2(k-1)-a\}}{T}\mathbf{1}_k. \tag{3.11}$$

Since the prior is improper, $\boldsymbol{\delta}^{GB}$ is generalized Bayes. From Corollary 3.1, $\boldsymbol{\delta}^{GB}$ is minimax if $0 \leq a < 2(k-1)$ for $k \geq 2$. It is interesting to note that for $a = k-1$, $\boldsymbol{\delta}^{GB}$ is identical to the empirical Bayes estimator $\boldsymbol{\delta}^{EB}$ given by (2.6).

[2] Let $\pi_2(\gamma) = \gamma^{-3}(1+\gamma)^{-a}$. Then, $g(\gamma) = (1+\gamma)^{-a}$, $\gamma g'(\gamma)/g(\gamma) = -a\gamma/(1+\gamma)$, and the resulting estimator is

$$\boldsymbol{\delta}^{GB} = \boldsymbol{\delta}^{BU} + \frac{1}{T}\frac{\int z^{2k-2}(1+z/T)^{-a}e^{-z/2}\mathrm{d}z}{\int z^{2k-3}(1+z/T)^{-a}e^{-z/2}\mathrm{d}z}\mathbf{1}_k. \tag{3.12}$$

The generalized Bayes estimator $\boldsymbol{\delta}^{GB}$ is minimax if $a \geq 0$ for $k \geq 2$.

As the condition for the minimaxity of $\boldsymbol{\delta}^{GB}$, Theorem 3.1 requires the monotonicity of $\gamma\pi_2'(\gamma)/\pi_2(\gamma)$ on the second stage prior $\pi_2(\gamma)$. Using the condition (2.14), however, we can derive a condition for the minimaxity without assuming the monotonicity. It is noted that the generalized Bayes estimator (3.3) can be expressed as

$$\boldsymbol{\delta}^{GB} = \boldsymbol{\delta}^{BU} - 2\nabla \log q_\pi(\boldsymbol{s}),$$

where $q_\pi(\boldsymbol{s}) = \int \gamma^{2k}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma$.

Proposition 3.1 *The generalized Bayes estimator $\boldsymbol{\delta}^{GB}$ is minimax if the second stage prior $\pi_2(\gamma)$ is an absolutely continuous function on $(0,\infty)$ satisfying the condition*

$$-\gamma\pi_2'(\gamma)/\pi_2(\gamma) \geq k+2. \tag{3.13}$$

Proof. We need to evaluate the l.h.s. of (2.14) as

$$(\boldsymbol{s}^{-1})^t \nabla q_\pi(\boldsymbol{s}) + \nabla^t \nabla q_\pi(\boldsymbol{s})$$
$$= -\frac{1}{2}\sum_{i=1}^k \frac{1}{s_i}\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma + \frac{k}{4}\int \gamma^{2k+2}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma,$$

which is less than or equal to

$$\frac{k}{2T}\left\{-k\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma + \frac{T}{2}\int \gamma^{2k+2}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma\right\}, \tag{3.14}$$

since $\sum_{i=1}^k s_i^{-1}\sum_{i=1}^k s_i \geq k^2$. From the equation (3.7), it is seen that (3.14) is equal to

$$\frac{k}{2T}\left\{-k\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma + (2k+2)\int \gamma^{2k+1}e^{-T\gamma/2}\pi_2(\gamma)\mathrm{d}\gamma\right.$$
$$\left. + \int \gamma^{2k+2}e^{-T\gamma/2}\pi_2'(\gamma)\mathrm{d}\gamma\right\}$$
$$= \frac{k}{2T}\int \left\{(k+2)\pi_2(\gamma) + \gamma\pi_2'(\gamma)\right\}\gamma^{2k+1}e^{-T\gamma/2}\mathrm{d}\gamma,$$

which proves Proposition 3.1 from (2.14). ■

As an example of Proposition 3.1, consider the function $\pi_2(\gamma) = \gamma^{-(a+3)}\exp\{2b\gamma - c\gamma^2/2\}$ for nonnegative constants a, b and c. Then it is observed that

$$-\gamma\frac{\mathrm{d}}{\mathrm{d}\gamma}\log\pi_2(\gamma) = a + 3 + c\gamma^2 - 2b\gamma,$$

which is a quadratic function of γ. For $c > 0$, this is larger than or equal to $a + 3 - b^2/c$. Hence from (3.13), the condition for the minimaxity is given by $b^2/c + k - 1 \leq a < 2(k-1)$. When $c = b = 0$, the condition is $k - 1 \leq a < 2(k-1)$, which is more restrictive than the condition given below (3.11).

The derivation of admissible and minimax estimators is one of the main goals in estimation in the decision-theoretic framework. Since we need substantial additional work to establish the admissibility of the generalized Bayes and minimax estimators given in this section, we do not discuss the admissibility problem in this paper. For this purpose, it would be interesting to find a proper prior such that the resulting Bayes estimator is minimax.

4 Improvement on the unbiased estimators

As stated in Section 1, the unbiased estimator $\boldsymbol{\delta}^{UB}$ of $\boldsymbol{\sigma}^{-2}$ is given by (1.2), and it is dominated by the best usual estimator $\boldsymbol{\delta}^{BU}$. Although $\boldsymbol{\delta}^{UB}$ is inferior to $\boldsymbol{\delta}^{BU}$, the problem of improving on $\boldsymbol{\delta}^{UB}$ has a very nice structure corresponding to the dominance results of Stein (1981) in the simultaneous estimation of a mean vector of a multivariate normal distribution. It is also interesting to note that $\boldsymbol{\delta}^{UB}$ can be dominated by expansion or shrinkage estimators and their combined methods, while $\boldsymbol{\delta}^{BU}$ is dominated by expansion estimators. Thus, in this section, we provide a summary of the dominance results over $\boldsymbol{\delta}^{UB}$.

For improving on $\boldsymbol{\delta}^{UB}$, consider estimators of the form $\boldsymbol{\delta}_g^* = \boldsymbol{\delta}^{UB} - \boldsymbol{g}(\boldsymbol{s})$ for $\boldsymbol{g}(\boldsymbol{s}) = (g_1(\boldsymbol{s}), \ldots, g_k(\boldsymbol{s}))^t$. Using the chi-square identity (2.1), we can write the risk function of the estimator $\boldsymbol{\delta}_g$ as

$$R(\boldsymbol{\delta}_g^*, \boldsymbol{\sigma}^{-2}) = R(\boldsymbol{\delta}^{UB}, \boldsymbol{\sigma}^{-2}) + E\left[\|\boldsymbol{g}(\boldsymbol{s})\|^2 + 4\nabla^t \boldsymbol{g}(\boldsymbol{s})\right],$$

which shows the following proposition.

Proposition 4.1 *The estimator $\boldsymbol{\delta}_g^*$ is better than $\boldsymbol{\delta}^{UB}$ if $\boldsymbol{g}(\boldsymbol{s})$ satisfies the inequality*

$$\|\boldsymbol{g}(\boldsymbol{s})\|^2 + 4\nabla^t \boldsymbol{g}(\boldsymbol{s}) \le 0.$$

Consider the case of $\boldsymbol{g}(\boldsymbol{s}) = 2\nabla \log f(\boldsymbol{s})$ where $f(\cdot)$ is a twice differentiable function and $f(\cdot) > 0$. From (2.10) and (2.13), it is observed that

$$\|\boldsymbol{g}(\boldsymbol{s})\|^2 + 4\nabla^t \boldsymbol{g}(\boldsymbol{s}) = 8\frac{\nabla^t\nabla f(\boldsymbol{s})}{f(\boldsymbol{s})} - 4\frac{\|\nabla f(\boldsymbol{s})\|^2}{\{f(\boldsymbol{s})\}^2} = \frac{16}{\sqrt{f(\boldsymbol{s})}}\nabla^t\nabla\sqrt{f(\boldsymbol{s})}, \tag{4.1}$$

which yields the following dominance result.

Proposition 4.2 *Assume that $\sqrt{f(\boldsymbol{s})}$ is superharmonic, namely, $\nabla^t\nabla\sqrt{f(\boldsymbol{s})} < 0$. Then the estimator*

$$\boldsymbol{\delta}_f^{GB*} = \boldsymbol{\delta}^{UB} - 2\nabla \log f(\boldsymbol{s}), \tag{4.2}$$

dominates the unbiased estimator $\boldsymbol{\delta}^{UB}$ under the loss (1.1). *The inequality $\nabla^t\nabla f(\boldsymbol{s}) < 0$ implies that $\nabla^t\nabla\sqrt{f(\boldsymbol{s})} < 0$.*

These propositions give us several classes of improved estimators as described below.

[1] Generalized Bayes estimator. Consider the generalized Bayes estimator against the prior $\pi(\boldsymbol{\eta}) = h(\boldsymbol{\eta})\prod_{i=1}^k \eta_i^{-2}$. Then the resulting generalized Bayes estimator can be expressed as

$$\boldsymbol{\delta}_\pi^{GB*} = \boldsymbol{\delta}^{UB} - 2\nabla \log f_\pi(\boldsymbol{s}), \tag{4.3}$$

where

$$f_\pi(\boldsymbol{s}) = \int \Big\{\prod_{i=1}^k \tau_i^{(n_i-2)/2-1} e^{-\tau_i/2}\Big\} h(\boldsymbol{\tau}/\boldsymbol{s}) \mathrm{d}\boldsymbol{\tau}.$$

Noting that $\nabla^t\nabla f_\pi(\boldsymbol{s}) = \int\{\prod_{i=1}^k \tau_i^{(n_i-2)/2-1} e^{-\tau_i/2}\}\nabla^t\nabla h(\boldsymbol{\tau}/\boldsymbol{s})\mathrm{d}\boldsymbol{\tau}$, we can see that if $\nabla^t\nabla h(\boldsymbol{\tau}/\boldsymbol{s}) \le 0$, then $\boldsymbol{\delta}_\pi^{GB*}$ dominates $\boldsymbol{\delta}^{UB}$ relative to the loss (1.1). Although the condition $\nabla^t\nabla h(\boldsymbol{\tau}/\boldsymbol{s}) \le 0$ is general, it is more restrictive than the condition $\nabla^t\nabla\sqrt{f_\pi(\boldsymbol{s})} < 0$, which will be checked below by a specific prior. The prior considered here is of the form

$$\pi(\boldsymbol{\eta}) = h(\boldsymbol{\eta})\prod_{i=1}^k \eta_i^{-2}, \quad \text{for} \quad h(\boldsymbol{\eta}) = \left(\prod_{i=1}^k \eta_i\right)^{-a}\left(\sum_{i=1}^k \eta_i\right)^{-b}, \tag{4.4}$$

where $a \geq 0$ and $b \geq 0$. Then, it can be shown that if $0 < a \leq 2(1-b)$ and $0 \leq b \leq 1$, then the generalized Bayes estimator $\boldsymbol{\delta}_\pi^{GB*}$ against the prior (4.4) dominates $\boldsymbol{\delta}^{UB}$ relative to the loss (1.1). For the details of the proof, see Tsukuma and Kubokawa (2007).

[2] James-Stein type estimator. Consider the expansion estimator

$$\boldsymbol{\delta}_\phi^{JS*} = \boldsymbol{\delta}^{UB} + \frac{\phi(\|\boldsymbol{s}\|^2)}{\|\boldsymbol{s}\|^2}\boldsymbol{s},$$

which yields that $\|\boldsymbol{g}\|^2 + 4\nabla^t\boldsymbol{g} = (\phi/\|\boldsymbol{s}\|^2)\{\phi - 4(k-2)\} - 8\phi'$. Hence, $\boldsymbol{\delta}_\phi^{JS*}$ dominates $\boldsymbol{\delta}^{UB}$ if $\phi(u)$ is a non-decreasing function such that $0 \leq \phi(u) \leq 4(k-2)$ for $k \geq 3$ and $u > 0$.

[3] Empirical Bayes estimator. Consider the shrinkage estimator

$$\boldsymbol{\delta}_\phi^{EB*} = \boldsymbol{\delta}^{UB} - \frac{\phi(T)}{T}\mathbf{1}_k,$$

which yields that $\|\boldsymbol{g}\|^2 + 4\nabla^t\boldsymbol{g} = k\phi(\phi-4)/T^2 + 4k\phi'/T$. Hence, $\boldsymbol{\delta}_\phi^{EB*}$ dominates $\boldsymbol{\delta}^{UB}$ if $\phi(t)$ is a non-increasing function such that $0 \leq \phi(t) \leq 4$ for $t > 0$. This is characterized as an empirical Bayes estimator as used in Section 2.3.

[4] Further improvement by combining expansion and shrinkage estimators. As described above, $\boldsymbol{\delta}_\phi^{EB*}$ is a shrinkage estimator while $\boldsymbol{\delta}_\phi^{JS*}$ is an expansion one. A new estimator proposed here can be derived by combining the shrinkage and expansion estimators. For example, consider the estimators

$$\boldsymbol{\delta}^{JS*} = \boldsymbol{\delta}^{UB} + \frac{\beta_1}{\|\boldsymbol{s}\|^2}\boldsymbol{s} = \boldsymbol{\delta}^{UB} - 2\nabla \log \|\boldsymbol{s}\|^{-\beta_1/2}, \tag{4.5}$$

$$\boldsymbol{\delta}^{EB*} = \boldsymbol{\delta}^{UB} - \frac{\beta_2}{T}\mathbf{1}_k = \boldsymbol{\delta}^{UB} - 2\nabla \log T^{\beta_2/2}, \tag{4.6}$$

for $\beta_1 = 2(k-2)$ and $\beta_2 = 2$. It is noted that these estimators are better than $\boldsymbol{\delta}^{UB}$. The proposed estimator is given by

$$\boldsymbol{\delta}^{C*} = \boldsymbol{\delta}^{UB} - 2\nabla \log\left(\|\boldsymbol{s}\|^{-\beta_1/2}T^{\beta_2/2}\right) = \boldsymbol{\delta}^{UB} + \frac{\beta_1}{\|\boldsymbol{s}\|^2}\boldsymbol{s} - \frac{\beta_2}{T}\mathbf{1}_k. \tag{4.7}$$

The following arguments guarantee that the combined estimator $\boldsymbol{\delta}^{C*}$ dominates both $\boldsymbol{\delta}^{JS*}$ and $\boldsymbol{\delta}^{EB*}$.

In general, let us define $f_{SH}(\boldsymbol{s})$ and $f_{EX}(\boldsymbol{s})$ as twice differentiable functions of $\boldsymbol{s}$ where $(\partial/\partial s_i)f_{SH}(\boldsymbol{s}) > 0$ and $(\partial/\partial s_i)f_{EX}(\boldsymbol{s}) < 0$ for $i = 1, \ldots, k$. Denote the shrinkage and the expansion estimators by $\boldsymbol{\delta}^{SH*} = \boldsymbol{\delta}^{UB} - 2\nabla \log f_{SH}(\boldsymbol{s})$ and $\boldsymbol{\delta}^{EX*} = \boldsymbol{\delta}^{UB} - 2\nabla \log f_{EX}(\boldsymbol{s})$, respectively. Define the combined estimator $\boldsymbol{\delta}^{C*}$ by

$$\boldsymbol{\delta}^{C*} = \boldsymbol{\delta}^{UB} - 2\nabla \log(f_{SH}(\boldsymbol{s})f_{EX}(\boldsymbol{s})),$$

which is also expressed as

$$\begin{aligned}\boldsymbol{\delta}^{C*} &= \boldsymbol{\delta}^{SH*} - 2\nabla \log f_{EX}(\boldsymbol{s}) \\ &= \boldsymbol{\delta}^{EX*} - 2\nabla \log f_{SH}(\boldsymbol{s}).\end{aligned}$$

Theorem 4.1 *Assume that* $\nabla^t\nabla\sqrt{f_{SH}(\boldsymbol{s})} < 0$ *and* $\nabla^t\nabla\sqrt{f_{EX}(\boldsymbol{s})} < 0$*. Then* $\boldsymbol{\delta}^{C*}$ *dominates both the shrinkage and the expansion estimators,* $\boldsymbol{\delta}^{SH*}$ *and* $\boldsymbol{\delta}^{EX*}$*, relative to the loss* (1.1)*.*

Proof. From (4.1), the difference in risk of $\boldsymbol{\delta}^{C*}$ and $\boldsymbol{\delta}^{SH*}$ can be represented as

$$\begin{aligned}
R(\boldsymbol{\delta}^{C*}&,\boldsymbol{\sigma}^{-2}) - R(\boldsymbol{\delta}^{SH*},\boldsymbol{\sigma}^{-2})\\
&= R(\boldsymbol{\delta}^{C*},\boldsymbol{\sigma}^{-2}) - R(\boldsymbol{\delta}^{UB},\boldsymbol{\sigma}^{-2}) - (R(\boldsymbol{\delta}^{SH*},\boldsymbol{\sigma}^{-2}) - R(\boldsymbol{\delta}^{UB},\boldsymbol{\sigma}^{-2}))\\
&= 16E\Big[\frac{\nabla^t\nabla\sqrt{f_{SH}(\boldsymbol{s})f_{EX}(\boldsymbol{s})}}{\sqrt{f_{SH}(\boldsymbol{s})f_{EX}(\boldsymbol{s})}} - \frac{\nabla^t\nabla\sqrt{f_{SH}(\boldsymbol{s})}}{\sqrt{f_{SH}(\boldsymbol{s})}}\Big]\\
&= 16E\Big[\frac{\nabla^t\nabla\sqrt{f_{EX}(\boldsymbol{s})}}{\sqrt{f_{EX}(\boldsymbol{s})}} + 2\frac{[\nabla\sqrt{f_{SH}(\boldsymbol{s})}]^t[\nabla\sqrt{f_{EX}(\boldsymbol{s})}]}{\sqrt{f_{SH}(\boldsymbol{s})f_{EX}(\boldsymbol{s})}}\Big],
\end{aligned}$$

which is negative since $\nabla^t\nabla\sqrt{f_{EX}(\boldsymbol{s})} < 0$ and

$$[\nabla\sqrt{f_{SH}(\boldsymbol{s})}]^t[\nabla\sqrt{f_{EX}(\boldsymbol{s})}] = \frac{1}{4\sqrt{f_{SH}(\boldsymbol{s})f_{EX}(\boldsymbol{s})}}\sum_{i=1}^{k}\Big[\frac{\partial f_{SH}(\boldsymbol{s})}{\partial s_i}\Big]\Big[\frac{\partial f_{EX}(\boldsymbol{s})}{\partial s_i}\Big] < 0.$$

Hence $\boldsymbol{\delta}^{C*}$ dominates $\boldsymbol{\delta}^{SH*}$ relative to the loss (1.1). Similarly, it can be shown that $R(\boldsymbol{\delta}^{C*},\boldsymbol{\sigma}^{-2}) - R(\boldsymbol{\delta}^{EX*},\boldsymbol{\sigma}^{-2}) < 0$. ■

It is noted that the above theorem can be applied to $\boldsymbol{\delta}^{JS*}$ and $\boldsymbol{\delta}^{GB*}$ to obtain their improved estimator

$$\boldsymbol{\delta}^{CB*} = \boldsymbol{\delta}^{GB*} + \{2(k-2)/\|\boldsymbol{s}\|^2\}\boldsymbol{s}, \tag{4.8}$$

where $\boldsymbol{\delta}^{GB*}$ is given by (4.3) against the prior (4.4) with $0 < a \le 2(1-b)$ and $0 \le b \le 1$.

5 Geometric interpretation

We first consider geometric representation of $\boldsymbol{\delta}^{UB}$, $\boldsymbol{\delta}^{BU}$, $\boldsymbol{\delta}^{ML}$ and $\boldsymbol{\sigma}^{-2}$. Let $n = n_1 = \cdots = n_k$ and let $\boldsymbol{a} = \boldsymbol{a}(\boldsymbol{s})$ and $\boldsymbol{b} = \boldsymbol{b}(\boldsymbol{s})$ be vector-valued functions of $\boldsymbol{s}$, respectively. Define inner product of $\boldsymbol{a}$ and $\boldsymbol{b}$ as

$$\langle \boldsymbol{a}, \boldsymbol{b}\rangle = E[\boldsymbol{a}^t\boldsymbol{b}] = \int \boldsymbol{a}^t\boldsymbol{b}\, p(\boldsymbol{s})\mathrm{d}\boldsymbol{s},$$

where $p(\boldsymbol{s})$ is the density function of $\boldsymbol{s}$. Denote $\|\boldsymbol{a}\|_I = \sqrt{\langle \boldsymbol{a}, \boldsymbol{a}\rangle}$. It is noted that $\langle \boldsymbol{\delta}^{BU} - \boldsymbol{\sigma}^{-2}, \boldsymbol{\delta}^{UB}\rangle = 0$, namely, $\boldsymbol{\delta}^{BU} - \boldsymbol{\sigma}^{-2}$ and $\boldsymbol{\delta}^{UB}$ intersect orthogonally. It is also seen that $\langle \boldsymbol{\delta}^{UB} - \boldsymbol{\sigma}^{-2}, \boldsymbol{\sigma}^{-2}\rangle = 0$ and $\langle \boldsymbol{\delta}^{ML} - \boldsymbol{\sigma}^{-2}, \boldsymbol{\sigma}^{-2}\rangle > 0$. Then $\boldsymbol{\delta}^{UB}$, $\boldsymbol{\delta}^{BU}$, $\boldsymbol{\delta}^{ML}$ and $\boldsymbol{\sigma}^{-2}$ are plotted as in Figure 1.

Let $\mathcal{S}(\boldsymbol{z}, r)$ be the closed sphere centered at the vector $\boldsymbol{z}$ with radius r. Denote by $\mathcal{S}^c(\boldsymbol{z}, r)$ complement of $\mathcal{S}(\boldsymbol{z}, r)$. Also let $\boldsymbol{O}$ be the zero vector. It is then reasonable that we seek an alternative estimator in the set

$$\mathcal{S}_{UB} = \mathcal{S}(\boldsymbol{O}, \|\boldsymbol{\delta}^{UB}\|_I) \cap \mathcal{S}(\boldsymbol{\sigma}^{-2}, \|\boldsymbol{\delta}^{UB} - \boldsymbol{\sigma}^{-2}\|_I)$$

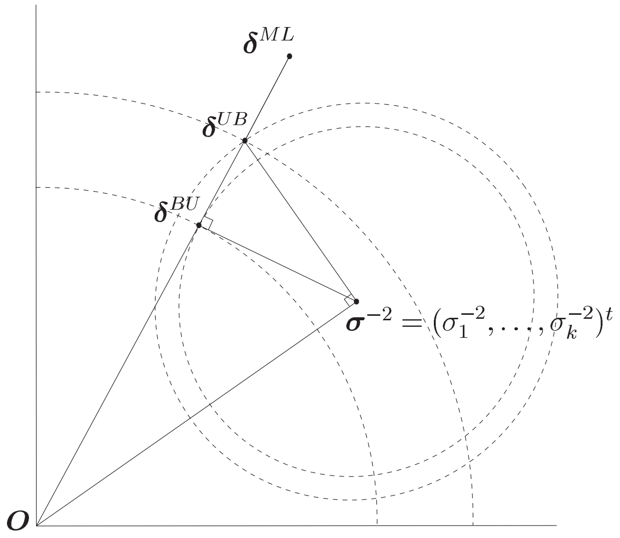


Figure 1: *Geometric interpretation of* $\boldsymbol{\delta}^{ML}$, $\boldsymbol{\delta}^{UB}$ *and* $\boldsymbol{\delta}^{BU}$ *in simultaneous estimation of normal precisions.*

or the set

$$\mathcal{S}_{BU} = \mathcal{S}^c(\boldsymbol{O}, \|\boldsymbol{\delta}^{BU}\|_I) \cap \mathcal{S}(\boldsymbol{\sigma}^{-2}, \|\boldsymbol{\delta}^{BU} - \boldsymbol{\sigma}^{-2}\|_I)$$

since both the sets include $\boldsymbol{\sigma}^{-2}$. Let $\boldsymbol{\delta}^A$ and $\boldsymbol{\delta}^B$ be the shrinkage estimators belonging to the set $\mathcal{S}_{UB}$. Since the set $\mathcal{S}_{UB}$ is convex, an estimator $\alpha\boldsymbol{\delta}^A + (1-\alpha)\boldsymbol{\delta}^B$ for $0 < \alpha < 1$ is also in the set, namely, $\alpha\boldsymbol{\delta}^A + (1-\alpha)\boldsymbol{\delta}^B$ is a shrinkage estimator with respect to $\boldsymbol{\delta}^{UB}$ and dominates $\boldsymbol{\delta}^{UB}$ relative to the loss (1.1). On the other hand, the set $\mathcal{S}_{BU}$ is not convex. Hence for the expansion estimators $\boldsymbol{\delta}^C$ and $\boldsymbol{\delta}^D$ belonging to $\mathcal{S}_{BU}$, $\alpha\boldsymbol{\delta}^C + (1-\alpha)\boldsymbol{\delta}^D$ is not always an expansion estimator with respect to $\boldsymbol{\delta}^{BU}$ and however the estimator dominates $\boldsymbol{\delta}^{BU}$ relative to the loss (1.1).

We next investigate the geometric relation among $\boldsymbol{\delta}^{UB}$, $\boldsymbol{\delta}^{JS*}$, $\boldsymbol{\delta}^{EB*}$, $\boldsymbol{\delta}^{C*}$ and $\boldsymbol{\sigma}^{-2}$, where $\boldsymbol{\delta}^{JS*}$, $\boldsymbol{\delta}^{EB*}$ and $\boldsymbol{\delta}^{C*}$ are given by (4.5), (4.6) and (4.7), respectively. Note that $\boldsymbol{\delta}^{JS*} \in \mathcal{S}_{UB}^c$ and $\boldsymbol{\delta}^{EB*} \in \mathcal{S}_{UB}$, namely $\boldsymbol{\delta}^{JS*}$ is an expansion estimator and $\boldsymbol{\delta}^{EB*}$ is a shrinkage estimator. It can be shown that

$$\langle \boldsymbol{\delta}^{JS*} - \boldsymbol{\delta}^{UB}, \boldsymbol{\delta}^{JS*} - \boldsymbol{\sigma}^{-2} \rangle = E\left[\frac{\beta_1\{\beta_1 - 2(k-2)\}}{\|\boldsymbol{s}\|^2}\right] = 0,$$

for $\beta_1 = 2(k-2)$, and that

$$\langle \boldsymbol{\delta}^{EB*} - \boldsymbol{\delta}^{UB}, \boldsymbol{\delta}^{EB*} - \boldsymbol{\sigma}^{-2} \rangle = E\left[\frac{k\beta_2(\beta_2 - 2)}{T^2}\right] = 0,$$

for $\beta_2 = 2$. Hence $\boldsymbol{\delta}^{JS*}$ and $\boldsymbol{\delta}^{EB*}$ lie on the surface of $\mathcal{S}((\boldsymbol{\delta}^{UB} + \boldsymbol{\sigma}^{-2})/2, \|\boldsymbol{\delta}^{UB} - \boldsymbol{\sigma}^{-2}\|_I/2)$. We also note that $\boldsymbol{\delta}^{BU}$ is on the surface and that $\boldsymbol{\delta}^{C*}$ belongs to the inner of the sphere. Since $\delta_i^{BU} = (n_i - 4)/s_i < (n_i - 2)/s_i - 2/\sum_{i=1}^k s_i = \delta_i^{EB*}$ with probability one, we can see that $\|\boldsymbol{\delta}^{BU}\|_I < \|\boldsymbol{\delta}^{EB*}\|_I$. These facts are described as in Figure 2.

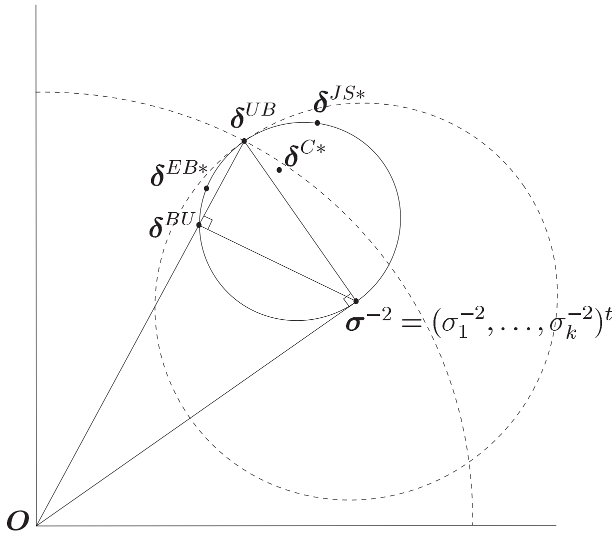


Figure 2: *Geometric relation among $\boldsymbol{\delta}^{UB}$, $\boldsymbol{\delta}^{BU}$, $\boldsymbol{\delta}^{JS*}$, $\boldsymbol{\delta}^{EB*}$, $\boldsymbol{\delta}^{C*}$ and $\boldsymbol{\sigma}^{-2}$.* The estimators $\boldsymbol{\delta}^{UB}$, $\boldsymbol{\delta}^{BU}$, $\boldsymbol{\delta}^{JS*}$ and $\boldsymbol{\delta}^{EB*}$ exist on the surface of the sphere centered at $(\boldsymbol{\delta}^{UB} + \boldsymbol{\sigma}^{-2})/2$ with radius $\|\boldsymbol{\delta}^{UB} - \boldsymbol{\sigma}^{-2}\|_I/2$, but $\boldsymbol{\delta}^{C*}$ lies in the inner of the sphere.

Similarly, we consider the geometric relation between $\boldsymbol{\delta}^{BU}$, $\boldsymbol{\delta}^{JS}$, $\boldsymbol{\delta}^{EB}$ and $\boldsymbol{\sigma}^{-2}$, where $\boldsymbol{\delta}^{JS}$ and $\boldsymbol{\delta}^{EB}$ are given by (2.3) and (2.6), respectively. It can be shown that

$$\langle \boldsymbol{\delta}^{JS} - \boldsymbol{\delta}^{BU}, \boldsymbol{\delta}^{JS} - \boldsymbol{\sigma}^{-2} \rangle = E\left[\frac{\alpha_1\{\alpha_1 - 4(k-1)\}}{\|\boldsymbol{s}\|^2}\right] = 0,$$

for $\alpha_1 = 4(k-1)$, and that

$$\begin{aligned}\langle \boldsymbol{\delta}^{EB} - \boldsymbol{\delta}^{BU}, \boldsymbol{\delta}^{EB} - \boldsymbol{\sigma}^{-2} \rangle =& \alpha_2 E\left[-\frac{2}{T}\sum_{i=1}^{k}\frac{1}{s_i} + \frac{2k}{T^2} + \frac{k\alpha_2}{T^2}\right] \\ \leq& E\left[\frac{k\alpha_2\{\alpha_2 - 2(k-1)\}}{T^2}\right] = 0,\end{aligned}$$

for $\alpha_2 = 2(k-1)$. Thus $\boldsymbol{\delta}^{JS}$ exists on the surface of $\mathcal{S}((\boldsymbol{\delta}^{BU} + \boldsymbol{\sigma}^{-2})/2, \|\boldsymbol{\delta}^{BU} - \boldsymbol{\sigma}^{-2}\|_I/2)$ and $\boldsymbol{\delta}^{EB}$ is located in the inner of the sphere. Noting that both $\boldsymbol{\delta}^{JS}$ and $\boldsymbol{\delta}^{EB}$ are expansion estimators, namely $\boldsymbol{\delta}^{JS} \in \mathcal{S}_{BU}$ and $\boldsymbol{\delta}^{EB} \in \mathcal{S}_{BU}$, we can see that $\boldsymbol{\delta}^{BU}$, $\boldsymbol{\delta}^{JS}$, $\boldsymbol{\delta}^{EB}$ and $\boldsymbol{\sigma}^{-2}$ are plotted as in Figure 3.

6 Numerical study

In this section we shall conduct risk comparison for the improved estimators of the unbiased and the best usual estimators in the case that the number of populations is three, namely, $k = 3$.

For the risk comparison, we first examine the following estimators:

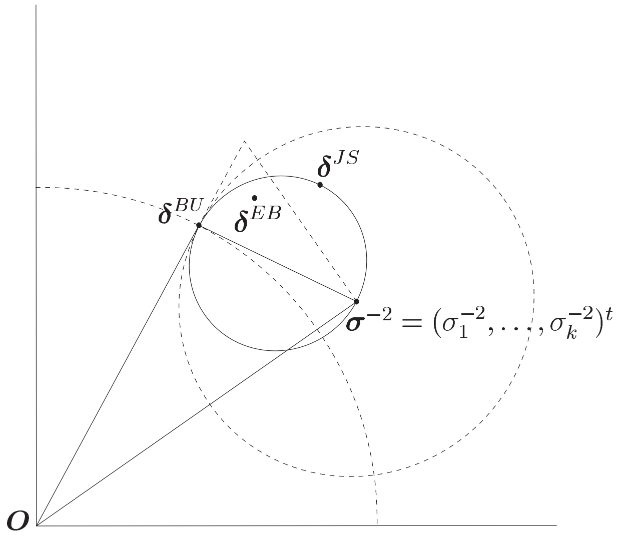


Figure 3: *Geometric relation among* $\boldsymbol{\delta}^{BU}$, $\boldsymbol{\delta}^{JS}$, $\boldsymbol{\delta}^{EB}$ *and* $\boldsymbol{\sigma}^{-2}$. $\boldsymbol{\delta}^{JS}$ is on the surface of the sphere centered at $(\boldsymbol{\delta}^{BU} + \boldsymbol{\sigma}^{-2})/2$ with radius $\|\boldsymbol{\delta}^{BU} - \boldsymbol{\sigma}^{-2}\|_I/2$, but $\boldsymbol{\delta}^{EB}$ lies in the inner of the sphere.

(1) $\boldsymbol{\delta}^{UB}$, given by (1.2),

(2) $\boldsymbol{\delta}^{BU}$, given by (1.3),

(3) $\boldsymbol{\delta}^{JS}$, given by (2.3),

(4) $\boldsymbol{\delta}^{EB}$, given by (2.6),

(5) $\boldsymbol{\delta}^{GB}$, given by (3.12) with $a = 2$,

(6) $\boldsymbol{\delta}^{JS*}$, given by (4.5),

(7) $\boldsymbol{\delta}^{EB*}$, given by (4.6),

(8) $\boldsymbol{\delta}^{C*}$, given by (4.7),

(9) $\boldsymbol{\delta}^{GB*}$, given by (4.3) against the prior (4.4) with $a = 1$ and $b = 1/2$.

(10) $\boldsymbol{\delta}^{CB*}$, given by (4.8) with $a = 1$ and $b = 1/2$.

It is noted that $\boldsymbol{\delta}^{JS}$, $\boldsymbol{\delta}^{EB}$ and $\boldsymbol{\delta}^{GB}$ dominate the best usual estimator $\boldsymbol{\delta}^{BU}$ relative to the loss (1.1). Also, note that $\boldsymbol{\delta}^{JS*}$, $\boldsymbol{\delta}^{EB*}$, $\boldsymbol{\delta}^{C*}$ and $\boldsymbol{\delta}^{GB*}$ dominate the unbiased estimator $\boldsymbol{\delta}^{UB}$ and that $\boldsymbol{\delta}^{C*}$ and $\boldsymbol{\delta}^{CB*}$ are combined estimators of the shrinkage and the expansion estimators and are superior, respectively, to both $\boldsymbol{\delta}^{EB*}$ and $\boldsymbol{\delta}^{JS*}$ and to both $\boldsymbol{\delta}^{GB*}$ and $\boldsymbol{\delta}^{JS*}$.

The estimates of risk values were computed by 10,000 independent replications. We set $(n_1, n_2, n_3) = (30, 30, 30)$ and $(5, 5, 5)$. For values of precision parameters $\boldsymbol{\sigma}^{-2}$, we chose $\boldsymbol{\sigma}^{-2} = (1, 1 + c, (1 + c)^{-1})^t$ for $c = 0, 1, 2, 3, 4$. This simulation results are given

Table 1: Risk values of various estimators for normal precisions with $\boldsymbol{\sigma}^{-2} = (1, 1+c, (1+c)^{-1})^t$ for $c = 0, 1, 2, 3, 4$.

	$(n_1, n_2, n_3) = (30, 30, 30)$					$(n_1, n_2, n_3) = (5, 5, 5)$				
c	0	1	2	3	4	0	1	2	3	4
$\boldsymbol{\delta}^{UB}$	0.23	0.41	0.78	1.33	2.02	5.96	9.09	16.64	27.51	41.58
$\boldsymbol{\delta}^{JS*}$	0.23	0.41	0.78	1.32	2.02	5.88	9.03	16.60	27.48	41.56
$\boldsymbol{\delta}^{EB*}$	0.23	0.41	0.78	1.32	2.02	5.87	9.02	16.59	27.48	41.55
$\boldsymbol{\delta}^{C*}$	0.23	0.40	0.78	1.32	2.02	5.66	8.86	16.48	27.40	41.49
$\boldsymbol{\delta}^{GB*}$	0.21	0.37	0.72	1.22	1.86	1.91	3.53	7.07	12.22	18.94
$\boldsymbol{\delta}^{CB*}$	0.20	0.36	0.71	1.21	1.86	1.38	3.14	6.80	12.03	18.80
$\boldsymbol{\delta}^{BU}$	0.22	0.38	0.72	1.22	1.87	1.99	3.34	6.34	10.63	16.19
$\boldsymbol{\delta}^{JS}$	0.19	0.36	0.72	1.22	1.86	0.87	2.50	5.77	10.23	15.89
$\boldsymbol{\delta}^{EB}$	0.21	0.37	0.71	1.21	1.85	1.21	2.43	5.29	9.47	14.94
$\boldsymbol{\delta}^{GB}$	0.21	0.36	0.70	1.20	1.84	1.00	2.07	4.74	8.79	14.15

in Table 1. From this result, it is observed that each risk value of $\boldsymbol{\delta}^{JS*}$, $\boldsymbol{\delta}^{EB*}$ and $\boldsymbol{\delta}^{C*}$ has a bit of reduction in risk than $\boldsymbol{\delta}^{UB}$ and, on the other hand, $\boldsymbol{\delta}^{JS}$, $\boldsymbol{\delta}^{EB}$ and $\boldsymbol{\delta}^{GB}$ have substantial risk reduction than $\boldsymbol{\delta}^{BU}$. The risk values of $\boldsymbol{\delta}^{CB*}$ is better than $\boldsymbol{\delta}^{BU}$ for $(n_1, n_2, n_3) = (30, 30, 30)$, but not for minimal sample size $(n_1, n_2, n_3) = (5, 5, 5)$ with scattered precisions ($c = 3, 4$).

We next confine our attention to the risk behavior of estimators $\boldsymbol{\delta}^{BU}$, $\boldsymbol{\delta}^{JS}$, $\boldsymbol{\delta}^{EB}$ and $\boldsymbol{\delta}^{GB}$. For these comparison, we shall describe the curves of the relative risks for each improved estimator $\boldsymbol{\delta}^a$ against $\boldsymbol{\delta}^{BU}$, that is, RR $= R(\boldsymbol{\delta}^a, \boldsymbol{\sigma}^{-2})/R(\boldsymbol{\delta}^{BU}, \boldsymbol{\sigma}^{-2})$. The curves are based on 10,000 independent replications. We took five sets of sample size $(n_1, n_2, n_3) = (30, 30, 30)$, $(5, 5, 5)$, $(30, 10, 50)$, $(30, 50, 10)$ and $(50, 30, 10)$. For values of precision $\boldsymbol{\sigma}^{-2}$, we assumed that $\boldsymbol{\sigma}^{-2} = (1, 1+c, (1+c)^{-1})^t$ for $0 \le c \le 4$. Note that the RR is a function of c and that an estimator $\boldsymbol{\delta}$ is better than $\boldsymbol{\delta}^{UB}$ if RR < 1. The results are given in Figure 4.

Figure 4 indicates the following important observations.

1. In the case that $(n_1, n_2, n_3) = (30, 30, 30)$ and $(5, 5, 5)$ and $c = 0$, namely, the n_i's are the same and each element of the precision parameter $\boldsymbol{\sigma}^{-2}$ is one, $\boldsymbol{\delta}^{JS}$ is better than both $\boldsymbol{\delta}^{EB}$ and $\boldsymbol{\delta}^{GB}$. On the other hand $\boldsymbol{\delta}^{GB}$ reduces risk than $\boldsymbol{\delta}^{JS}$ and $\boldsymbol{\delta}^{EB}$ if c is near zero for the cases $(n_1, n_2, n_3) = (30, 10, 50)$, $(30, 50, 10)$ and $(50, 30, 10)$.

2. For minimal sample size $(n_1, n_2, n_3) = (5, 5, 5)$ with $c = 0$, the RR of $\boldsymbol{\delta}^{JS}$ is about 0.45. In other words, the risk of $\boldsymbol{\delta}^{JS}$ is less than half of the risk of $\boldsymbol{\delta}^{BU}$.

3. When the elements of $\boldsymbol{\sigma}^{-2}$ are much different, $\boldsymbol{\delta}^{GB}$ is excellent except $(n_1, n_2, n_3) = (30, 50, 10)$.

4. The risk value of $\boldsymbol{\delta}^{EB}$ is inferior to that of either $\boldsymbol{\delta}^{JS}$ or $\boldsymbol{\delta}^{GB}$.

5. The case that $(n_1, n_2, n_3) = (30, 10, 50)$ and c is large implies that the data is sufficiently obtained from a population with poor precision and not from one with good precision. In this case, $\boldsymbol{\delta}^{GB}$ is the best. Conversely, in the case that $(n_1, n_2, n_3) = (30, 50, 10)$ and c is large, $\boldsymbol{\delta}^{JS}$ is superior.

References

[1] Berger, J. (1980). Improving on inadmissible estimators in continuous exponential families with applications to simultaneous estimation of gamma scale parameters, *Ann. Statist.*, **8**, 545–571.

[2] Bilodeau, M. (1988). On the simultaneous estimation of scale parameters, *Canad. J. Statist.*, **16**, 169-174.

[3] DasGupta, A. (1986). Simultaneous estimation in the multiparameter gamma distribution under weighted quadratic losses, *Ann. Statist.*, **14**, 206-219.

[4] DasGupta, A. (1989). A general theorem on decision theory for nonnegative functionals: with applications, *Ann. Statist.*, **17**, 1360-1374.

[5] Ghosh, M. and Parsian, A. (1980). Admissible and minimax multiparameter estimation in exponential families, *J. Multivariate Anal.*, **10**, 551-564.

[6] Haff, L.R. and Johnson, R.W. (1986). The superharmonic condition for simultaneous estimation of means in exponential families, *Canad. J. Statist.*, **14**, 43–54.

[7] Hudson, H.M. (1978). A natural identity for exponential families with applications in multiparameter estimation. *Ann. Statist.*, **6**, 478-484.

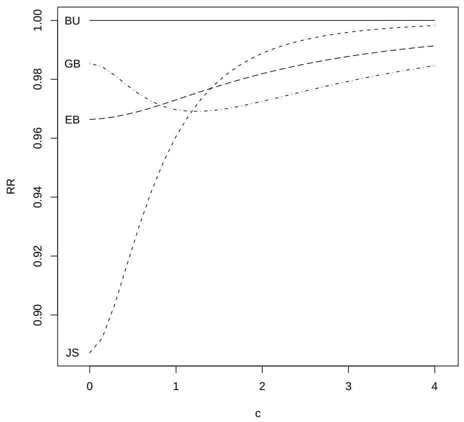
[8] James, W. and Stein, C. (1961). Estimation with quadratic loss. In *Proc. Fourth Berkeley Symp. Math. Statist. Probab.*, **1**, 361–379. University of California Press, Berkeley.

[9] Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proc. Third Berkeley Symp. Math. Statist. Probab.*, **1**, 197-206. University of California University, Berkeley.

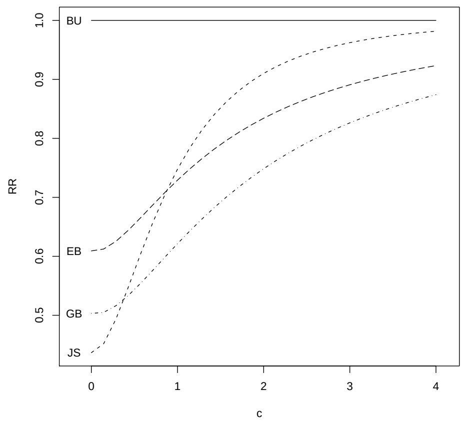
[10] Stein, C. (1973). Estimation of the mean of a multivariate normal distribution. In *Proc. Prague Symp. Asymptotic Statist.*, 345-381.

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

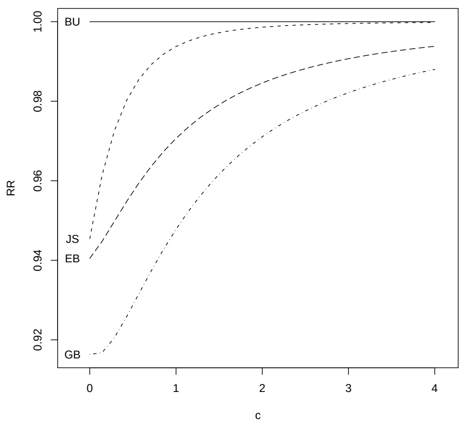
[12] Tsukuma, H. and Kubokawa, T. (2007). The Stein phenomenon in simultaneous estimation of normal precisions. Discussion Paper Series, CIRJE-F-473. Faculty of Economics, University of Tokyo.



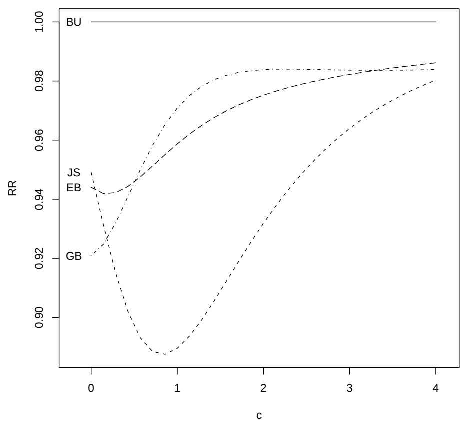
$(n_1, n_2, n_3) = (30, 30, 30)$



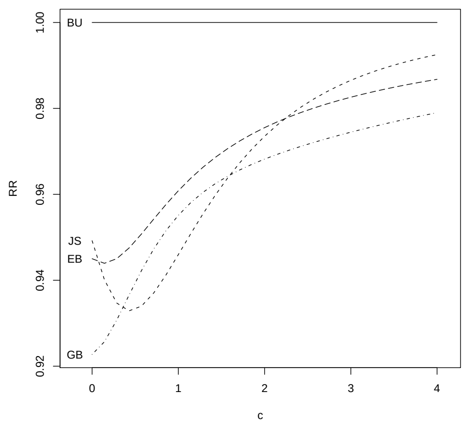
$(n_1, n_2, n_3) = (5, 5, 5)$



$(n_1, n_2, n_3) = (30, 10, 50)$



$(n_1, n_2, n_3) = (30, 50, 10)$



$(n_1, n_2, n_3) = (50, 30, 10)$

Figure 4: The relative risks of improved estimators with respect to the best usual estimator for $\boldsymbol{\sigma}^{-2} = (1, 1+c, (1+c)^{-1})^t$, $0 \le c \le 4$.