Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 99 (2008) 141 – 164
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Stein's phenomenon in estimation of means restricted to a polyhedral convex cone

Hisayuki Tsukuma[a,*], Tatsuya Kubokawa[b]

[a] *Faculty of Medicine, Toho University, 5-21-16 Omori-Nishi, Ota-ku, Tokyo 143-8540, Japan*
[b] *Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

Received 20 February 2006
Available online 21 November 2006

Abstract

This paper treats the problem of estimating the restricted means of normal distributions with a known variance, where the means are restricted to a polyhedral convex cone which includes various restrictions such as positive orthant, simple order, tree order and umbrella order restrictions. In the context of the simultaneous estimation of the restricted means, it is of great interest to investigate decision-theoretic properties of the generalized Bayes estimator against the uniform prior distribution over the polyhedral convex cone. In this paper, the generalized Bayes estimator is shown to be minimax. It is also proved that it is admissible in the one- or two-dimensional case, but is improved on by a shrinkage estimator in the three- or more-dimensional case. This means that the so-called Stein phenomenon on the minimax generalized Bayes estimator can be extended to the case where the means are restricted to the polyhedral convex cone. The risk behaviors of the estimators are investigated through Monte Carlo simulation, and it is revealed that the shrinkage estimator has a substantial risk reduction.

AMS 2000 subject classification: primary 62C1062F30; secondary 62C1562C20

Keywords: Admissibility; Decision theory; Generalized Bayes estimator; Inadmissibility; James–Stein estimator; Minimaxity; Polyhedral convex cone; Restricted parameters; Shrinkage estimation; Simple order restriction; Simultaneous estimation; Tree order restriction; Umbrella order restriction

1. Introduction

Statistical inference under restriction of a parameter space has been extensively studied from practical and theoretical points of view. Most results have been devoted to testing issues against

* Corresponding author.
E-mail addresses: tsukuma@med.toho-u.ac.jp (H. Tsukuma), tatsuya@e.u-tokyo.ac.jp (T. Kubokawa).

0047-259X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2006.10.002

ordered alternatives (See Barlow et al. [1]). Compared with these developments, few papers have presented a theoretical investigation of point estimation under the restriction of means such as the simple order and tree order restrictions. Katz [13] and Farrell [7] studied admissibility and minimaxity in estimation of a restricted parameter in the one-dimensional case. In multi-dimensional cases, Hwang and Peddada [10] derived a graphical condition which ensures that the unrestricted estimators of means are dominated by the restricted maximum likelihood (REML) or isotonic regression estimators. Chang [5,6] and Sengupta and Sen [17] showed that the REML estimator can be improved on by a shrinkage estimator in the context of the simultaneous estimation of vector of the means for the dimension being more than or equal to three. These were extensions of the results of James and Stein [11] and Stein [18] to the case of the restricted means. Although the REML estimator is practically useful, it has a theoretical drawback of the inadmissibility, since truncated estimators such as the REML are, in general, inadmissible. An alternative approach is to employ a generalized Bayes procedure, and the aim of our study is to investigate and clarify the properties of the generalized Bayes estimators in a decision-theoretic framework.

To specify the problem considered here, let $\boldsymbol{X} = (X_1, \dots, X_p)^t$ be a random vector having p-variate normal distribution $\mathcal{N}_p(\boldsymbol{\theta}, \boldsymbol{I}_p)$ for unknown $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)^t$. It is supposed that:

(P) the means θ_i's or the mean vector $\boldsymbol{\theta}$ are restricted to a polyhedral convex cone of the form

$$P = \{\boldsymbol{\theta} | \boldsymbol{r}_i^t \boldsymbol{\theta} \leqslant \alpha_i,\ i = 1, \dots, q\} = \{\boldsymbol{\theta} | \boldsymbol{R}\boldsymbol{\theta} \leqslant \boldsymbol{\alpha}\}, \tag{1.1}$$

where $\boldsymbol{R} = (\boldsymbol{r}_1, \dots, \boldsymbol{r}_q)^t$ is a $q \times p$ known matrix with full low rank q ($\leqslant p$) and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_q)^t$ is known.

The general assumption (P) on the means is practically important in statistical inferences, because the polyhedral convex cone (1.1) includes several restrictions such as

(A) $A = \{\boldsymbol{\theta} | \theta_i \geqslant 0 \text{ for } i = 1, 2, \dots, p\}$;
(B) $B = \{\boldsymbol{\theta} | \theta_1 \leqslant \theta_2 \leqslant \cdots \leqslant \theta_p\}$;
(C) $C = \{\boldsymbol{\theta} | \theta_1 \leqslant \theta_i \text{ for } i = 2, 3, \dots, p\}$;
(D) $D = \{\boldsymbol{\theta} | \theta_1 \leqslant \cdots \leqslant \theta_k \geqslant \cdots \geqslant \theta_p\}$;
(E) $E = \{\boldsymbol{\theta} | \theta_i + \alpha_i \leqslant \theta_{i+1} \text{ for } i = 1, \dots, p-1, \text{ where the } \alpha_i\text{'s are known and nonnegative}\}$;
(F) $F = \{\boldsymbol{\theta} | \theta_{i+1} - \theta_i \leqslant \theta_{i+2} - \theta_{i+1} \text{ for } i = 1, \dots, p-2\}$.

The restrictions (A), (B), (C) and (D) are called the positive orthant, simple order, tree order and umbrella order restrictions, respectively. In this paper, we want to estimate the restricted mean vector $\boldsymbol{\theta}$ based on $\boldsymbol{X}$ relative to the quadratic loss function

$$L(\boldsymbol{\theta}, \boldsymbol{\delta}(\boldsymbol{X})) = \|\boldsymbol{\delta}(\boldsymbol{X}) - \boldsymbol{\theta}\|^2 = \sum_{i=1}^{p} (\delta_i(\boldsymbol{X}) - \theta_i)^2, \tag{1.2}$$

where $\boldsymbol{\delta} = \boldsymbol{\delta}(\boldsymbol{X}) = (\delta_1(\boldsymbol{X}), \dots, \delta_p(\boldsymbol{X}))^t$ is an estimator of $\boldsymbol{\theta}$. Every estimator is evaluated by the risk function $R(\boldsymbol{\theta}, \boldsymbol{\delta}) = E[L(\boldsymbol{\theta}, \boldsymbol{\delta}(\boldsymbol{X}))]$.

A major method for estimating the restricted means is the REML estimator, though it has the theoretical defect of the inadmissibility as noted above. An alternative method is a Bayesian procedure. A basic approach to the Bayes estimation of the restricted means is to suppose the uniform prior distribution over the polyhedral convex cone P, and the resulting generalized Bayes estimator against the uniform prior is

$$\boldsymbol{\delta}^{\mathrm{GB}} = \int_{\boldsymbol{\theta}\in P} \boldsymbol{\theta} \exp(-\|\boldsymbol{X}-\boldsymbol{\theta}\|^2/2)\,\mathrm{d}\boldsymbol{\theta} \Big/ \int_{\boldsymbol{\theta}\in P} \exp(-\|\boldsymbol{X}-\boldsymbol{\theta}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}. \tag{1.3}$$

An interesting query from a practical aspect is whether the generalized Bayes estimator is preferable to the REML estimator as an estimator of the restricted means. For this query, Section 5 demonstrates the practical difference between the two estimators through an example of Robertson et al. [16]. While the simple order restriction is supposed in the example, the unbiased estimates do not preserve the order of the corresponding means and some elements of the REML estimates take the same value. On the other hand, the generalized Bayes estimator gives natural estimates since all the estimates are distinct and keep the increasing order. This may suggest that the generalized Bayes estimator $\boldsymbol{\delta}^{\mathrm{GB}}$ is preferable to the REML as an estimator of the restricted means. However, no theoretical properties of $\boldsymbol{\delta}^{\mathrm{GB}}$ have been developed except $p = 1$. When the means are not restricted, the generalized Bayes estimator over the whole space is $\boldsymbol{X}$, and its minimaxity and the Stein phenomenon, namely admissibility for $p = 1, 2$ and inadmissibility for $p \geqslant 3$ have been studied in the literature. Can these theoretical properties be inherited by the generalized Bayes estimator $\boldsymbol{\delta}^{\mathrm{GB}}$ over the restricted parameter space P? The main aim of this paper is to resolve this interesting question and to establish the minimaxity and the Stein phenomenon for $\boldsymbol{\delta}^{\mathrm{GB}}$.

Section 2 handles the minimax issue of the generalized Bayes estimator $\boldsymbol{\delta}^{\mathrm{GB}}$. It is shown that the unrestricted and unbiased estimator $\boldsymbol{\delta}^{\mathrm{UB}} = \boldsymbol{X}$ remains minimax under the restriction (P). The proof will be done by using a method based on modification of Girshick and Savage's [8] argument. This result implies that the REML estimator is minimax, since it is superior to $\boldsymbol{\delta}^{\mathrm{UB}}$. Moreover, from the minimaxity of $\boldsymbol{\delta}^{\mathrm{UB}}$ and the result of Hartigan [9], it follows that the generalized Bayes estimator $\boldsymbol{\delta}^{\mathrm{GB}}$ is minimax. His argument also gives us exact expressions of the risk functions of the generalized Bayes estimator $\boldsymbol{\delta}^{\mathrm{GB}}$ for the restrictions (A)–(C).

The admissibility and inadmissibility of the generalized Bayes estimator $\boldsymbol{\delta}^{\mathrm{GB}}$ are discussed in Sections 3 and 4. When the dimension p is larger than or equal to three, the so-called Stein effect will work in general in the context of the simultaneous estimation of the means. Thus, it may be imaginable that $\boldsymbol{\delta}^{\mathrm{GB}}$ is inadmissible and dominated by a shrinkage estimator given by

$$\boldsymbol{\delta}^{\mathrm{SH}} = \boldsymbol{\alpha}_* + \left\{1 - \frac{p-2}{\|\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{X}) - \boldsymbol{\alpha}_*\|^2}\right\}(\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{X}) - \boldsymbol{\alpha}_*),$$

where $\boldsymbol{\alpha}_* = \boldsymbol{R}^t(\boldsymbol{R}\boldsymbol{R}^t)^{-1}\boldsymbol{\alpha}$. However, the domain region P of the integrals in $\boldsymbol{\delta}^{\mathrm{GB}}$ is very complicated, and it is technically hard to evaluate the risk function of $\boldsymbol{\delta}^{\mathrm{SH}}$. The key to showing the dominance result lies in the case of $q < p$, and it is noticed that $\boldsymbol{R}$ satisfies the condition

$$\boldsymbol{R}\boldsymbol{1}_p = \boldsymbol{0}_q. \tag{1.4}$$

In fact, Eq. (1.4) is fulfilled by the restrictions (B)–(F). As shown in Section 3, the generalized Bayes estimator given in (1.3) can be expressed by

$$\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{R}^t\boldsymbol{q}(\boldsymbol{x}), \tag{1.5}$$

where $\boldsymbol{q}(\boldsymbol{x})$ is a $q \times 1$ vector-valued function of $\boldsymbol{x}$. Combining the expression (1.5) and the condition (1.4) yields the identity $\boldsymbol{1}_p^t\boldsymbol{\delta}^{\mathrm{GB}} = \boldsymbol{1}_p^t\boldsymbol{x}$, namely,

$$\sum_{j=1}^{p} x_j = \sum_{j=1}^{p} \delta_j^{\mathrm{GB}} \tag{1.6}$$

for $q < p$, where δ_j^{GB} denotes the jth element of $\boldsymbol{\delta}^{\mathrm{GB}}$. Using the identity (1.6), we can establish the dominance property of $\boldsymbol{\delta}^{\mathrm{SH}}$ over $\boldsymbol{\delta}^{\mathrm{GB}}$ under the condition (1.4). Based on this idea, in Section 3, we shall prove the stronger result that $\boldsymbol{\delta}^{\mathrm{SH}}$ dominates $\boldsymbol{\delta}^{\mathrm{GB}}$ under the general restriction (P) of the polyhedral convex cone for any full rank matrix $\boldsymbol{R}$ in the case of $q \leqslant p$.

In the one-dimensional case, namely, $p = 1$, for the positive orthant restriction (A), the admissibility of the generalized Bayes estimator is well known [13]. Similarly, in the case of $p = 2$, it is expected that the generalized Bayes estimator $\boldsymbol{\delta}^{\mathrm{GB}}$ is admissible for the quadratic loss. In fact, the admissibility of $\boldsymbol{\delta}^{\mathrm{GB}}$ is established in Section 4 for the restrictions (A) and (B) in the case $p = 2$. However, it seems technically difficult to show the admissibility for the general restriction (P). The results obtained in Sections 3 and 4 mean that the Stein phenomenon of the minimax generalized Bayes estimator, which is known in estimation of the unrestricted means, can be extended to the restricted cases.

As stated above, Section 5 presents the practical difference between estimators through an example of Robertson et al. [16]. Section 5 also gives the results of Monte Carlo simulations to evaluate the risks of the shrinkage and generalized Bayes estimators $\boldsymbol{\delta}^{\mathrm{SH}}$ and $\boldsymbol{\delta}^{\mathrm{GB}}$ for the restrictions (A)–(C). In the simulations we also calculate the risks of the REML estimator and the James–Stein type estimator improving on the REML. The finding in Section 5 is that $\boldsymbol{\delta}^{\mathrm{SH}}$ reduces the risks favorably over the others. In particular, the risk reduction is substantial when each of the true means is near to zero.

In Appendix A we give the proofs of non-essential theorem and lemma.

We conclude this section with stating some remarks. It is noted that the minimaxity of the generalized Bayes estimators given in Section 2 can be extended to the case of a known and positive definite matrix $\boldsymbol{\Sigma}$ where $\boldsymbol{X} \sim \mathcal{N}_p(\boldsymbol{\theta}, \boldsymbol{\Sigma})$. The inadmissibility result given in Section 3 remains true in the case that $\boldsymbol{\Sigma}$ is a known and positive definite matrix.

It is interesting to note that we have a different story if the restricted parameter space is bounded. For instance, in the case that $\boldsymbol{\theta}$ is restricted to the compact set $\{\boldsymbol{\theta} | \|\boldsymbol{\theta}\| \leqslant r\}$ for a positive $r > 0$, the Bayes estimator against the uniform prior over the restricted space is admissible for any dimension, but is not minimax as studied by Casella and Strawderman [4] and Berry [2].

2. Minimaxity

In this section, we shall establish the minimaxity of the generalized Bayes estimator $\boldsymbol{\delta}^{\mathrm{GB}}$ against the uniform prior over the polyhedral convex cone P given by (1.1). To this end, we first show that the unrestricted and unbiased estimator $\boldsymbol{\delta}^{\mathrm{UB}} = \boldsymbol{X}$ remains minimax under the restriction (P).

Theorem 2.1. *The unrestricted estimator $\boldsymbol{\delta}^{\mathrm{UB}} = \boldsymbol{X}$ is minimax under the restriction* (P) *relative to the loss* (1.2).

Proof. The minimaxity is proved based on the classic method of Girshick and Savage [8] who treated an unrestricted case in a univariate location family. Kubokawa [14] recently demonstrated that the method of Girshick and Savage was still useful for showing minimaxity in a univariate restricted case.

We first prove the minimaxity in the case of $q < p$. Let

$$\begin{aligned} P_k &= \{\boldsymbol{\theta} | -k \leqslant \boldsymbol{r}_i^t \boldsymbol{\theta} - \alpha_i \leqslant 0,\ -k/2 \leqslant (\boldsymbol{r}_j^*)^t \boldsymbol{\theta} \leqslant k/2 \text{ for } i = 1, \ldots, q,\ j = 1, \ldots, p-q\} \\ &= \{\boldsymbol{\theta} | -k\boldsymbol{1}_q \leqslant \boldsymbol{R}\boldsymbol{\theta} - \boldsymbol{\alpha} \leqslant \boldsymbol{0}_q,\ -(k/2)\boldsymbol{1}_{p-q} \leqslant \boldsymbol{R}_* \boldsymbol{\theta} \leqslant (k/2)\boldsymbol{1}_{p-q}\}, \end{aligned}$$

where $\boldsymbol{R}_* = (\boldsymbol{r}_1^*, \ldots, \boldsymbol{r}_{p-q}^*)^t$ is a $(p-q) \times p$ matrix such that $\boldsymbol{R}_*\boldsymbol{R}^t = \boldsymbol{0}_{(p-q)\times q}$. It is noted that

$$\int_{\boldsymbol{\theta}\in P_k} \mathrm{d}\boldsymbol{\theta} = k^p |\boldsymbol{R}\boldsymbol{R}^t|^{-1/2} |\boldsymbol{R}_*\boldsymbol{R}_*^t|^{-1/2},$$

so that we consider the sequence of the prior distributions

$$\pi_k(\boldsymbol{\theta}) = \begin{cases} k^{-p}|\boldsymbol{R}\boldsymbol{R}^t|^{1/2}|\boldsymbol{R}_*\boldsymbol{R}_*^t|^{1/2} & \text{if } \boldsymbol{\theta} \in P_k, \\ 0 & \text{otherwise.} \end{cases}$$

The corresponding Bayes estimators are given by

$$\boldsymbol{\delta}_k^\pi = \boldsymbol{\delta}_k^\pi(\boldsymbol{X}) = \int_{\boldsymbol{a}\in P_k} \boldsymbol{a} \exp(-\|\boldsymbol{a}-\boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{a} \Bigg/ \int_{\boldsymbol{a}\in P_k} \exp(-\|\boldsymbol{a}-\boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{a}$$

with the Bayes risk function

$$r(\pi_k, \boldsymbol{\delta}_k^\pi) = \frac{c|\boldsymbol{R}\boldsymbol{R}^t|^{1/2}|\boldsymbol{R}_*\boldsymbol{R}_*^t|^{1/2}}{k^p} \int_{\boldsymbol{\theta}\in P_k} \int \|\boldsymbol{\delta}_k^\pi(\boldsymbol{x}) - \boldsymbol{\theta}\|^2 \exp(-\|\boldsymbol{x}-\boldsymbol{\theta}\|^2/2)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{\theta},$$

where $c = (2\pi)^{-p/2}$. Since $r(\pi_k, \boldsymbol{\delta}_k^\pi) \leqslant r(\pi_k, \boldsymbol{\delta}^{\mathrm{UB}}) = p$, it is sufficient to show that $\liminf_{k\to\infty} r(\pi_k, \boldsymbol{\delta}_k^\pi) \geqslant p$. The transformations $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{\theta}$ and $\boldsymbol{t} = \boldsymbol{a} - \boldsymbol{\theta}$ give that

$$\begin{aligned} \boldsymbol{\delta}_k^\pi(\boldsymbol{x}) - \boldsymbol{\theta} &= \frac{\int_{\boldsymbol{a}\in P_k} (\boldsymbol{a}-\boldsymbol{\theta}) \exp(-\|\boldsymbol{a}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{a}}{\int_{\boldsymbol{a}\in P_k} \exp(-\|\boldsymbol{a}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{a}} \\ &= \frac{\int_{\boldsymbol{t}+\boldsymbol{\theta}\in P_k} \boldsymbol{t} \exp(-\|\boldsymbol{t}-\boldsymbol{z}\|^2/2)\,\mathrm{d}\boldsymbol{t}}{\int_{\boldsymbol{t}+\boldsymbol{\theta}\in P_k} \exp(-\|\boldsymbol{t}-\boldsymbol{z}\|^2/2)\,\mathrm{d}\boldsymbol{t}}. \end{aligned}$$

An important point in the proof of the minimaxity is to consider the transformation from $\boldsymbol{\theta}$ to $\boldsymbol{\xi} = (\boldsymbol{\xi}_{(q)}^t, \boldsymbol{\xi}_{(p-q)}^t)^t$, where

$$\boldsymbol{\xi}_{(q)} = 2(\boldsymbol{R}\boldsymbol{\theta} - \boldsymbol{\alpha} + (k/2)\boldsymbol{1}_q)/k, \quad \boldsymbol{\xi}_{(p-q)} = 2\boldsymbol{R}_*\boldsymbol{\theta}/k.$$

Through the transformation, the region P_k is rewritten as

$$\begin{aligned} P_k^* = \{\boldsymbol{t} | &-(k/2)(\boldsymbol{1}_q + \boldsymbol{\xi}_{(q)}) \leqslant \boldsymbol{R}\boldsymbol{t} \leqslant (k/2)(\boldsymbol{1}_q - \boldsymbol{\xi}_{(q)}) \text{ and} \\ &-(k/2)(\boldsymbol{1}_{p-q} + \boldsymbol{\xi}_{(p-q)}) \leqslant \boldsymbol{R}_*\boldsymbol{t} \leqslant (k/2)(\boldsymbol{1}_{p-q} - \boldsymbol{\xi}_{(p-q)})\} \end{aligned}$$

and the term $\boldsymbol{\delta}_k^\pi(\boldsymbol{x}) - \boldsymbol{\theta}$ is rewritten as

$$\boldsymbol{\delta}_k^\pi(\boldsymbol{x}) - \boldsymbol{\theta} = \frac{\int_{\boldsymbol{t}\in P_k^*} \boldsymbol{t} \exp(-\|\boldsymbol{t}-\boldsymbol{z}\|^2/2)\,\mathrm{d}\boldsymbol{t}}{\int_{\boldsymbol{t}\in P_k^*} \exp(-\|\boldsymbol{t}-\boldsymbol{z}\|^2/2)\,\mathrm{d}\boldsymbol{t}} = \boldsymbol{\delta}_k^*(\boldsymbol{z}|\boldsymbol{\xi}) \quad \text{say.}$$

Since the Jacobian of the transformation from $\boldsymbol{\theta}$ to $\boldsymbol{\xi}$ is

$$k^p|\boldsymbol{R}\boldsymbol{R}^t|^{-1/2}|\boldsymbol{R}_*\boldsymbol{R}_*^t|^{-1/2}/2^p,$$

the Bayes risk of $\boldsymbol{\delta}_k^{\pi}$ is expressed as

$$r(\pi_k, \boldsymbol{\delta}_k^{\pi}) = \frac{c}{2^p} \int_{\boldsymbol{\xi} \in U} \int \|\boldsymbol{\delta}_k^*(\boldsymbol{z}|\boldsymbol{\xi})\|^2 \exp(-\|\boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{z}\, \mathrm{d}\boldsymbol{\xi},$$

where $U = \{\boldsymbol{\xi} | |\xi_i| \leqslant 1 \text{ for } i = 1, \dots, p\}$. For a small $\varepsilon > 0$, let $U_\varepsilon = \{\boldsymbol{\xi} | |\xi_i| < 1 - \varepsilon \text{ for } i = 1, \dots, p\}$. Then, the Bayes risk is evaluated as

$$r(\pi_k, \boldsymbol{\delta}_k^{\pi}) \geqslant \frac{c}{2^p} \int_{\boldsymbol{\xi} \in U_\varepsilon} \int \|\boldsymbol{\delta}_k^*(\boldsymbol{z}|\boldsymbol{\xi})\|^2 \exp(-\|\boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{z}\, \mathrm{d}\boldsymbol{\xi}.$$

From the fact that $|\xi_i| < 1 - \varepsilon$, it is seen that $1 + \xi_i > 0$ and $1 - \xi_i > 0$, which imply that P_k^* tends to $\mathbb{R}^p$ as $k \to \infty$ and then $\boldsymbol{\delta}_k^*(\boldsymbol{z}|\boldsymbol{\xi})$ converges $\boldsymbol{\delta}^{\mathrm{UB}}(\boldsymbol{z}) = \boldsymbol{z}$. Using the Fatou lemma, we obtain that

$$\begin{aligned}
\liminf_{k\to\infty} r(\pi_k, \boldsymbol{\delta}_k^{\pi}) &\geqslant \frac{c}{2^p} \liminf_{k\to\infty} \int_{\boldsymbol{\xi} \in U_\varepsilon} \int \|\boldsymbol{\delta}_k^*(\boldsymbol{z}|\boldsymbol{\xi})\|^2 \exp(-\|\boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{z}\, \mathrm{d}\boldsymbol{\xi} \\
&\geqslant \frac{c}{2^p} \int_{\boldsymbol{\xi} \in U_\varepsilon} \int \left\| \liminf_{k\to\infty} \boldsymbol{\delta}_k^*(\boldsymbol{z}|\boldsymbol{\xi}) \right\|^2 \exp(-\|\boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{z}\, \mathrm{d}\boldsymbol{\xi} \\
&= \frac{c}{2^p} \int_{\boldsymbol{\xi} \in U_\varepsilon} \mathrm{d}\boldsymbol{\xi} \int \|\boldsymbol{\delta}^{\mathrm{UB}}(\boldsymbol{z})\|^2 \exp(-\|\boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{z} \\
&= (1 - \varepsilon)^p p.
\end{aligned}$$

Since ε is arbitrary, it follows that $\liminf_{k\to\infty} r(\pi_k, \boldsymbol{\delta}_k^{\pi}) \geqslant p$, which establishes the minimaxity of $\boldsymbol{\delta}^{\mathrm{UB}} = \boldsymbol{X}$.

The minimaxity result for $q = p$ can be proved along the same line as used above. As a sequence of prior distributions of $\boldsymbol{\theta}$, we consider the form $\pi_k(\boldsymbol{\theta}) = k^{-p}|\boldsymbol{R}\boldsymbol{R}^t|^{1/2}$ if $\boldsymbol{\theta} \in P_k$ and 0 otherwise, where

$$P_k = \{\boldsymbol{\theta} | -k\mathbf{1}_p \leqslant \boldsymbol{R}\boldsymbol{\theta} - \boldsymbol{\alpha} \leqslant \mathbf{0}_p\}.$$

Then, we can employ the same arguments as in the above proof to establish the minimaxity of $\boldsymbol{\delta}^{\mathrm{UB}}$ in the case of $q = p$. □

Although the minimaxity of the unrestricted estimator $\boldsymbol{\delta}^{\mathrm{UB}} = \boldsymbol{X}$ is verified in the restricted case (P), $\boldsymbol{\delta}^{\mathrm{UB}}$ may be useless, because it takes values outside the restricted space. An alternative is to take the REML estimator which projects $\boldsymbol{X}$ onto the restricted space P. Since the REML estimator is known to dominate the unrestricted one $\boldsymbol{X}$ in terms of risk, Theorem 2.1 implies that the REML estimator is minimax under the loss (1.2). In general, however, such truncated estimators like the REML one are known to be inadmissible. Another approach is to consider the generalized Bayes estimator against the uniform prior on the restricted space P, given by $\boldsymbol{\delta}^{\mathrm{GB}} = \boldsymbol{X} + \boldsymbol{g}$, where

$$\boldsymbol{g} = \boldsymbol{g}(\boldsymbol{X}) = \int_{\boldsymbol{\theta} \in P} (\boldsymbol{\theta} - \boldsymbol{X}) \exp(-\|\boldsymbol{\theta} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\theta} \Big/ \int_{\boldsymbol{\theta} \in P} \exp(-\|\boldsymbol{\theta} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\theta}.$$

Hartigan [9] showed that the generalized Bayes estimator against the uniform prior over a closed convex set dominates $\boldsymbol{\delta}^{\mathrm{UB}}$ under the quadratic loss function. This yields the following corollary:

Corollary 2.1. *For the restriction* (P), *the generalized Bayes estimator* $\boldsymbol{\delta}^{\mathrm{GB}}$ *dominates the unrestricted estimator* $\boldsymbol{\delta}^{\mathrm{UB}}$ *and hence* $\boldsymbol{\delta}^{\mathrm{GB}}$ *is minimax.*

The generalized Bayes estimators $\boldsymbol{\delta}^{\mathrm{GB}}$ for the restrictions (A), (B) and (C) are denoted by $\boldsymbol{\delta}_A^{\mathrm{GB}}$, $\boldsymbol{\delta}_B^{\mathrm{GB}}$ and $\boldsymbol{\delta}_C^{\mathrm{GB}}$, respectively. Exact expressions of their risk functions of the generalized Bayes estimators are given in the following theorem.

Theorem 2.2. *For the restrictions* (A), (B) *and* (C), *the risks of the generalized Bayes estimators* $\boldsymbol{\delta}_A^{\mathrm{GB}}$, $\boldsymbol{\delta}_B^{\mathrm{GB}}$ *and* $\boldsymbol{\delta}_C^{\mathrm{GB}}$ *are expressed as follows*:

$$R(\boldsymbol{\theta}, \boldsymbol{\delta}_A^{\mathrm{GB}}) = p - \sum_{i=1}^{p} E\left[\frac{\theta_i \exp(-X_i^2/2)}{\int_0^\infty \exp(-(\xi_i - X_i)^2/2)\,\mathrm{d}\xi_i}\right], \tag{2.1}$$

$$R(\boldsymbol{\theta}, \boldsymbol{\delta}_B^{\mathrm{GB}}) = p + \sum_{i=1}^{p-1} (\theta_i - \theta_{i+1}) E\left[\frac{\int_{\boldsymbol{\xi}_{-i} \in B_{-i}} \psi(\xi_{i+1}|X_i) \prod_{j \neq i} \psi(\xi_j|X_j)\,\mathrm{d}\boldsymbol{\xi}_{-i}}{\int_{\boldsymbol{\xi} \in B} \exp(-\|\boldsymbol{\xi} - \boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}\right], \tag{2.2}$$

$$R(\boldsymbol{\theta}, \boldsymbol{\delta}_C^{\mathrm{GB}}) = p + \sum_{i=2}^{p} (\theta_1 - \theta_i) E\left[\frac{\int_{\boldsymbol{\xi}_{-i} \in C_{-i}} \psi(\xi_1|X_i) \prod_{j \neq i} \psi(\xi_j|X_j)\,\mathrm{d}\boldsymbol{\xi}_{-i}}{\int_{\boldsymbol{\xi} \in C} \exp(-\|\boldsymbol{\xi} - \boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}\right], \tag{2.3}$$

where $\boldsymbol{\xi}_{-i} = (\xi_1, \ldots, \xi_{i-1}, \xi_{i+1}, \ldots, \xi_p)^t$, $B_{-i} = \{\boldsymbol{\xi}_{-i} | -\infty < \xi_1 \leqslant \cdots \leqslant \xi_{i-1} \leqslant \xi_{i+1} \leqslant \cdots \leqslant \xi_p < \infty\}$ *and* $C_{-i} = \{\boldsymbol{\xi}_{-i} | \xi_1 \leqslant \xi_j \text{ for } j \neq 1, i\}$, *and* $\psi(\xi_j|x_i)$ *is defined by*

$$\psi(\xi_j|x_i) = \exp\{-(\xi_j - x_i)^2/2\}. \tag{2.4}$$

We put the proof of Theorem 2.2 in Section 6.1. Theorem 2.2 provides the exact expressions of the risks of the generalized Bayes estimators $\boldsymbol{\delta}_\lambda^{\mathrm{GB}}$. From the expressions, it is seen that $R(\boldsymbol{\theta}, \boldsymbol{\delta}_A^{\mathrm{GB}}) = R(\boldsymbol{\theta}, \boldsymbol{X}) = p$ at $\boldsymbol{\theta} = \boldsymbol{0}$ and that $R(\boldsymbol{\theta}, \boldsymbol{\delta}_B^{\mathrm{GB}}) = R(\boldsymbol{\theta}, \boldsymbol{X})$ and $R(\boldsymbol{\theta}, \boldsymbol{\delta}_C^{\mathrm{GB}}) = R(\boldsymbol{\theta}, \boldsymbol{X})$ at $\theta_1 = \cdots = \theta_p$.

Remark 2.1. It is remarked that the results given by Theorems 2.1 and 2.2 can be extended to the general case that $\boldsymbol{X} \sim \mathcal{N}_p(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ for a positive definite and fully known matrix $\boldsymbol{\Sigma}$. In this general situation, it can be shown that $\boldsymbol{\delta}_\Sigma^{\mathrm{UB}} = \boldsymbol{X}$ is minimax in estimation of $\boldsymbol{\theta}$ restricted to the polyhedral convex cone P under the quadratic loss of the form $\|\boldsymbol{\delta} - \boldsymbol{\theta}\|_\Sigma^2 = (\boldsymbol{\delta} - \boldsymbol{\theta})^t \boldsymbol{\Sigma}^{-1} (\boldsymbol{\delta} - \boldsymbol{\theta})$, where $\boldsymbol{\delta}$ is an estimator of $\boldsymbol{\theta}$. It can be also verified that the generalized Bayes estimator

$$\boldsymbol{\delta}_\Sigma^{\mathrm{GB}} = \int_{\boldsymbol{\theta} \in P} \boldsymbol{\theta} \exp(-\|\boldsymbol{\theta} - \boldsymbol{X}\|_\Sigma^2/2)\,\mathrm{d}\boldsymbol{\theta} \Bigg/ \int_{\boldsymbol{\theta} \in P} \exp(-\|\boldsymbol{\theta} - \boldsymbol{X}\|_\Sigma^2/2)\,\mathrm{d}\boldsymbol{\theta}$$

dominates $\boldsymbol{\delta}_\Sigma^{\mathrm{UB}}$ and is minimax.

3. Inadmissibility

We shall discuss the inadmissibility and admissibility of the generalized Bayes estimator $\boldsymbol{\delta}^{\mathrm{GB}}$ in Sections 3 and 4. When the dimension p is larger than or equal to three, it may be possible that the generalized Bayes estimator against the uniform prior over the restricted space P is inadmissible since the Stein effect can work in the context of the simultaneous estimation of the means. In this section, we shall show that the generalized Bayes estimator $\boldsymbol{\delta}^{\mathrm{GB}} = \boldsymbol{X} + \boldsymbol{g}(\boldsymbol{X})$ is improved on by a shrinkage estimator of the form

$$\boldsymbol{\delta}^{\mathrm{SH}} = \boldsymbol{\delta}^{\mathrm{SH}}(\boldsymbol{X}) = \boldsymbol{\alpha}_* + \left\{ 1 - \frac{a}{\|\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{X}) - \boldsymbol{\alpha}_*\|^2} \right\} (\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{X}) - \boldsymbol{\alpha}_*), \tag{3.1}$$

where $\boldsymbol{\alpha}_* = \boldsymbol{R}^t(\boldsymbol{R}\boldsymbol{R}^t)^{-1}\boldsymbol{\alpha}$ and a is a positive constant.

To establish the dominance result, we first derive interesting and useful properties of the generalized Bayes estimator $\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{g}(\boldsymbol{x})$.

Lemma 3.1. *The generalized Bayes estimator under the restriction* (P) *is expressed by*

$$\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{R}^t \boldsymbol{q}(\boldsymbol{x}), \tag{3.2}$$

where $\boldsymbol{q}(\boldsymbol{x})$ is a $q \times 1$ vector-valued function of $\boldsymbol{x}$.

The proof of Lemma 3.1 is postponed to Section 6.2. Lemma 3.1 implies that $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{R}^t\boldsymbol{q}(\boldsymbol{x})$. Using Lemma 3.1, we immediately obtain the following lemma:

Lemma 3.2. *Assume that $q < p$. If $\boldsymbol{R}\boldsymbol{1}_p = \boldsymbol{0}_q$, then the identity of the form $\boldsymbol{1}_p^t\boldsymbol{x} = \boldsymbol{1}_p^t\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{x})$, namely,*

$$\sum_{j=1}^{p} x_j = \sum_{j=1}^{p} \delta_j^{\mathrm{GB}}(\boldsymbol{x}),$$

holds under the restriction (P).

It is noted that the condition that $\boldsymbol{R}\boldsymbol{1}_p = \boldsymbol{0}_q$ in Lemma 3.2 is satisfied by the order restrictions (B)–(F). The identity given in Lemma 3.2 is the key to showing Lemma 3.4 which is essential for establishing the dominance result. To show Lemma 3.4, we use the following lemma due to Karlin and Rinott [12].

Lemma 3.3. *Let f_1 and f_2 be probability densities on $\mathcal{X} \subseteq R^p$ satisfying for all $\boldsymbol{\theta}, \boldsymbol{\xi} \in \mathcal{X}$*

$$f_1(\boldsymbol{\theta}) f_2(\boldsymbol{\xi}) \leqslant f_2(\boldsymbol{\theta} \vee \boldsymbol{\xi}) f_1(\boldsymbol{\theta} \wedge \boldsymbol{\xi}), \tag{3.3}$$

where

$$\boldsymbol{\theta} \vee \boldsymbol{\xi} = (\max(\theta_1, \xi_1), \ldots, \max(\theta_p, \xi_p))^t,$$
$$\boldsymbol{\theta} \wedge \boldsymbol{\xi} = (\min(\theta_1, \xi_1), \ldots, \min(\theta_p, \xi_p))^t$$

for $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)^t$ and $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)^t$. Then for any increasing ϕ on $\mathcal{X}$,

$$\int_{\boldsymbol{\theta} \in \mathcal{X}} \phi(\boldsymbol{\theta}) f_1(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta} \leqslant \int_{\boldsymbol{\theta} \in \mathcal{X}} \phi(\boldsymbol{\theta}) f_2(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta}.$$

Lemma 3.4. *Let $\boldsymbol{H} = \boldsymbol{H}(\boldsymbol{x})$ be a $p \times p$ matrix with the (i, j)-element being $H_{ij}(\boldsymbol{x}) = \partial g_j(\boldsymbol{x})/\partial x_i$. Then $\boldsymbol{H}$ is negative semi-definite.*

Proof. The proof will be done by considering the three cases: (Case 1) $\boldsymbol{R}$ satisfies $\boldsymbol{R}\boldsymbol{1}_p = \boldsymbol{0}_q$ for $q < p$, (Case 2) $\boldsymbol{R}$ is any full rank matrix for $q < p$ and (Case 3) $\boldsymbol{R}$ is any full rank matrix for

$q = p$. Note that the element H_{ij} is written by

$$\begin{aligned}\frac{\partial g_j(\boldsymbol{x})}{\partial x_i} &= \frac{\partial}{\partial x_i}\left\{\frac{\int_{\boldsymbol{\xi}\in P} \xi_j \exp(-\|\boldsymbol{\xi}\|^2/2+\boldsymbol{x}^t\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi}\in P} \exp(-\|\boldsymbol{\xi}\|^2/2+\boldsymbol{x}^t\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}} - x_j\right\}\\ &= \frac{\int_{\boldsymbol{\xi}\in P} \xi_i\xi_j \exp(-\|\boldsymbol{\xi}\|^2/2+\boldsymbol{x}^t\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi}\in P} \exp(-\|\boldsymbol{\xi}\|^2/2+\boldsymbol{x}^t\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}}\\ &\quad - \frac{\int_{\boldsymbol{\xi}\in P} \xi_i \exp(-\|\boldsymbol{\xi}\|^2/2+\boldsymbol{x}^t\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi}\in P} \exp(-\|\boldsymbol{\xi}\|^2/2+\boldsymbol{x}^t\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}} \times \frac{\int_{\boldsymbol{\xi}\in P} \xi_j \exp(-\|\boldsymbol{\xi}\|^2/2+\boldsymbol{x}^t\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi}\in P} \exp(-\|\boldsymbol{\xi}\|^2/2+\boldsymbol{x}^t\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}} - \delta_{ij},\end{aligned}$$

where δ_{ij} denotes Kronecker's delta. Hence $\boldsymbol{H}$ is symmetric, i.e., $H_{ij}(\boldsymbol{x}) = H_{ji}(\boldsymbol{x})$ for $i < j$.

Case 1: When $\boldsymbol{R}$ satisfies $\boldsymbol{R}\boldsymbol{1}_p = \boldsymbol{0}_q$ for $q < p$, Lemma 3.2 implies that $\boldsymbol{1}_p^t\boldsymbol{x} = \boldsymbol{1}_p^t\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{x})$. From Lemma 3.2, we observe that

$$0 = \frac{\partial}{\partial x_i}\left\{\boldsymbol{1}_p^t(\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{x}) - \boldsymbol{x})\right\} = \frac{\partial}{\partial x_i}\sum_{j=1}^{p} g_j(\boldsymbol{x}) = \sum_{j=1}^{p} H_{ij}(\boldsymbol{x}), \quad i = 1, \ldots, p.$$

Thus, the diagonal elements of $\boldsymbol{H}$ can be rewritten as

$$\begin{aligned}H_{11}(\boldsymbol{x}) &= -\sum_{j=2}^{p} H_{1j}(\boldsymbol{x}),\\ H_{22}(\boldsymbol{x}) &= -\sum_{j\neq 2} H_{2j}(\boldsymbol{x}) = -H_{12}(\boldsymbol{x}) - \sum_{j=3}^{p} H_{2j}(\boldsymbol{x}),\\ &\;\;\vdots\\ H_{pp}(\boldsymbol{x}) &= -\sum_{j\neq p} H_{pj}(\boldsymbol{x}) = -\sum_{i=1}^{p-1} H_{ip}(\boldsymbol{x}),\end{aligned}$$

which implies that

$$\begin{aligned}\boldsymbol{H} &= \begin{pmatrix} -\sum_{j=2}^{p} H_{1j}(\boldsymbol{x}) & H_{12}(\boldsymbol{x}) & \cdots & H_{1p}(\boldsymbol{x})\\ H_{12}(\boldsymbol{x}) & -H_{12}(\boldsymbol{x}) - \sum_{j=3}^{p} H_{2j}(\boldsymbol{x}) & \cdots & H_{2p}(\boldsymbol{x})\\ \vdots & \vdots & \ddots & \vdots\\ H_{1p}(\boldsymbol{x}) & H_{2p}(\boldsymbol{x}) & \cdots & -\sum_{i=1}^{p-1} H_{ip}(\boldsymbol{x})\end{pmatrix}\\ &= -\sum_{i<j} H_{ij}(\boldsymbol{x})\boldsymbol{P}_{ij}, \end{aligned} \tag{3.4}$$

where $\boldsymbol{P}_{ij}$ is a $p \times p$ matrix such that the (i, i)- and (j, j)-elements are ones, the (i, j)- and (j, i)-elements are minus ones and the others are zeros. It is noted that $\boldsymbol{z}^t\boldsymbol{P}_{ij}\boldsymbol{z} = (z_i - z_j)^2$ for any vector $\boldsymbol{z} = (z_1, \ldots, z_p)^t$. Then from (3.4), we observe that

$$\boldsymbol{z}^t\boldsymbol{H}\boldsymbol{z} = -\sum_{i<j} H_{ij}(\boldsymbol{x})(z_i - z_j)^2, \tag{3.5}$$

which implies that $\boldsymbol{H}$ is negative semi-definite if $H_{ij}(\boldsymbol{x}) \geqslant 0$ for $i < j$. Hence it is sufficient to show the inequality $H_{ij}(\boldsymbol{x}) \geqslant 0$ for $i < j$.

Lemma 3.3 is used to verify the inequality $H_{ij}(\boldsymbol{x}) \geqslant 0$. For the restriction (P), let f and f_ε be probability densities on $\mathcal{X} = P$ with

$$f(\xi) = \frac{\exp(-\|\xi\|^2/2 + \boldsymbol{x}^t \xi)}{c(\boldsymbol{x})}, \quad f_\varepsilon(\boldsymbol{\eta}) = \frac{\exp(-\|\xi\|^2/2 + (\boldsymbol{x} + \boldsymbol{\epsilon}_i)^t \xi)}{c(\boldsymbol{x} + \boldsymbol{\epsilon}_i)},$$

where $\boldsymbol{x} + \boldsymbol{\epsilon}_i = (x_1, \ldots, x_{i-1}, x_i + \varepsilon, x_{i+1}, \ldots, x_p)^t$, $\varepsilon > 0$, and

$$c(\boldsymbol{x}) = \int_{\xi \in P} \exp(-\|\xi\|^2/2 + \boldsymbol{x}^t \xi)\, \mathrm{d}\boldsymbol{\xi}.$$

It is easy to check that a pair of f and f_ε satisfies the inequality (3.3). Then from Lemma 3.3, it follows that for any $\varepsilon > 0$,

$$\int_{\xi \in P} \xi_j f(\xi)\, \mathrm{d}\boldsymbol{\xi} \leqslant \int_{\xi \in P} \xi_j f_\varepsilon(\xi)\, \mathrm{d}\boldsymbol{\xi},$$

which means that $\delta_j^{\mathrm{GB}} = \int_{\xi \in P} \xi_j f(\xi)\, \mathrm{d}\boldsymbol{\xi}$ is nondecreasing of x_i, namely,

$$H_{ij}(\boldsymbol{x}) = \frac{\partial g_j(\boldsymbol{x})}{\partial x_i} = \frac{\partial \delta_j^{\mathrm{GB}}}{\partial x_i} \geqslant 0$$

for $i \neq j$. From (3.5), it is seen that $\boldsymbol{H}$ is negative semi-definite, and Lemma 3.4 is proved in Case 1.

Case 2: Let $\boldsymbol{R}$ be any full rank matrix for $q < p$. For any vector $\boldsymbol{\theta} \in R^p$, let us consider the linear mapping $\varphi : \boldsymbol{\theta} \to \boldsymbol{R\theta}$. Since the rank of $\boldsymbol{R}$ is q $(< p)$, the dimension of the kernel of φ is $p - q$. Hence there exists a vector $\boldsymbol{\beta}$ such that $\boldsymbol{R\beta} = \boldsymbol{0}_q$. Let $\boldsymbol{\Gamma}$ be a $p \times p$ orthogonal matrix such that $\boldsymbol{\Gamma\beta} = \boldsymbol{1}_p$ where $\boldsymbol{\beta}$ is a vector in the kernel of φ, namely, $\boldsymbol{\beta}$ belongs to the intersection of hyperplanes $\boldsymbol{r}_i^t\boldsymbol{\theta} = 0,\ i = 1, \ldots, q$, if $q \geqslant 2$ or to the hyperplane $\boldsymbol{R\theta} = 0$ if $q = 1$. Noting that $\boldsymbol{R\beta} = \boldsymbol{0}_q$, we can see that

$$\boldsymbol{0}_q = \boldsymbol{R\beta} = \boldsymbol{R\Gamma}^t\boldsymbol{\Gamma\beta} = \boldsymbol{R\Gamma}^t\boldsymbol{1}_p.$$

It is noted that $\boldsymbol{H}$ is written as

$$\begin{aligned}\boldsymbol{H} = &\frac{\int_{\xi \in P} \boldsymbol{\xi\xi}^t \exp(-\|\xi\|^2/2 + \boldsymbol{x}^t \xi)\, \mathrm{d}\boldsymbol{\xi}}{\int_{\xi \in P} \exp(-\|\xi\|^2/2 + \boldsymbol{x}^t \xi)\, \mathrm{d}\boldsymbol{\xi}} \\ &- \left\{ \frac{\int_{\xi \in P} \boldsymbol{\xi} \exp(-\|\xi\|^2/2 + \boldsymbol{x}^t \xi)\, \mathrm{d}\boldsymbol{\xi}}{\int_{\xi \in P} \exp(-\|\xi\|^2/2 + \boldsymbol{x}^t \xi)\, \mathrm{d}\boldsymbol{\xi}} \right\} \left\{ \frac{\int_{\xi \in P} \boldsymbol{\xi}^t \exp(-\|\xi\|^2/2 + \boldsymbol{x}^t \xi)\, \mathrm{d}\boldsymbol{\xi}}{\int_{\xi \in P} \exp(-\|\xi\|^2/2 + \boldsymbol{x}^t \xi)\, \mathrm{d}\boldsymbol{\xi}} \right\} \\ &- \boldsymbol{I}_p.\end{aligned}$$

Making the transformation $\boldsymbol{\zeta} = \boldsymbol{\Gamma}\boldsymbol{\xi}$ and letting $\boldsymbol{z} = \boldsymbol{\Gamma}\boldsymbol{x}$, we can write $\boldsymbol{H}$ as

$$\begin{aligned}\boldsymbol{H} &= \boldsymbol{\Gamma}^t \left(\frac{\int_{\boldsymbol{\zeta}\in P^*} \boldsymbol{\zeta}\boldsymbol{\zeta}^t \exp(-\|\boldsymbol{\zeta}\|^2/2 + \boldsymbol{z}^t\boldsymbol{\zeta})\,\mathrm{d}\boldsymbol{\zeta}}{\int_{\boldsymbol{\zeta}\in P^*} \exp(-\|\boldsymbol{\zeta}\|^2/2 + \boldsymbol{z}^t\boldsymbol{\zeta})\,\mathrm{d}\boldsymbol{\zeta}} \right. \\ &\quad \left. - \left\{ \frac{\int_{\boldsymbol{\zeta}\in P^*} \boldsymbol{\zeta} \exp(-\|\boldsymbol{\zeta}\|^2/2 + \boldsymbol{z}^t\boldsymbol{\zeta})\,\mathrm{d}\boldsymbol{\zeta}}{\int_{\boldsymbol{\zeta}\in P^*} \exp(-\|\boldsymbol{\zeta}\|^2/2 + \boldsymbol{z}^t\boldsymbol{\zeta})\,\mathrm{d}\boldsymbol{\zeta}} \right\} \left\{ \frac{\int_{\boldsymbol{\zeta}\in P^*} \boldsymbol{\zeta}^t \exp(-\|\boldsymbol{\zeta}\|^2/2 + \boldsymbol{z}^t\boldsymbol{\zeta})\,\mathrm{d}\boldsymbol{\zeta}}{\int_{\boldsymbol{\zeta}\in P^*} \exp(-\|\boldsymbol{\zeta}\|^2/2 + \boldsymbol{z}^t\boldsymbol{\zeta})\,\mathrm{d}\boldsymbol{\zeta}} \right\} - \boldsymbol{I}_p \right) \boldsymbol{\Gamma} \\ &= \boldsymbol{\Gamma}^t \left(\frac{g_j^*(\boldsymbol{z})}{\partial z_i} \right) \boldsymbol{\Gamma},\end{aligned}$$

where $P^* = \{\boldsymbol{\zeta} | \boldsymbol{R}^*\boldsymbol{\zeta} \leqslant \boldsymbol{\alpha}\}$ for $\boldsymbol{R}^* = \boldsymbol{R}\boldsymbol{\Gamma}^t$ and

$$g_j^*(\boldsymbol{z}) = \frac{\int_{\boldsymbol{\zeta}\in P^*} (\zeta_j - z_j) \exp(-\|\boldsymbol{\zeta}\|^2/2 + \boldsymbol{z}^t\boldsymbol{\zeta})\,\mathrm{d}\boldsymbol{\zeta}}{\int_{\boldsymbol{\zeta}\in P^*} \exp(-\|\boldsymbol{\zeta}\|^2/2 + \boldsymbol{z}^t\boldsymbol{\zeta})\,\mathrm{d}\boldsymbol{\zeta}}.$$

Using the fact that $\boldsymbol{R}^*\boldsymbol{1}_p = \boldsymbol{0}_q$ and the result in Case 1, we can see that the matrix $(\partial g_j^*(\boldsymbol{u})/\partial u_i)$ is negative semi-definite, which implies that $\boldsymbol{H}$ is negative semi-definite in Case 2.

Case 3: When $q = p$ and $\boldsymbol{R}$ is any full rank matrix, let $\boldsymbol{\lambda} = (\boldsymbol{\xi}^t, \xi_{p+1})^t$ and $\boldsymbol{y} = (\boldsymbol{x}^t, x_{p+1})^t$ for any ξ_{p+1} and x_{p+1}. Denote

$$\begin{aligned}P_0 &= \{\boldsymbol{\lambda} | \boldsymbol{R}\boldsymbol{\xi} \leqslant \boldsymbol{\alpha},\ -\infty < \xi_{p+1} < \infty\} \\ &= \{\boldsymbol{\lambda} | \boldsymbol{R}_0\boldsymbol{\lambda} \leqslant \boldsymbol{\alpha}\},\end{aligned}$$

where $\boldsymbol{R}_0$ is $p \times (p+1)$ matrix of the form $\boldsymbol{R}_0 = [\boldsymbol{R}, \boldsymbol{0}_p]$. The set P_0 is a polyhedral convex cone in $\mathbb{R}^{p+1}$. Let $\boldsymbol{H}_0 = (\partial g_j^{**}(\boldsymbol{y})/\partial y_i)$, where

$$g_j^{**}(\boldsymbol{y}) = \frac{\int_{\boldsymbol{\lambda}\in P_0} (\lambda_j - y_j) \exp(-\|\boldsymbol{\lambda}\|^2/2 + \boldsymbol{y}^t\boldsymbol{\lambda})\,\mathrm{d}\boldsymbol{\lambda}}{\int_{\boldsymbol{\lambda}\in P_0} \exp(-\|\boldsymbol{\lambda}\|^2/2 + \boldsymbol{y}^t\boldsymbol{\lambda})\,\mathrm{d}\boldsymbol{\lambda}}, \quad j = 1, \ldots, p+1.$$

It is noted that $g_{p+1}^{**}(\boldsymbol{y}) = 0$ and $g_j^{**}(\boldsymbol{y}) = g_j(\boldsymbol{x})$ if $j = 1, \ldots, p$ and that

$$\boldsymbol{H}_0 = \begin{pmatrix} \boldsymbol{H} & \boldsymbol{0}_p \\ \boldsymbol{0}_p^t & 0 \end{pmatrix}.$$

From the result in Case 2, $\boldsymbol{H}_0$ is negative semi-definite, and so is $\boldsymbol{H}$. Therefore, the negative semi-definiteness of $\boldsymbol{H}$ is proved for all the cases (Cases 1–3). $\square$

The following divergence theorem is useful for evaluating the risk function of $\boldsymbol{\delta}^{\mathrm{SH}}$. Let $\mathcal{C}$ be a closed convex set in R^p and let $\partial\mathcal{C}$ be the boundary of $\mathcal{C}$. Denote by $\boldsymbol{v}(\boldsymbol{\xi})$ an outward unit normal vector at a point $\boldsymbol{\xi}$ on $\partial\mathcal{C}$ and by $\ell_{\mathcal{C}}(\boldsymbol{\xi})$ the Lebesgue measure on $\partial\mathcal{C}$.

Lemma 3.5 (*Gauss' divergence theorem*). *For $i = 1, \ldots, p$, let $\phi_i(\boldsymbol{\xi})$ be a differentiable function on $\mathcal{C}$. Then for $\boldsymbol{v}(\boldsymbol{\xi}) = (v_1(\boldsymbol{\xi}), \ldots, v_p(\boldsymbol{\xi}))^t$,*

$$\int_{\boldsymbol{\xi}\in\mathcal{C}} \sum_{i=1}^{p} \frac{\partial}{\partial \xi_i} \phi_i(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} = \int_{\boldsymbol{\xi}\in\partial\mathcal{C}} \sum_{i=1}^{p} v_i(\boldsymbol{\xi})\phi_i(\boldsymbol{\xi})\,\mathrm{d}\ell_{\mathcal{C}}(\boldsymbol{\xi}).$$

Theorem 3.1. *Assume that $0 < a \leqslant 2(p-2)$ for $p \geqslant 3$. Under the restriction* (P), *$\boldsymbol{\delta}^{\mathrm{SH}}$ given by* (3.1) *dominates $\boldsymbol{\delta}^{\mathrm{GB}}$ relative to the loss* (1.2).

Proof. It is noted that the estimation problem and the estimators $\boldsymbol{\delta}^{\mathrm{SH}}$ and $\boldsymbol{\delta}^{\mathrm{GB}}$ are invariant under the transformations $\boldsymbol{X} \to \boldsymbol{X}+\boldsymbol{b}$ and $\boldsymbol{\theta} \to \boldsymbol{\theta}+\boldsymbol{b}$ with $\boldsymbol{\alpha} \to \boldsymbol{\alpha}+\boldsymbol{R}\boldsymbol{b}$ for any vector $\boldsymbol{b}$. Hence we assume without any loss of generality that $\boldsymbol{\alpha} = \mathbf{0}_p$.

The risk difference of the estimators $\boldsymbol{\delta}^{\mathrm{SH}}$ and $\boldsymbol{\delta}^{\mathrm{GB}}$ is written as

$$
\begin{aligned}
\Delta &= R(\boldsymbol{\theta}, \boldsymbol{\delta}^{\mathrm{SH}}) - R(\boldsymbol{\theta}, \boldsymbol{\delta}^{\mathrm{GB}}) \\
&= E\left[-2a \sum_{i=1}^{p} (X_i - \theta_i + g_i(\boldsymbol{X})) \frac{X_i + g_i(\boldsymbol{X})}{\|\boldsymbol{X}+\boldsymbol{g}(\boldsymbol{X})\|^2} + \frac{a^2}{\|\boldsymbol{X}+\boldsymbol{g}(\boldsymbol{X})\|^2} \right].
\end{aligned} \tag{3.6}
$$

From the Stein identity, the cross product term can be evaluated as

$$
\begin{aligned}
&\sum_{i=1}^{p} E\left[\frac{(X_i - \theta_i)(X_i + g_i(\boldsymbol{X}))}{\|\boldsymbol{X}+\boldsymbol{g}(\boldsymbol{X})\|^2} \right] \\
&\quad = \sum_{i=1}^{p} E\left[\frac{\partial}{\partial X_i} \frac{X_i + g_i(\boldsymbol{X})}{\|\boldsymbol{X}+\boldsymbol{g}(\boldsymbol{X})\|^2} \right] \\
&\quad = E\left[\frac{p-2}{\|\boldsymbol{X}+\boldsymbol{g}(\boldsymbol{X})\|^2} + \frac{\sum_{i=1}^{p} \partial g_i(\boldsymbol{X})/\partial X_i}{\|\boldsymbol{X}+\boldsymbol{g}(\boldsymbol{X})\|^2} \right. \\
&\qquad \left. -2 \frac{\sum_{i,j=1}^{p} (X_i + g_i(\boldsymbol{X}))(X_j + g_j(\boldsymbol{X}))(\partial g_j(\boldsymbol{X})/\partial X_i)}{\|\boldsymbol{X}+\boldsymbol{g}(\boldsymbol{X})\|^4} \right].
\end{aligned} \tag{3.7}
$$

Substituting (3.7) into (3.6) yields

$$
\begin{aligned}
\Delta &= E\left[\frac{a^2 - 2a(p-2)}{\|\boldsymbol{X}+\boldsymbol{g}(\boldsymbol{X})\|^2} \right] \\
&\quad -2aE\left[\frac{\sum_{i=1}^{p} \{\partial g_i(\boldsymbol{X})/\partial X_i + g_i(\boldsymbol{X})(X_i + g_i(\boldsymbol{X}))\}}{\|\boldsymbol{X}+\boldsymbol{g}(\boldsymbol{X})\|^2} \right] \\
&\quad +4aE\left[\frac{(\boldsymbol{X}+\boldsymbol{g}(\boldsymbol{X}))^t \boldsymbol{H} (\boldsymbol{X}+\boldsymbol{g}(\boldsymbol{X}))}{\|\boldsymbol{X}+\boldsymbol{g}(\boldsymbol{X})\|^4} \right],
\end{aligned} \tag{3.8}
$$

where $\boldsymbol{H}$ is $p \times p$ matrix whose (i, j)-element is given by $\partial g_j(\boldsymbol{X})/\partial X_i$.

For the second expectation in the r.h.s. of (3.8), we note that

$$
\begin{aligned}
&\sum_{i=1}^{p} \left\{ \frac{\partial g_i(\boldsymbol{x})}{\partial x_i} + g_i(\boldsymbol{x})(x_i + g_i(\boldsymbol{x})) \right\} \\
&\quad = \sum_{i=1}^{p} \frac{(\partial/\partial x_i) \int_{\boldsymbol{\xi} \in P} (\xi_i - x_i) \exp(-\|\boldsymbol{\xi}\|^2/2 + \boldsymbol{x}^t \boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi} \in P} \exp(-\|\boldsymbol{\xi}\|^2/2 + \boldsymbol{x}^t \boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi}}.
\end{aligned}
$$

Using Lemma 3.5, we observe that

$$
\begin{aligned}
&\sum_{i=1}^{p} \frac{\partial}{\partial x_i} \int_{\xi \in P} (\xi_i - x_i) \exp(-\|\boldsymbol{\xi}\|^2/2 + \boldsymbol{x}^t \boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} \\
&\quad = -\int_{\xi \in P} \sum_{i=1}^{p} \frac{\partial}{\partial \xi_i} \{\xi_i \exp(-\|\boldsymbol{\xi}\|^2/2 + \boldsymbol{x}^t \boldsymbol{\xi})\} \mathrm{d}\boldsymbol{\xi} \\
&\quad = -\int_{\xi \in \partial P} \boldsymbol{v}(\boldsymbol{\xi})^t \boldsymbol{\xi} \exp(-\|\boldsymbol{\xi}\|^2/2 + \boldsymbol{x}^t \boldsymbol{\xi})\, \mathrm{d}\ell_P(\boldsymbol{\xi}) \\
&\quad = 0,
\end{aligned}
$$

where ∂P and $\ell_P(\boldsymbol{\xi})$, respectively, denote the boundary of P and the Lebesgue measure on ∂P, and $\boldsymbol{v}(\boldsymbol{\xi})$ is an outward unit normal vector at a point $\boldsymbol{\xi}$ on the boundary. The last equality follows from the fact that the inner product $\boldsymbol{v}(\boldsymbol{\xi})^t \boldsymbol{\xi}$ is equal to zero since the origin belongs to the intersection of hyperplanes $\boldsymbol{r}_i^t \boldsymbol{\xi} = 0,\ i = 1, \ldots, q$, for $\boldsymbol{R} = (\boldsymbol{r}_1, \ldots, \boldsymbol{r}_q)^t$.

Finally, applying Lemma 3.4 to the third term of the r.h.s. in (3.8), we complete the proof of Theorem 3.1. □

We conclude this section by stating some remarks. It is noted that the results given in Section 3 can be extended to the model $\boldsymbol{X} \sim \mathcal{N}_p(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma}$ being any known, positive definite matrix.

Remark 3.1. The interesting relation given in Lemma 3.2 for the restriction (P) with satisfying $\boldsymbol{R}\mathbf{1}_p = \mathbf{0}_q$ implies the following: denote a projection matrix onto $\mathbf{1}_p$ by $\mathcal{P}_1 = \mathbf{1}_p(\mathbf{1}_p^t\mathbf{1}_p)^{-1}\mathbf{1}_p^t$. We then note that $\boldsymbol{\delta}^{\mathrm{GB}} - \boldsymbol{X} = \boldsymbol{g}$ belongs to the orthogonal complement of $\mathbf{1}_p$ since $\mathcal{P}_1\boldsymbol{X} = \mathcal{P}_1\boldsymbol{\delta}^{\mathrm{GB}}$.

Fig. 1 is a rough sketch of each estimator for the simple order restriction (B). Under this restriction, the matrix R satisfies $\boldsymbol{R}\mathbf{1}_p = \mathbf{0}_q$. The left side in Fig. 1 is the case where the observation $\boldsymbol{X}$ is inside $B_x = \{\boldsymbol{X} | X_1 \leqslant X_2 \leqslant \cdots \leqslant X_p\}$ and the right is the case where $\boldsymbol{X}$ is in the set $\{\boldsymbol{X} | X_1 \geqslant X_2 \geqslant \cdots \geqslant X_p\}$. In Fig. 1, the estimator $\boldsymbol{\delta}_B^{\mathrm{ML}}$ is the REML estimator of the form

$$
\boldsymbol{\delta}_B^{\mathrm{ML}} = \mathcal{P}_{B_x}\boldsymbol{X},
$$

where $\mathcal{P}_{B_x}\boldsymbol{X}$ is the orthogonal projection of $\boldsymbol{X}$ onto B_x. Also the estimator $\boldsymbol{\delta}_B^{\mathrm{JS}}$ is the James–Stein type estimator of the form

$$
\boldsymbol{\delta}_B^{\mathrm{JS}} = \mathcal{P}_{B_x}\boldsymbol{X} - \frac{p-2}{\|\boldsymbol{X}\|^2}\boldsymbol{X} I_{B_x}(\boldsymbol{X}),
$$

where $I_{B_x}(\boldsymbol{X})$ is the indicator function of the region B_x. Note that $\boldsymbol{\delta}_B^{\mathrm{JS}}$ dominates $\boldsymbol{\delta}_B^{\mathrm{ML}}$ with respect to the quadratic loss function.

Remark 3.2. The matrix $\boldsymbol{H}$ in Lemma 3.4 can be represented as

$$
\begin{aligned}
\boldsymbol{H} &= E[(\boldsymbol{\xi} - E[\boldsymbol{\xi}|\boldsymbol{X}])(\boldsymbol{\xi} - E[\boldsymbol{\xi}|\boldsymbol{X}])^t|\boldsymbol{X}] - \boldsymbol{I}_p \\
&= \mathbf{Cov}(\boldsymbol{\xi}|\boldsymbol{X}) - \boldsymbol{I}_p.
\end{aligned}
$$

It is noted that $\mathbf{Cov}(\boldsymbol{\xi}|\boldsymbol{X})$ is a covariance matrix of the conditional distribution of $\boldsymbol{\xi}$ given $\boldsymbol{X}$ and that the conditional distribution of $\boldsymbol{\xi}$ given $\boldsymbol{X}$ is a truncated multivariate normal distribution.

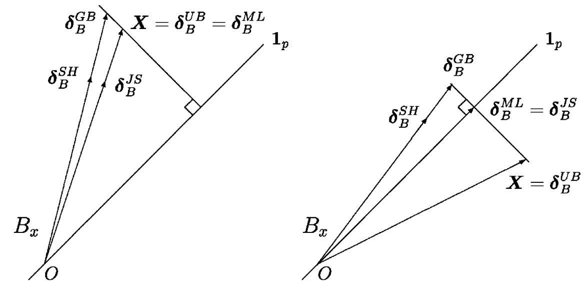

Fig. 1. Rough sketch of each estimator for the simple order restriction (B) (left: $\boldsymbol{X} \in \{\boldsymbol{X}|X_1 \leqslant \cdots \leqslant X_p\}$, right: $\boldsymbol{X} \in \{\boldsymbol{X}|X_1 \geqslant \cdots \geqslant X_p\}$).

It is conceivable that the covariance matrix of non-truncated normal vector, i.e., $\boldsymbol{I}_p$, is larger than $\mathbf{Cov}(\xi|\boldsymbol{X})$. This matter really holds for the truncation on the set P since $\boldsymbol{H}$ is negative semi-definite.

Remark 3.3. As seen in Lehmann and Casella [15], the James–Stein estimator can be improved on by the positive-part Stein estimator when the parameter space is not restricted. However, it seems technically hard to establish the same property in our restricted problem. We thus claim the conjecture that the shrinkage estimator $\boldsymbol{\delta}^{\mathrm{SH}}$ may be dominated by its positive-part estimator $\boldsymbol{\delta}^{\mathrm{PP}} = \boldsymbol{\alpha}_* + \max\{0, 1 - a\|\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{X}) - \boldsymbol{\alpha}_*\|^{-2}\}(\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{X}) - \boldsymbol{\alpha}_*)$.

4. Admissibility

For $p = 2$, we expect that the generalized Bayes estimator is admissible. In this section, we prove the admissibility results for the positive orthant and the simple order restrictions (A) and (B).

Using Lemma 3.3, we first show the admissibility of the generalized Bayes estimator for the positive orthant restriction (A).

Theorem 4.1. *In the estimation issue under the positive orthant restriction* (A), *the generalized Bayes estimator* $\delta_{\mathrm{A}}^{\mathrm{GB}}$ *is admissible for* $p = 2$.

Proof. The method of Brown and Hwang [3] is employed for the proof. Consider a sequence of the prior distributions $\pi_k(\boldsymbol{\theta}) = \{h_k(\|\boldsymbol{\theta}\|)\}^2,\ \boldsymbol{\theta} \in A$, where

$$h_k(t) = \begin{cases} 1 & \text{if } 0 \leqslant t \leqslant 1, \\ 1 - \log t / \log k & \text{if } 1 \leqslant t \leqslant k, \\ 0 & \text{if } k < t. \end{cases} \tag{4.1}$$

The resulting sequence of the Bayes estimators has the form

$$\boldsymbol{\delta}_k^{\pi^*} = \boldsymbol{\delta}_k^{\pi^*}(\boldsymbol{X}) = \frac{\int_{\boldsymbol{\theta} \in A} \boldsymbol{\theta}\{h_k(\|\boldsymbol{\theta}\|)\}^2 \exp(-\|\boldsymbol{\theta} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\theta}}{\int_{\boldsymbol{\theta} \in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2 \exp(-\|\boldsymbol{\theta} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\theta}}.$$

The difference of the Bayes risk functions of the two estimators $\boldsymbol{\delta}_A^{\mathrm{GB}}$ and $\boldsymbol{\delta}_k^{\pi^*}$ is written by

$$\begin{aligned}\Delta_k &= c\int_{\boldsymbol{\theta}\in A}\int \{\|\boldsymbol{\delta}_A^{\mathrm{GB}}-\boldsymbol{\theta}\|^2-\|\boldsymbol{\delta}_k^{\pi^*}-\boldsymbol{\theta}\|^2\}\\ &\quad\times\exp(-\|\boldsymbol{x}-\boldsymbol{\theta}\|^2/2)\,\mathrm{d}\boldsymbol{x}\{h_k(\|\boldsymbol{\theta}\|)\}^2\,\mathrm{d}\boldsymbol{\theta}\\ &= c\int \|\boldsymbol{\delta}_A^{\mathrm{GB}}-\boldsymbol{\delta}_k^{\pi^*}\|^2\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{x} \end{aligned}\tag{4.2}$$

for $c=1/(2\pi)$. Let f_1 and f_2 be probability densities on $\mathcal{X}=A$ with

$$f_1(\boldsymbol{\theta})=\frac{\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)}{\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}},$$
$$f_2(\boldsymbol{\eta})=\frac{\exp(-\|\boldsymbol{\eta}-\boldsymbol{x}\|^2/2)}{\int_{\boldsymbol{\eta}\in A}\exp(-\|\boldsymbol{\eta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\eta}}.$$

Noting that $\{h_k(\|\boldsymbol{\theta}\wedge\boldsymbol{\eta}\|)\}^2\geqslant\{h_k(\|\boldsymbol{\theta}\|)\}^2$ since $\{h_k(t)\}^2$ is nonincreasing of t, we can see that $f_1(\boldsymbol{\theta})f_2(\boldsymbol{\eta})\leqslant f_2(\boldsymbol{\theta}\vee\boldsymbol{\eta})f_1(\boldsymbol{\theta}\wedge\boldsymbol{\eta})$. Hence, it follows from Lemma 3.3 that

$$(\boldsymbol{\delta}_k^{\pi^*})_i=\int_{\boldsymbol{\theta}\in A}\theta_i f_1(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}\leqslant\int_{\boldsymbol{\theta}\in A}\theta_i f_2(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}=\delta_i^{\mathrm{GB}},\quad i=1,\ 2.\tag{4.3}$$

Using the integration by parts shows that

$$\delta_1^{\mathrm{GB}}-x_1=\frac{\int_{\boldsymbol{\theta}\in A}(\theta_1-x_1)\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}{\int_{\boldsymbol{\theta}\in A}\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}=\frac{\exp(-x_1^2/2)}{\int_0^\infty\exp(-(\theta_1-x_1)^2/2)\,\mathrm{d}\theta_1}\tag{4.4}$$

and that

$$\begin{aligned}&(\boldsymbol{\delta}_k^{\pi^*})_1-x_1\\ &\quad=\frac{\int_{\boldsymbol{\theta}\in A}(\theta_1-x_1)\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}{\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}\\ &\quad=\frac{\exp(-x_1^2/2)\int_0^k\{h_k(\theta_2)\}^2\exp(-(\theta_2-x_2)^2/2)\,\mathrm{d}\theta_2}{\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}\\ &\qquad-2\frac{\int_{\boldsymbol{\theta}\in A,\,1\leqslant\|\boldsymbol{\theta}\|\leqslant k}h_k(\|\boldsymbol{\theta}\|)\{\theta_1/(\|\boldsymbol{\theta}\|^2\log k)\}\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}{\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}.\end{aligned}\tag{4.5}$$

Combining (4.3)–(4.5) gives that

$$\begin{aligned}&0\leqslant\delta_1^{\mathrm{GB}}-x_1-(\boldsymbol{\delta}_k^{\pi^*})_1+x_1\\ &\quad=\frac{\exp(-x_1^2/2)}{\int_0^\infty\exp(-(\theta_1-x_1)^2/2)\,\mathrm{d}\theta_1}\\ &\qquad-\frac{\exp(-x_1^2/2)\int_0^k\{h_k(\theta_2)\}^2\exp(-(\theta_2-x_2)^2/2)\,\mathrm{d}\theta_2}{\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}\\ &\qquad+2\frac{\int_{\boldsymbol{\theta}\in A,\,1\leqslant\|\boldsymbol{\theta}\|\leqslant k}\theta_1\|\boldsymbol{\theta}\|^{-2}(\log k)^{-1}h_k(\|\boldsymbol{\theta}\|)\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}{\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}.\end{aligned}\tag{4.6}$$

Note that

$$h_k(\|\boldsymbol{\theta}\|)I(\boldsymbol{\theta}\in A,\ 0\leqslant\|\boldsymbol{\theta}\|\leqslant k)\leqslant h_k(\theta_2)I(0\leqslant\theta_1\leqslant k,\ 0\leqslant\theta_2\leqslant k),$$

where $I(\mathcal{A})$ denotes one if $\mathcal{A}$ is true and zero otherwise. Then we observe that

$$\begin{aligned}
&\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}\\
&\quad\leqslant\int_{0\leqslant\theta_1\leqslant k,\,0\leqslant\theta_2\leqslant k}\{h_k(\theta_2)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}\\
&\quad\leqslant\int_0^\infty\exp(-(\theta_1-x_1)^2/2)\,\mathrm{d}\theta_1\int_0^k\{h_k(\theta_2)\}^2\exp(-(\theta_2-x_2)^2/2)\,\mathrm{d}\theta_2,
\end{aligned}$$

which is used to evaluate the second term in the r.h.s. of Eq. (4.6), and we can see that

$$\begin{aligned}
0&\leqslant\delta_1^{\mathrm{GB}}-x_1-(\boldsymbol{\delta}_k^{\pi^*})_1+x_1\\
&\leqslant 2\frac{\int_{\boldsymbol{\theta}\in A,\,1\leqslant\|\boldsymbol{\theta}\|\leqslant k}\theta_1\|\boldsymbol{\theta}\|^{-2}(\log k)^{-1}h_k(\|\boldsymbol{\theta}\|)\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}{\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}.
\end{aligned}\tag{4.7}$$

Similarly, we have

$$\begin{aligned}
0&\leqslant\delta_2^{\mathrm{GB}}-x_2-(\boldsymbol{\delta}_k^{\pi^*})_2+x_2\\
&\leqslant 2\frac{\int_{\boldsymbol{\theta}\in A,\,1\leqslant\|\boldsymbol{\theta}\|\leqslant k}\theta_2\|\boldsymbol{\theta}\|^{-2}(\log k)^{-1}h_k(\|\boldsymbol{\theta}\|)\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}{\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}.
\end{aligned}$$

Therefore, the term $\|\boldsymbol{\delta}_A^{\mathrm{GB}}-\boldsymbol{\delta}_k^{\pi^*}\|^2$ in (4.2) is evaluated as

$$\begin{aligned}
&\|\boldsymbol{\delta}_A^{\mathrm{GB}}-\boldsymbol{\delta}_k^{\pi^*}\|^2\\
&\quad\leqslant 4\left\{\frac{\int_{\boldsymbol{\theta}\in A,\,1\leqslant\|\boldsymbol{\theta}\|\leqslant k}\theta_1\|\boldsymbol{\theta}\|^{-2}(\log k)^{-1}h_k(\|\boldsymbol{\theta}\|)\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}{\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}\right\}^2\\
&\qquad+4\left\{\frac{\int_{\boldsymbol{\theta}\in A,\,1\leqslant\|\boldsymbol{\theta}\|\leqslant k}\theta_2\|\boldsymbol{\theta}\|^{-2}(\log k)^{-1}h_k(\|\boldsymbol{\theta}\|)\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}{\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}\right\}^2\\
&\quad\leqslant 4\frac{\int_{\boldsymbol{\theta}\in A,\,1\leqslant\|\boldsymbol{\theta}\|\leqslant k}\|\boldsymbol{\theta}\|^{-2}(\log k)^{-2}\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}}{\int_{\boldsymbol{\theta}\in A}\{h_k(\|\boldsymbol{\theta}\|)\}^2\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}},
\end{aligned}$$

where the second inequality follows from Schwarz's inequality. Finally, we obtain that

$$\begin{aligned}
\Delta_k&\leqslant 4c\int\int_{\boldsymbol{\theta}\in A,\,1\leqslant\|\boldsymbol{\theta}\|\leqslant k}\|\boldsymbol{\theta}\|^{-2}(\log k)^{-2}\exp(-\|\boldsymbol{\theta}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{x}\\
&=4\int_{\boldsymbol{\theta}\in A,\,1\leqslant\|\boldsymbol{\theta}\|\leqslant k}\frac{1}{\|\boldsymbol{\theta}\|^2(\log k)^2}\,\mathrm{d}\boldsymbol{\theta}\\
&=2\pi\int_1^k\frac{1}{r(\log k)^2}\,\mathrm{d}r=2\pi(\log k)^{-1}\to 0,\quad k\to\infty,
\end{aligned}$$

which completes the proof for the restriction (A) with $p=2$. □

Theorem 4.2. *In the estimation issue under the simple order restriction* (B), *the generalized Bayes estimator* δ_B^{GB} *is admissible for* $p=2$.

Proof. Take the sequence of the prior distributions $\pi_k(\boldsymbol{\theta}) = \{h_k(\|\boldsymbol{\theta}\|)\}^2, \boldsymbol{\theta} \in B$, where $h_k(t)$ is defined as (4.1). Then we denote by $\boldsymbol{\delta}_k^{\pi^*} = \boldsymbol{\delta}_k^{\pi^*}(\boldsymbol{X})$ the sequence of the corresponding Bayes estimators. The difference of the Bayes risk functions of the two estimators $\boldsymbol{\delta}_B^{\mathrm{GB}}$ and $\boldsymbol{\delta}_k^{\pi^*}$ is given by

$$\Delta_k = c \int \|\boldsymbol{\delta}_B^{\mathrm{GB}} - \boldsymbol{\delta}_k^{\pi^*}\|^2 \int_{\boldsymbol{\theta} \in B} \{h_k(\|\boldsymbol{\theta}\|)\}^2 \exp(-\|\boldsymbol{\theta} - \boldsymbol{x}\|^2/2)\, \mathrm{d}\boldsymbol{\theta}\, \mathrm{d}\boldsymbol{x},$$

where $c = (2\pi)^{-1}$.

Let $\boldsymbol{\Gamma}$ be the orthogonal matrix such as

$$\boldsymbol{\Gamma} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Making the orthogonal transformations $\boldsymbol{\xi} = \boldsymbol{\Gamma}\boldsymbol{\theta}$ and $\boldsymbol{z} = \boldsymbol{\Gamma}\boldsymbol{x}$, we can write the difference of the Bayes risk functions as

$$\Delta_k = c \int \|\boldsymbol{\delta}_B^{\mathrm{GB}}(\boldsymbol{z}) - \boldsymbol{\delta}_k^{\pi^*}(\boldsymbol{z})\|^2 \int_{\boldsymbol{\xi} \in P_b} \{h_k(\|\boldsymbol{\xi}\|)\}^2 \exp(-\|\boldsymbol{\xi} - \boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}\, \mathrm{d}\boldsymbol{z}, \tag{4.8}$$

where $P_b = \{\boldsymbol{\xi} | -\infty < \xi_1 < \infty,\ 0 \leqslant \xi_2 < \infty\}$ and

$$\boldsymbol{\delta}_B^{\mathrm{GB}}(\boldsymbol{z}) = \frac{\int_{\boldsymbol{\xi} \in P_b} \boldsymbol{\xi} \exp(-\|\boldsymbol{\xi} - \boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi} \in P_b} \exp(-\|\boldsymbol{\xi} - \boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}},$$

$$\boldsymbol{\delta}_k^{\pi^*}(\boldsymbol{z}) = \frac{\int_{\boldsymbol{\xi} \in P_b} \boldsymbol{\xi} \{h_k(\|\boldsymbol{\xi}\|)\}^2 \exp(-\|\boldsymbol{\xi} - \boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi} \in P_b} \{h_k(\|\boldsymbol{\xi}\|)\}^2 \exp(-\|\boldsymbol{\xi} - \boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}.$$

By the integration of parts, each element of $\boldsymbol{\delta}_B^{\mathrm{GB}}(\boldsymbol{z})$ and $\boldsymbol{\delta}_k^{\pi^*}(\boldsymbol{z})$ can be rewritten as

$$\delta_1^{\mathrm{GB}}(\boldsymbol{z}) = z_1,$$

$$\delta_2^{\mathrm{GB}}(\boldsymbol{z}) = z_2 + \frac{\exp(-z_2^2/2)}{\int_0^\infty \exp(-(\xi_2 - z_2)^2/2)\, \mathrm{d}\xi_2},$$

$$\{\boldsymbol{\delta}_k^{\pi^*}(\boldsymbol{z})\}_1 = z_1,$$

$$\begin{aligned}\{\boldsymbol{\delta}_k^{\pi^*}(\boldsymbol{z})\}_2 = z_2 &+ \frac{\exp(-z_2^2/2) \int_{0 \leqslant |\xi_1| \leqslant k} \{h_k(|\xi_1|)\}^2 \exp(-(\xi_1 - z_1)^2/2)\, \mathrm{d}\xi_1}{\int_{\boldsymbol{\xi} \in P_b} \{h_k(\|\boldsymbol{\xi}\|)\}^2 \exp(-\|\boldsymbol{\xi} - \boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}} \\ &- 2\frac{\int_{\boldsymbol{\xi} \in P_b,\, 1 \leqslant \|\boldsymbol{\xi}\| \leqslant k} \xi_2 \|\boldsymbol{\xi}\|^{-2} (\log k)^{-1} h_k(\|\boldsymbol{\xi}\|) \exp(-\|\boldsymbol{\xi} - \boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi} \in P_b} \{h_k(\|\boldsymbol{\xi}\|)\}^2 \exp(-\|\boldsymbol{\xi} - \boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}.\end{aligned}$$

We here show that $\delta_2^{\mathrm{GB}}(\boldsymbol{z}) - \{\boldsymbol{\delta}_k^{\pi^*}(\boldsymbol{z})\}_2 \geqslant 0$ by using Lemma 3.3. Let f_1 and f_2 be probability densities on $\mathcal{X} = P_b$ with

$$f_1(\xi_2) = \frac{\exp(-(\xi_2 - z_2)^2/2) \int_{-\infty}^{\infty} \{h_k(\|\boldsymbol{\xi}\|)\}^2 \exp(-(\xi_1 - z_1)^2/2)\, \mathrm{d}\xi_1}{\int_{\boldsymbol{\xi} \in P_b} \{h_k(\|\boldsymbol{\xi}\|)\}^2 \exp(-\|\boldsymbol{\xi} - \boldsymbol{z}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}},$$

$$f_2(\eta_2) = \frac{\exp(-(\eta_2 - z_2)^2/2)}{\int_0^\infty \exp(-(z_2 - \eta_2)^2/2)\, \mathrm{d}\eta_2}.$$

Now, we can see that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\xi_2}\int_{-\infty}^{\infty}\{h_k(\|\xi\|)\}^2\exp(-(\xi_1-z_1)^2/2)\,\mathrm{d}\xi_1\\
&\quad=\begin{cases}-2\displaystyle\int_{-\sqrt{k^2-\xi_2^2}}^{\sqrt{k^2-\xi_2^2}}\frac{\xi_2}{\|\xi\|^2\log k}h_k(\|\xi\|)\exp(-(\xi_1-z_1)^2/2)\,\mathrm{d}\xi_1 & \text{if } 0\leqslant\xi_2\leqslant k,\\ 0 & \text{otherwise,}\end{cases}
\end{aligned}$$

which implies that $\int_{-\infty}^{\infty}\{h_k(\|\xi\|)\}^2\exp(-(\xi_1-z_1)^2/2)\,\mathrm{d}\xi_1$ is nonincreasing of ξ_2. Then it holds that $f_1(\xi_2)f_2(\eta_2)\leqslant f_2(\max(\xi_2,\ \eta_2))f_1(\min(\xi_2,\ \eta_2))$, which yields from Lemma 3.3 that

$$\{\boldsymbol{\delta}_k^{\pi^*}(z)\}_2=\int_0^{\infty}\xi_2 f_1(\xi_2)\,\mathrm{d}\xi_2\leqslant\int_0^{\infty}\xi_2 f_2(\xi_2)\,\mathrm{d}\xi_2=\delta_2^{\mathrm{GB}}(z). \tag{4.9}$$

Using the inequality (4.9) and the evaluation method similar to (4.7), we obtain that

$$\begin{aligned}
0&\leqslant\delta_2^{\mathrm{GB}}(z)-\{\boldsymbol{\delta}_k^{\pi^*}(z)\}_2\\
&\leqslant 2\frac{\int_{\boldsymbol{\xi}\in P_b,\,1\leqslant\|\boldsymbol{\xi}\|\leqslant k}\xi_2\|\boldsymbol{\xi}\|^{-2}(\log k)^{-1}h_k(\|\boldsymbol{\xi}\|)\exp(-\|\boldsymbol{\xi}-z\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi}\in P_b}\{h_k(\|\boldsymbol{\xi}\|)\}^2\exp(-\|\boldsymbol{\xi}-z\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}.
\end{aligned}$$

Then the term $\|\boldsymbol{\delta}_B^{\mathrm{GB}}(z)-\boldsymbol{\delta}_k^{\pi^*}(z)\|^2$ in (4.8) can be evaluated as

$$\begin{aligned}
\|\boldsymbol{\delta}_B^{\mathrm{GB}}(z)-\boldsymbol{\delta}_k^{\pi^*}(z)\|^2&=(\delta_2^{\mathrm{GB}}(z)-\{\boldsymbol{\delta}_k^{\pi^*}(z)\}_2)^2\\
&\leqslant 4\frac{\int_{\boldsymbol{\xi}\in P_b,\,1\leqslant\|\boldsymbol{\xi}\|\leqslant k}\xi_2^2\|\boldsymbol{\xi}\|^{-4}(\log k)^{-2}\exp(-\|\boldsymbol{\xi}-z\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi}\in P_b}\{h_k(\|\boldsymbol{\xi}\|)\}^2\exp(-\|\boldsymbol{\xi}-z\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}.
\end{aligned}$$

Hence from (4.8), we observe that

$$\Delta_k\leqslant 4\int_{\boldsymbol{\xi}\in P_b,\,1\leqslant\|\boldsymbol{\xi}\|\leqslant k}\frac{\xi_2^2}{\|\boldsymbol{\xi}\|^4(\log k)^2}\mathrm{d}\boldsymbol{\xi}=2\pi(\log k)^{-1}\to 0,\quad k\to\infty,$$

which proves Theorem 4.2. □

Remark 4.1. It would be interesting to know why the generalized Bayes estimator over the restricted space is not admissible when $p\geqslant 3$. Although it is difficult to give a convinced answer to the query, an intuitive reason of the inadmissibility is that the restricted space expands as spaciously as the unrestricted space for $p\geqslant 3$ when $\|\boldsymbol{\theta}\|$ goes to infinity. From a technical aspect, the similar argument of the proof of Theorem 4.1 gives that the difference of the Bayes risk function like (4.2) is evaluated as, for $p\geqslant 3$,

$$\begin{aligned}
\Delta_k&\leqslant 4\int_{\boldsymbol{\theta}\in A,\,1\leqslant\|\boldsymbol{\theta}\|\leqslant k}\frac{1}{\|\boldsymbol{\theta}\|^2(\log k)^2}\,\mathrm{d}\boldsymbol{\theta}=(\text{const.})\times\int_{1\leqslant r\leqslant k}\frac{r^{p-3}}{(\log k)^2}\,\mathrm{d}r\\
&=\frac{(\text{const.})}{p-2}\times\frac{k^{p-2}-1}{(\log k)^2},
\end{aligned}$$

which does not converge to zero as $k\to\infty$. See also Brown and Hwang [3].

5. Numerical studies

5.1. The pituitary fissure data

We first apply the generalized Bayes and its shrinkage procedures to the practical data in Example 1.1.1 of Robertson et al. [16]. The data are the size of the pituitary fissure and are measured with respect to four kinds of age. It is reasonable to assume that the pituitary fissure increases with age, namely, the simple order restriction. For more details, see Examples 1.1.1 and 1.2.2 of Robertson et al. [16].

Table 1 shows the results of five estimators; the unbiased estimator, the REML estimator, the James–Stein estimator, the generalized Bayes estimator, and its shrinkage estimator, denoted by UB, ML, JS, GB and SH, respectively. The four estimators ML, JS, GB and SH are imposed on the simple order restriction (B), respectively, where ML and JS are defined as Remark 3.1. Table 1 suggests that UB violates the order corresponding to subjects aged 8, 10, and 12 and, in the same cases, ML gives the same estimates, namely 22.22. For this data, JS is equivalent to ML. On the other hand, GB gives natural estimates since the estimates for four ages are distinct and keep the increasing order with age. In this case, SH gives slightly smaller values than GB.

5.2. Risk comparison via Monte Carlo simulations

In this section, we investigate the risk performances of estimators for the restrictions (A)–(C) via Monte Carlo simulations. The risk values of the estimators are derived from 10,000 replications and these replications are generated from $\mathcal{N}_p(\boldsymbol{\theta}, \boldsymbol{I}_p)$ with $p = 3$. For a nonnegative value c, we consider the three cases: $\boldsymbol{\theta} = (c, c, c)^t$, $\boldsymbol{\theta} = (-c, 0, c)^t$ and $\boldsymbol{\theta} = (-c, c, c)^t$ which correspond to the positive orthant restriction (A), the simple order restriction (B) and the tree order restriction (C), respectively. Figs. 2–4 are the simulation results in the cases of the positive orthant, the simple order, and the tree order restrictions, respectively. In Figs. 2–4, UB, GB, and SH stand for $\boldsymbol{\delta}^{\mathrm{UB}}$, $\boldsymbol{\delta}_\lambda^{\mathrm{GB}}$, and $\boldsymbol{\delta}_\lambda^{\mathrm{SH}}$ with $a = p - 2$, respectively.

Moreover, we calculated the risk values of the REML estimator and the James–Stein type estimator for each restriction on the means. Let $A_x = \{\boldsymbol{X} | X_i \geqslant 0 \text{ for } i = 1, \ldots, p\}$, $B_x = \{\boldsymbol{X} | X_1 \leqslant X_2 \leqslant \cdots \leqslant X_p\}$ and $C_x = \{\boldsymbol{X} | X_1 \leqslant X_i \text{ for } i = 2, \ldots, p\}$. For the restricted space $\lambda = A_x, B_x$ and C_x, the REML estimator and the James–Stein type estimator are given by

$$\boldsymbol{\delta}_\lambda^{\mathrm{ML}} = \mathcal{P}_\lambda \boldsymbol{X},$$

$$\boldsymbol{\delta}_\lambda^{\mathrm{JS}} = \mathcal{P}_\lambda \boldsymbol{X} - \frac{p-2}{\|\boldsymbol{X}\|^2} \boldsymbol{X} I_\lambda(\boldsymbol{X}),$$

Table 1
Estimates for the pituitary fissure data

Age	8	10	12	14
UB	22.50	23.33	20.83	24.25
ML	22.22	22.22	22.22	24.25
JS	22.22	22.22	22.22	24.25
GB	20.68	22.08	23.03	25.15
SH	20.66	22.06	23.01	25.13

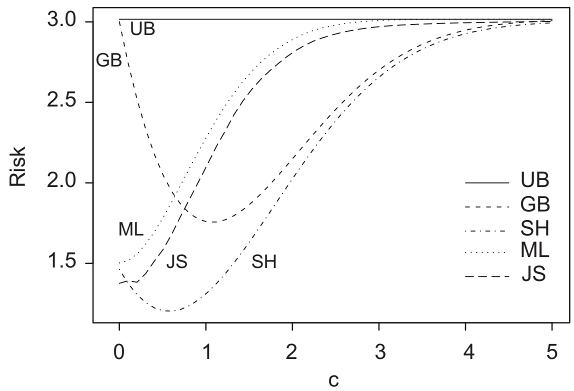

Fig. 2. Risk values in estimation under positive orthant restriction $\boldsymbol{\theta} = (c,\ c,\ c)^t,\ 0 \leqslant c \leqslant 5$.

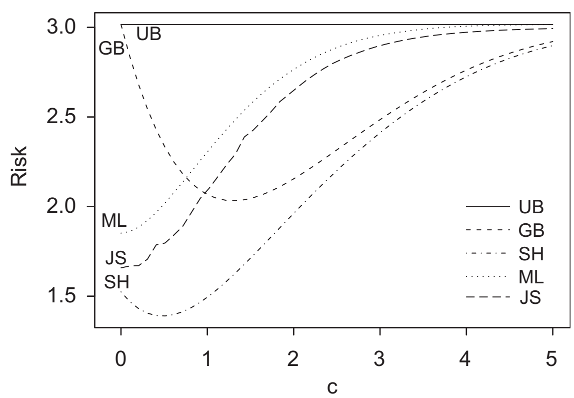

Fig. 3. Risk values in estimation under simple order restriction $\boldsymbol{\theta} = (-c,\ 0,\ c)^t,\ 0 \leqslant c \leqslant 5$.

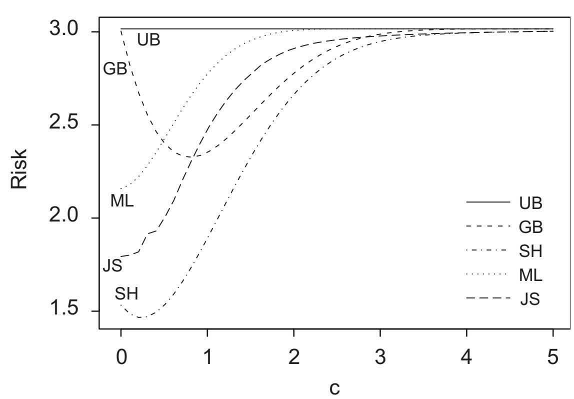

Fig. 4. Risk values in estimation under tree order restriction $\boldsymbol{\theta} = (-c,\ c,\ c)^t,\ 0 \leqslant c \leqslant 5$.

where $\mathcal{P}_\lambda \boldsymbol{X}$ is the orthogonal projection of $\boldsymbol{X}$ onto λ, and $I_\lambda(\boldsymbol{X})$ is the indicator function of the region λ. In Figs. 2–4, $\boldsymbol{\delta}_\lambda^{\mathrm{ML}}$ and $\boldsymbol{\delta}_\lambda^{\mathrm{JS}}$ are abbreviated as ML and JS, respectively.

We summarize the results of Figs. 2–4 as follows:

(1) Under simple order and tree order restrictions, SH $\succ$ JS $\succ$ ML $\succ$ GB $\succ$ UB when $\boldsymbol{\theta}$ is near to zero vector, where $\delta^a \succ \delta^b$ means that δ^a is better than δ^b. Under positive orthant restriction, JS $\succ$ SH $\succ$ ML $\succ$ GB $\succ$ UB when $\boldsymbol{\theta}$ is near to zero vector.
(2) Under each restriction, SH $\succ$ GB $\succ$ JS $\succ$ ML $\succ$ UB when c is larger than one. As c increases, namely, $\boldsymbol{\theta}$ is far from zero vector, the risk values of SH, GB, JS, and ML approach that of UB.
(3) GB has the same risk value as UB has in the case where $\boldsymbol{\theta}$ is zero vector. This fact analytically follows from (2.1), (2.2) and (2.3) in Section 2.
(4) As shown in Theorem 3.1, SH reduces the risk favorably over GB, in particular, when c is small. For the simulation set-up considered here, the figures show that SH is superior to ML.

Appendix A. Proofs

A.1 Proof of Theorem 2.2

The arguments used in Hartigan [9] are useful for the proof. Let

$$\boldsymbol{g}_\lambda = (g_1^\lambda(\boldsymbol{X}), \ldots, g_p^\lambda(\boldsymbol{X}))^t = \frac{\int_{\xi\in\lambda} (\boldsymbol{\xi} - \boldsymbol{X}) \exp(-\|\boldsymbol{\xi} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}{\int_{\xi\in\lambda} \exp(-\|\boldsymbol{\xi} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}$$

for $\lambda = A, B$ and C. For each λ, the difference of the generalized Bayes and the unbiased estimators is represented as

$$\begin{aligned}\Delta_\lambda &= E[\|\boldsymbol{X} + \boldsymbol{g}_\lambda - \boldsymbol{\theta}\|^2 - \|\boldsymbol{X} - \boldsymbol{\theta}\|^2] \\ &= E[2(\boldsymbol{X} - \boldsymbol{\theta})^t \boldsymbol{g}_\lambda + \|\boldsymbol{g}_\lambda\|^2] \\ &= E[(\boldsymbol{X} - \boldsymbol{\theta})^t \boldsymbol{g}_\lambda + \|\boldsymbol{g}_\lambda\|^2] + E[(\boldsymbol{X} - \boldsymbol{\theta})^t \boldsymbol{g}_\lambda].\end{aligned}$$

Applying the Stein identity to $E[(\boldsymbol{X} - \boldsymbol{\theta})^t \boldsymbol{g}_\lambda]$, we have

$$\begin{aligned}E[(\boldsymbol{X} - \boldsymbol{\theta})^t \boldsymbol{g}_\lambda] &= E\left[\sum_{i=1}^p \frac{\partial g_i^\lambda(\boldsymbol{X})}{\partial X_i}\right] \\ &= -p + E\left[\sum_{i=1}^p \left[\frac{\int_{\xi\in\lambda} (\xi_i - X_i)^2 \exp(-\|\boldsymbol{\xi} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}{\int_{\xi\in\lambda} \exp(-\|\boldsymbol{\xi} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}\right.\right. \\ &\qquad \left.\left. - \left\{\frac{\int_{\xi\in\lambda} (\xi_i - X_i) \exp(-\|\boldsymbol{\xi} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}{\int_{\xi\in\lambda} \exp(-\|\boldsymbol{\xi} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}\right\}^2\right]\right] \\ &= -p - \|\boldsymbol{g}_\lambda\|^2 + E\left[\sum_{i=1}^p \frac{\int_{\xi\in\lambda} (\xi_i - X_i)^2 \exp(-\|\boldsymbol{\xi} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}{\int_{\xi\in\lambda} \exp(-\|\boldsymbol{\xi} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}\right],\end{aligned}$$

which yields that

$$\Delta_\lambda = E\left[(\boldsymbol{X} - \boldsymbol{\theta})^t \boldsymbol{g}_\lambda - p + \sum_{i=1}^p \frac{\int_{\xi\in\lambda} (\xi_i - X_i)^2 \exp(-\|\boldsymbol{\xi} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}{\int_{\xi\in\lambda} \exp(-\|\boldsymbol{\xi} - \boldsymbol{X}\|^2/2)\, \mathrm{d}\boldsymbol{\xi}}\right]. \tag{A.1}$$

From the fact that $(X_i-\theta_i)(\xi_i-X_i)=(\xi_i-\theta_i)(\xi_i-X_i)-(\xi_i-X_i)^2$, the term $E[(\boldsymbol{X}-\boldsymbol{\theta})^t\boldsymbol{g}_\lambda]$ is rewritten as

$$\begin{aligned}E[(\boldsymbol{X}-\boldsymbol{\theta})^t\boldsymbol{g}_\lambda] &= E\left[\sum_{i=1}^p \frac{\int_{\boldsymbol{\xi}\in\lambda}(X_i-\theta_i)(\xi_i-X_i)\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi}\in\lambda}\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}\right]\\ &= E\left[\sum_{i=1}^p \frac{\int_{\boldsymbol{\xi}\in\lambda}(\xi_i-\theta_i)(\xi_i-X_i)\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi}\in\lambda}\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}\right.\\ &\quad\left.-\sum_{i=1}^p \frac{\int_{\boldsymbol{\xi}\in\lambda}(\xi_i-X_i)^2\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi}\in\lambda}\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}\right],\end{aligned}$$

which, from (A.1), gives that

$$\Delta_\lambda = E\left[\sum_{i=1}^p \frac{\int_{\boldsymbol{\xi}\in\lambda}(\xi_i-\theta_i)(\xi_i-X_i)\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi}\in\lambda}\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}\right]-p. \tag{A.2}$$

It is noted that

$$\begin{aligned}&\frac{\partial}{\partial\xi_i}\{(\xi_i-\theta_i)\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2)\}\\ &\quad=\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2)-(\xi_i-\theta_i)(\xi_i-X_i)\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2).\end{aligned}$$

Then, it follows from (A.2) and the above expression that

$$\Delta_\lambda = E\left[\sum_{i=1}^p \frac{\int_{\boldsymbol{\xi}\in\lambda}(\partial/\partial\xi_i)\{(\theta_i-\xi_i)\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2)\}\mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi}\in\lambda}\exp(-\|\boldsymbol{\xi}-\boldsymbol{X}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}\right]. \tag{A.3}$$

Finally, we shall evaluate the r.h.s. of Eq. (A.3) directly for each restricted space.

For the simple order restriction (B), it is seen that

$$\begin{aligned}&\sum_{i=1}^p\int_{\boldsymbol{\xi}\in B}\frac{\partial}{\partial\xi_i}\left\{(\theta_i-\xi_i)\exp(-\|\boldsymbol{\xi}-\boldsymbol{x}\|^2/2)\right\}\mathrm{d}\boldsymbol{\xi}\\ &\quad=\sum_{i=1}^p\int_{\boldsymbol{\xi}_{-i}\in B_{-i}}\left[(\theta_i-\xi_i)\exp(-\|\boldsymbol{\xi}-\boldsymbol{x}\|^2/2)\right]_{\xi_i=\xi_{i-1}}^{\xi_{i+1}}\mathrm{d}\boldsymbol{\xi}_{-i}\\ &\quad=\int_{\boldsymbol{\xi}_{-1}\in B_{-1}}(\theta_1-\xi_2)\psi(\xi_2|x_1)\prod_{j\neq 1}\psi(\xi_j|x_j)\,\mathrm{d}\boldsymbol{\xi}_{-1}\\ &\qquad+\sum_{i=2}^{p-1}\int_{\boldsymbol{\xi}_{-i}\in B_{-i}}\{(\theta_i-\xi_{i+1})\psi(\xi_{i+1}|x_i)-(\theta_i-\xi_{i-1})\psi(\xi_{i-1}|x_i)\}\prod_{j\neq i}\psi(\xi_j|x_j)\,\mathrm{d}\boldsymbol{\xi}_{-i}\\ &\qquad-\int_{\boldsymbol{\xi}_{-p}\in B_{-p}}(\theta_p-\xi_{p-1})\psi(\xi_{p-1}|x_p)\prod_{j\neq p}\psi(\xi_j|x_j)\,\mathrm{d}\boldsymbol{\xi}_{-p},\end{aligned} \tag{A.4}$$

where $\xi_0 = -\infty$ and $\xi_{p+1} = \infty$, and B_{-i} is defined in Theorem 2.2. It is noted that for any integrable function $h(\cdot)$ on $\mathbb{R}$,

$$\begin{aligned}&\int_{\boldsymbol{\xi}_{-i}\in B_{-i}} h(\xi_{i+1})\psi(\xi_{i+1}|x_i)\prod_{j\neq i}\psi(\xi_j|x_j)\,\mathrm{d}\boldsymbol{\xi}_{-i}\\&\quad=\int_{\boldsymbol{\xi}_{-(i+1)}\in B_{-(i+1)}} h(\xi_i)\psi(\xi_i|x_{i+1})\prod_{j\neq i+1}\psi(\xi_j|x_j)\,\mathrm{d}\boldsymbol{\xi}_{-(i+1)},\end{aligned}$$

which can be used to simplify (A.4) as

$$\begin{aligned}&\sum_{i=1}^{p}\int_{\boldsymbol{\xi}\in B}\frac{\partial}{\partial\xi_i}\left\{(\theta_i-\xi_i)\exp(-\|\boldsymbol{\xi}-\boldsymbol{x}\|^2/2)\right\}\mathrm{d}\boldsymbol{\xi}\\&\quad=\sum_{i=1}^{p-1}(\theta_i-\theta_{i+1})\int_{\boldsymbol{\xi}_{-i}\in B_{-i}}\psi(\xi_{i+1}|x_i)\prod_{j\neq i}\psi(\xi_j|x_j)\,\mathrm{d}\boldsymbol{\xi}_{-i}.\end{aligned}$$

Substituting this into (A.3) gives the risk expression (2.2).

Using the similar way also gives the risk expressions (2.1) and (2.3), respectively, and the derivations are omitted. □

A.2 Proof of Lemma 3.1

For the case of $q = p$, we take $\boldsymbol{q}(\boldsymbol{x}) = (\boldsymbol{R}^t)^{-1}\boldsymbol{g}(\boldsymbol{x})$, which gives the expression (3.2). In the case of $q < p$, let $\boldsymbol{\Gamma} = [\boldsymbol{R}^t(\boldsymbol{R}\boldsymbol{R}^t)^{-1/2},\ \boldsymbol{R}_*^t]^t$ be an orthogonal matrix where $\boldsymbol{R}_*\boldsymbol{R}^t = \boldsymbol{0}_{(p-q)\times q}$. Making the transformation $\boldsymbol{\eta} = \boldsymbol{\Gamma}\boldsymbol{\xi}$ and letting $\boldsymbol{z} = \boldsymbol{\Gamma}\boldsymbol{x}$, we get

$$\begin{aligned}\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{x})-\boldsymbol{x}&=\frac{\int_{\boldsymbol{\xi}\in P}(\boldsymbol{\xi}-\boldsymbol{x})\exp(-\|\boldsymbol{\xi}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}{\int_{\boldsymbol{\xi}\in P}\exp(-\|\boldsymbol{\xi}-\boldsymbol{x}\|^2/2)\,\mathrm{d}\boldsymbol{\xi}}\\&=[\boldsymbol{R}^t(\boldsymbol{R}\boldsymbol{R}^t)^{-1/2},\ \boldsymbol{R}_*^t]\frac{\int_{\boldsymbol{\eta}\in P_\eta}(\boldsymbol{\eta}-\boldsymbol{z})\exp(-\|\boldsymbol{\eta}-\boldsymbol{z}\|^2/2)\,\mathrm{d}\boldsymbol{\eta}}{\int_{\boldsymbol{\eta}\in P_\eta}\exp(-\|\boldsymbol{\eta}-\boldsymbol{z}\|^2/2)\,\mathrm{d}\boldsymbol{\eta}},\end{aligned}$$

where $P_\eta = \{\boldsymbol{\eta}|\boldsymbol{R}\boldsymbol{\Gamma}^t\boldsymbol{\eta}\leqslant\boldsymbol{\alpha}\}$. Noting that $\boldsymbol{R}\boldsymbol{\Gamma}^t = [(\boldsymbol{R}\boldsymbol{R}^t)^{1/2},\ \boldsymbol{0}_{q\times(p-q)}]$, we can rewrite the integral domain P_η as

$$P_\eta = \{\boldsymbol{\eta}|(\boldsymbol{R}\boldsymbol{R}^t)^{1/2}\boldsymbol{\eta}_{(q)}\leqslant\boldsymbol{\alpha},\ -\infty<\eta_{q+1}<\infty,\ldots,-\infty<\eta_p<\infty\}$$

for $\boldsymbol{\eta}_{(q)} = (\eta_1,\ldots,\eta_q)^t$. Hence, it follows that for $i = q+1,\ldots,p$,

$$\begin{aligned}\frac{\int_{\boldsymbol{\eta}\in P_\eta}(\eta_i-z_i)\exp(-\|\boldsymbol{\eta}-\boldsymbol{z}\|^2/2)\,\mathrm{d}\boldsymbol{\eta}}{\int_{\boldsymbol{\eta}\in P_\eta}\exp(-\|\boldsymbol{\eta}-\boldsymbol{z}\|^2/2)\,\mathrm{d}\boldsymbol{\eta}}&=\frac{\int_{-\infty}^{\infty}(\eta_i-z_i)\exp(-(\eta_i-z_i)^2/2)\,\mathrm{d}\eta_i}{\int_{-\infty}^{\infty}\exp(-(\eta_i-z_i)^2/2)\,\mathrm{d}\eta_i}\\&=0,\end{aligned}$$

which yields that

$$\begin{aligned}\boldsymbol{\delta}^{\mathrm{GB}}(\boldsymbol{x})-\boldsymbol{x}&=\boldsymbol{R}^t(\boldsymbol{R}\boldsymbol{R}^t)^{-1/2}\frac{\int_{\boldsymbol{\eta}_{(q)}\in P_\eta^*}(\boldsymbol{\eta}_{(q)}-\boldsymbol{z}_{(q)})\exp(-\|\boldsymbol{\eta}_{(q)}-\boldsymbol{z}_{(q)}\|^2/2)\,\mathrm{d}\boldsymbol{\eta}_{(q)}}{\int_{\boldsymbol{\eta}_{(q)}\in P_\eta^*}\exp(-\|\boldsymbol{\eta}_{(q)}-\boldsymbol{z}_{(q)}\|^2/2)\,\mathrm{d}\boldsymbol{\eta}_{(q)}}\\&\equiv\boldsymbol{R}^t\boldsymbol{q}(\boldsymbol{x}),\end{aligned}$$

where $P_\eta^* = \{\boldsymbol{\eta}_{(q)}|(\boldsymbol{R}\boldsymbol{R}^t)^{1/2}\boldsymbol{\eta}_{(q)}\leqslant\boldsymbol{\alpha}\}$ and $\boldsymbol{z}_{(q)} = (\boldsymbol{R}\boldsymbol{R}^t)^{-1/2}\boldsymbol{R}\boldsymbol{x}$. Hence we get the expression (3.2) for $q < p$. □

Acknowledgments

The research of the first author was supported in part by a Grant-in-Aid for JSPS Fellows No. 1610018. The research of the second author was supported in part by grants from the Ministry of Education, Japan, Nos. 13680371, 15200021 and 15200022 and in part by a grant from the 21st Century COE Program at Faculty of Economics, University of Tokyo.

References

[1] R.E. Barlow, D.J. Bartholomew, J.M. Bremmer, H.D. Brunk, Statistical Inference Under Order Restrictions, Wiley, New York, 1972.
[2] J.C. Berry, Minimax estimation of a bounded normal mean vector, J. Multivariate Anal. 35 (1990) 130–139.
[3] L.D. Brown, J.T. Hwang, A unified admissibility proof, in: S.S. Gupta, J. Berger (Eds.), Statistical Decision Theory and Related Topics, vol. III, Academic Press, New York, 1982, pp. 205–230.
[4] G. Casella, W.E. Strawderman, Estimating a bounded mean, Ann. Statist. 9 (1981) 870–879.
[5] Y.-T. Chang, Stein-type estimators for parameters restricted by linear inequalities, Keio Sci. Tech. Rep. 34 (1981) 83–95.
[6] Y.-T. Chang, Stein-type estimators of parameters in truncated spaces, Keio Sci. Tech. Rep. 35 (1982) 185–193.
[7] R.H. Farrell, Estimators of a location parameter in the absolutely continuous case, Ann. Math. Statist. 35 (1964) 949–998.
[8] M.A. Girshick, L.J. Savage, Bayes and minimax estimates for quadratic loss functions, in: Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, University of California Press, Berkeley, 1951, pp. 53–74.
[9] J. Hartigan, Uniform priors on convex sets improve risk, Statist. Probab. Lett. 67 (2004) 285–288.
[10] J.T.G. Hwang, S.D. Peddada, Confidence interval estimation subject to order restrictions, Ann. Statist. 22 (1994) 67–93.
[11] W. James, C. Stein, Estimation with quadratic loss, in: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, University of California Press, Berkeley, 1961, pp. 361–379.
[12] S. Karlin, Y. Rinott, Classes of orderings of measures and related correlation inequalities, I. Multivariate totally positive distributions, J. Multivariate Anal. 10 (1980) 467–498.
[13] M. Katz, Admissible and minimax estimates of parameters in truncated spaces, Ann. Math. Statist. 32 (1961) 136–142.
[14] T. Kubokawa, Minimaxity in estimation of restricted parameters, J. Japan Statist. Soc. 34 (2004) 229–253.
[15] E.L. Lehmann, G. Casella, Theory of Point Estimation, second ed., Springer, New York, 1998.
[16] T. Robertson, F.T. Wright, R.L. Dykstra, Order Restricted Statistical Inference, Wiley, New York, 1988.
[17] D. Sengupta, P.K. Sen, Shrinkage estimation in a restricted parameter space, Sankhya A 53 (1991) 389–411.
[18] C. Stein, Estimation of the mean of a multivariate normal distribution, Ann. Statist. 9 (1981) 1135–1151.