CIRJE-F-786

A Unified Approach to Non-minimaxity of Sets of Linear Combinations of Restricted Location Estimators

Tatsuya Kubokawa
University of Tokyo

William E. Strawderman
Rutgers University

January 2011

A Unified Approach to Non-minimaxity of Sets of Linear Combinations of Restricted Location Estimators

Tatsuya Kubokawa *and William E. Strawderman†
University of Tokyo and *Rutgers University*

January 26, 2011

This paper studies minimaxity of estimators of a set of linear combinations of location parameters μ_i, $i = 1, \dots, k$ under quadratic loss. When each location parameter is known to be positive, previous results about minimaxity or non-minimaxity are extended from the case of estimating a single linear combination, to estimating any number of linear combinations. Necessary and/or sufficient conditions for minimaxity of general estimators are derived. Particular attention is paid to the generalized Bayes estimator with respect to the uniform distribution and to the truncated version of the unbiased estimator (which is the maximum likelihood estimator for symmetric unimodal distributions). A necessary and sufficient condition for minimaxity of the uniform prior generalized Bayes estimator is particularly simple; If one estimates $\boldsymbol{\theta} = \boldsymbol{A}\boldsymbol{\mu}$ where $\boldsymbol{A}$ is an $\ell \times k$ known matrix, the estimator is minimax if and only if $(\boldsymbol{A}\boldsymbol{A}^t)_{ij} \leq 0$ for any i and j, $(i \neq j)$. This condition is also sufficient (but not necessary) for minimaxity of the MLE.

Key words and phrases: Decision theory, generalized Bayes, linear combination, location parameter, location-scale family, maximum likelihood estimator, minimaxity, restricted parameter, restricted estimator, truncated estimator, quadratic loss.

1 Introduction

Estimation of restricted parameters has received much attention in the literature, and interesting studies have been developed from a decision-theoretic point of view since Katz (1961) and Farrell (1964). For recent developments, see Marchand and Strawderman (2004), Oono and Shinozaki (2005), van Eeden (2006) and Tsukuma and Kubokawa (2008). It is especially interesting to note that in the estimation of means of normal distributions, minimax properties of the uniform prior generalized Bayes estimator and

*Faculty of Economics, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan E-Mail: tatsuya@e.u-tokyo.ac.jp
†Department of Statistics, Rutgers University, Piscataway, New Jersey, 08854-8019, E-mail: straw@stat.rutgers.edu

the maximum likelihood estimator (MLE) in unrestricted estimation problems are not necessarily inherited in the restricted problems. One example of non-minimaxity is the case of estimating a bounded mean; Casella and Strawderman (1981) showed that the uniform prior generalized Bayes estimator is not minimax and Marchand and Perron (2001) demonstrated that minimaxity of the MLE is limitted. Another example is the case of estimating the sum of positively restricted means; Kubokawa (2010) showed the uniform prior generalized Bayes estimator is not minimax (if $k \geq 2$) and Kubokawa and Strawderman (2011) demonstrated that minimaxity of the MLE is limited. In this paper, we consider the extension of the results of these two papers to the simultaneous estimation of a set of linear combinations of means which are restricted to be positive, and derive necessary and/or sufficient conditions for minimaxity of general types of estimators. Quadratic loss is considered throughout the paper.

To explain instructively the problem treated here, consider the following problem: Let $X_1, \ldots, X_k$ be independent random variables such that X_i has a density $f_i(x - \mu_i)$ with a location parameter μ_i. Assume that $E[X_i^2] < \infty$ and that the location parameter μ_i is restricted to the positive real numbers $\{\mu_i > 0\}$ for $i = 1, \ldots, k$. Let $c_i = \int_{-\infty}^{\infty} z f_i(z) \mathrm{d}z$. Then, an unbiased estimator of μ_i is

$$\widehat{\mu}_i^U = X_i - c_i,$$

which is minimax under the squared error loss. We consider a set of linear combinations $\sum_{j=1}^k a_{ij}\mu_j$ for $i = 1, \ldots, \ell$, which are expressed as $\boldsymbol{a}_i^t \boldsymbol{\mu}$, where $\boldsymbol{a}_i^t = (a_{i1}, \ldots, a_{ik})$, $\boldsymbol{\mu}^t = (\mu_1, \ldots, \mu_k)$ and $\boldsymbol{\mu}^t$ denotes the transpose of $\boldsymbol{\mu}$. When $\boldsymbol{\mu}$ is restricted to $\Omega = \{\boldsymbol{\mu} | \mu_i > 0, i = 1, \ldots, k\}$, we want to estimate the set of the linear combinations

$$\boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_\ell \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_1^t \boldsymbol{\mu} \\ \vdots \\ \boldsymbol{a}_\ell^t \boldsymbol{\mu} \end{pmatrix} = \boldsymbol{A}^t \boldsymbol{\mu},$$

where $\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell)$, a $k \times \ell$ matrix.

In the case of $\ell = 1$, Kubokawa (2010) verified that the linear combination of unbiased estimators $\theta_1^U = \sum_{j=1}^k a_{1j} \widehat{\mu}_j^U$ is a minimax estimator with a constant risk. Then it was shown that the minimaxity of the linear combination of uniform prior generalized Bayes estimators $\hat{\theta}_1^{GB} = \sum_{j=1}^k a_{1j} \widehat{\mu}_j^{GB}$ is quite limited. Here,

$$\widehat{\mu}_j^{GB} = \frac{\int_0^\infty \mu_j f_j(X_j - \mu_j) \mathrm{d}\mu_j}{\int_0^\infty f_j(X_j - \mu_j) \mathrm{d}\mu_j} = X_j - \frac{\int_{-\infty}^{X_j} z f_j(z) \mathrm{d}z}{\int_{-\infty}^{X_j} f_j(z) \mathrm{d}z}. \tag{1.1}$$

In particular, $\hat{\theta}_1^{GB}$ is minimax for $k = 1$, but not minimax for $k \geq 3$. In the case of $k = 2$, it is minimax when $a_{11}a_{12} \leq 0$, but not minimax when $a_{11}a_{12} > 0$. Also in the case of $\ell = 1$, Kubokawa and Strawderman (2011) treated the truncated estimator

$$\widehat{\mu}_j^{TR} = \max\{\widehat{\mu}_j^U, 0\} = X_j - \min\{X_j, c_j\},$$

which is, as well, the MLE of μ_j for symmetric unimodal distributions and for some other distributions. Although $\widehat{\mu}_j^{TR}$ is minimax in estimation of the single location μ_j, it was shown that the minimaxity of the linear combination $\hat{\theta}_1^{TR} = \sum_{j=1}^{k} a_{1j}\widehat{\mu}_j^{TR}$ is also limited in the context of estimating θ_1. When $\ell = k$ and $\boldsymbol{A} = \boldsymbol{I}_k$, the indentity matrix, on the other hand, it can be verified that the uniform prior generalized Bayes estimator and the MLE are minimax. Thus, the problem treated in this paper fills in gaps between the above results for $\ell = 1$ and $\ell = k$.

In this paper we give a general necessary and sufficient condition for minimaxity of a general estimator of the form $\boldsymbol{A}^t\widehat{\boldsymbol{\mu}}$ where each $\widehat{\mu}_i(X_i)$ depends only on X_i. We show the condition is also sufficient when each $\widehat{\mu}_i$ is either the truncated estimator or the uniform prior generalized Bayes estimator on $(0, \infty)$. The condition takes on the very simple form, i.e., all off-diagonal elements of $\boldsymbol{A}\boldsymbol{A}^t$ are non-positive, for the uniform prior generalized Bayes estimator. This condition is also sufficient (but not necessary) for minimaxity of the truncated estimators as well. The sufficiency of the general necessary condition is also demonstrated for certain other minimax estimator $\widehat{\mu}_i$.

The paper is organized as follows: A general class of estimators, denoted by $\widehat{\boldsymbol{\theta}}_\phi$, of $\boldsymbol{\theta} = \boldsymbol{A}^t\boldsymbol{\mu}$ is handled throughout the paper. In Section 2, a general necessary condition (NC) and a sufficient condition (SC) for minimaxity of $\widehat{\boldsymbol{\theta}}_\phi$ are derived. Additionally, it is shown that the sufficient condition (SC) is also necessary for minimaxity of the uniform prior generalized Bayes estimator. Also, some examples of matrices $\boldsymbol{A}$ satisfying the sufficient condition (SC) are given.

In Section 3, a general condition is derived under which the necessary condition (NC) becomes sufficient for minimaxity of $\widehat{\boldsymbol{\theta}}_\phi$. In particular, it is shown that the truncated estimators (which are MLE for symmetric unimodal distributions) are governed by this result, and the necessary and sufficient condition for their minimaxity is given.

In Section 4, we consider the specific case wherein the underlying distributions are normal. In Section 4.1, we provide a unified approach to necessary and sufficient conditions for minimaxity of the uniform prior generalized Bayes estimator and the MLE. In Section 4.2, we extend the results to the unknown variance case and show that similar dominance results hold. Finally, the proof of minimaxity of the unrestricted generalized Bayes estimator of $\boldsymbol{\theta}$ is given for a location-scale family. In particular, this implies that the unbiased estimator of $\boldsymbol{\theta}$ is minimax in normal distributions with a common unknown variance.

2 Conditions for Minimaxity and Non-minimaxity

2.1 A necessary condition for minimaxity

Consider the problem of estimating a vector of linear combinations $\boldsymbol{\theta} = \boldsymbol{A}^t\boldsymbol{\mu}$ relative to the squared error loss function $\|\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|^2$ for an estimator $\widehat{\boldsymbol{\theta}}$ where $\|\boldsymbol{u}\|^2 = \sum_{i=1}^{\ell} u_i^2$ for $\boldsymbol{u} = (u_1, \ldots, u_\ell)^t$. Let $\boldsymbol{X} = (X_1, \ldots, X_k)^t$ and $\boldsymbol{c} = (c_1, \ldots, c_k)^t$. An unbiased estimator of $\boldsymbol{\theta}$ is $\widehat{\boldsymbol{\theta}}^U = \boldsymbol{A}^t\widehat{\boldsymbol{\mu}}^U$ for

$$\widehat{\boldsymbol{\mu}}^U = (\widehat{\mu}_1^U, \ldots, \widehat{\mu}_k^U)^t = \boldsymbol{X} - \boldsymbol{c},$$

and the same arguments as in Kubokawa (2010) can be used to show that $\widehat{\boldsymbol{\theta}}^U$ is a minimax estimator with a constant risk. See also the proof given in the appendix.

We begin by considering a general type of estimator

$$\widehat{\mu}_{\phi,j} = X_j - \phi_j(X_j)$$

of μ_j for $j = 1, \ldots, k$, where $E[\phi_j^2(X_j)] < \infty$. Let

$$\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}} = \boldsymbol{A}^t \widehat{\boldsymbol{\mu}}_{\boldsymbol{\phi}}, \quad \text{for} \quad \widehat{\boldsymbol{\mu}}_{\boldsymbol{\phi}} = (\widehat{\mu}_{\phi,1}, \ldots, \widehat{\mu}_{\phi,k})^t. \tag{2.1}$$

Then the risk function of the estimator $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}}$ is written as

$$\begin{aligned} R(\boldsymbol{\mu}, \widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}}) =& E[\|\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}} - \boldsymbol{\theta}\|^2] = E[(\widehat{\boldsymbol{\mu}}_{\boldsymbol{\phi}} - \boldsymbol{\mu})^t \boldsymbol{A}\boldsymbol{A}^t (\widehat{\boldsymbol{\mu}}_{\boldsymbol{\phi}} - \boldsymbol{\mu})] \\ =& \sum_{i=1}^k (\boldsymbol{A}\boldsymbol{A}^t)_{ii} R_i(\mu_i) + \sum_{i=1}^k \sum_{j=1, j\neq i}^k (\boldsymbol{A}\boldsymbol{A}^t)_{ij} B_i(\mu_i) B_j(\mu_j), \end{aligned}$$

where $(\boldsymbol{A}\boldsymbol{A}^t)_{ij}$ denotes the (i,j)-th element of $\boldsymbol{A}\boldsymbol{A}^t$, and $R_i(\mu_i)$ and $B_i(\mu_i)$ are, respectively, the risk function and the bias of the estimator $\widehat{\mu}_{\phi,i} = X_i - \phi_i(X_i)$, given by

$$\begin{aligned} B_i(\mu_i) =& E[\widehat{\mu}_{\phi,i} - \mu_i] = E[X_i - \phi_i(X_i) - \mu_i], \\ R_i(\mu_i) =& E[(\widehat{\mu}_{\phi,i} - \mu_i)^2] = E[(X_i - \phi_i(X_i) - \mu_i)^2]. \end{aligned}$$

These can be also expressed as

$$\begin{aligned} B_i(\mu_i) =& E[c_i - \phi_i(X_i)], \\ R_i(\mu_i) =& E[(X_i - c_i - \mu_i)^2] - D_i(\mu_i), \end{aligned} \tag{2.2}$$

where the risk difference can be expressed as

$$D_i(\mu_i) = E[\{c_i - \phi_i(X_i)\}\{c_i + \phi_i(X_i) - 2(X_i - \mu_i)\}]. \tag{2.3}$$

Then, the difference between the risk functions of $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}}$ and the minimax estimator $\widehat{\boldsymbol{\theta}}^U$ is

$$\Delta(\boldsymbol{\mu}) = -\sum_{i=1}^k (\boldsymbol{A}\boldsymbol{A}^t)_{ii} D_i(\mu_i) + \sum_{i=1}^k \sum_{j=1, j\neq i}^k (\boldsymbol{A}\boldsymbol{A}^t)_{ij} B_i(\mu_i) B_j(\mu_j), \tag{2.4}$$

and it is seen that the minimaxity of $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}}$ is equivalent to $\Delta(\boldsymbol{\mu}) \leq 0$.

We first derive a necessary condition for the minimaxity of $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}}$. For this purpose, we assume that $\lim_{\mu_i \to \infty} B_i(\mu_i) = 0$ and $\lim_{\mu_i \to \infty} D_i(\mu_i) = 0$. As shown below, this assumption can be guaranteed when $\phi_i(w)$ converges to c_i as $w \to \infty$. Let $\Lambda = \{1, \ldots, k\}$ and let C be any subset of Λ. If $\mu_i \to 0$ for all $i \in C$, and if $\mu_j \to \infty$ for all $j \in \Lambda \backslash C$, then the risk difference $\Delta(\boldsymbol{\mu})$ converges to

$$-\sum_{i \in C} (\boldsymbol{A}\boldsymbol{A}^t)_{ii} D_i(0) + \sum_{i \in C} \sum_{j \in C, j \neq i} (\boldsymbol{A}\boldsymbol{A}^t)_{ij} B_i(0) B_j(0).$$

Proposition 2.1 *Assume that* $\lim_{\mu_i\to\infty} B_i(\mu_i) = 0$ *and* $\lim_{\mu_i\to\infty} D_i(\mu_i) = 0$. *If the estimator* $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}}$ *is minimax, then for all subset* C *of* $\{1, \ldots, k\}$,

$$(\mathbf{NC}) \ : \qquad -\sum_{i\in C}(\boldsymbol{A}\boldsymbol{A}^t)_{ii}D_i(0) + \sum_{i\in C}\sum_{j\in C, j\neq i}(\boldsymbol{A}\boldsymbol{A}^t)_{ij}B_i(0)B_j(0) \le 0. \tag{2.5}$$

The assumption given in Proposition 2.1 is satisfied if the function $\phi_i(w)$ satisfies the following conditions: (i) $\lim_{w\to\infty}\phi_i(w) = c_i$ and (ii) there exists a function $\Phi_i(z)$ such that $\sup_{\mu_i>0}|\phi_i(z+\mu_i)| \le \Phi_i(z)$, $\Phi_i(z)$ is independent of μ_i, and $E[|\Phi_i(Z)|^2] < \infty$, where Z has a density $f_i(z)$. In fact, using the Lebesgue's dominated convergence theorem, we can see that $\lim_{\mu_i\to\infty} B_i(\mu_i) = c_i - \lim_{\mu_i\to\infty} E[\phi_i(Z+\mu_i)] = 0$ and $\lim_{\mu_i\to\infty} D_i(\mu_i) = \lim_{\mu_i\to\infty} E[\{c_i - \phi_i(Z+\mu_i)\}\{c_i + \phi_i(Z+\mu_i) - 2Z\}] = 0$.

If $\phi_i(w)$ satisfies the condition

(A1) $\phi_i(w)$ is nondecreasing and $\lim_{w\to\infty}\phi_i(w) = c_i$ for $i = 1, \ldots, k$,

then it can be seen that $\sup_{\mu_i>0}|\phi_i(z+\mu_i)| \le |\phi_i(z)| + |c_i|$, so that the above condition (ii) is satisfied by $E[|\phi_i(Z)|^2] < \infty$. That is, if $\phi_i(w)$ satisfies the condition (A1), then the assumption of Proposition 2.1 holds, namely, $B_i(\mu_i) \ge 0$, $\lim_{\mu_i\to\infty} B_i(\mu_i) = 0$ and $\lim_{\mu_i\to\infty} D_i(\mu_i) = 0$.

2.2 A sufficient condition for minimaxity

To get sufficient conditions for minimaxity, we need to find conditions such that $\Delta(\boldsymbol{\mu}) \le 0$ for any $\boldsymbol{\mu}$. In this subsection we derive a general sufficient condition and show that it is also a necessary condition for minimaxity of the uniform prior generalized Bayes estimator.

[1] A general sufficient condition. If $D_i(\mu_i) \ge 0$ and $B_i(\mu_i) \ge 0$ for $\mu_i > 0$ and $i = 1, \ldots, k$, then it follow from (2.4) that $\Delta(\boldsymbol{\mu}) \le 0$ if all off-diagonal elements of $\boldsymbol{A}\boldsymbol{A}^t$ satisfy the condition $(\boldsymbol{A}\boldsymbol{A}^t)_{ij} \le 0$ for all $i,\ j\ (i \neq j)$.

Proposition 2.2 *Assume the following conditions:*

$$(\mathbf{SC}) \ : \ \begin{cases} (\mathbf{SC1}) & D_i(\mu_i) \ge 0 \ \textit{and} \ B_i(\mu_i) \ge 0 \ \textit{for any} \ \mu_i > 0 \ \textit{and} \ i = 1, \ldots, k \\ (\mathbf{SC2}) & (\boldsymbol{A}\boldsymbol{A}^t)_{ij} \le 0 \ \textit{for all} \ i, j \ (i \neq j). \end{cases}$$

Then, the estimator $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}}$ *is minimax.*

It is noted that $D_i(\mu_i) \ge 0$ for any $\mu_i > 0$ implies that the estimator $\widehat{\mu}_{\phi,i}$ dominates $\widehat{\mu}_i^U$, or it is minimax. As verified in Kubokawa (2010), the estimator $\widehat{\mu}_{\phi,i} = X_i - \phi_i(X_i)$ is minimax if $\phi_i(w)$ satisfies the condition (A1) and the condition given by

(A2) $\phi_i(w) \ge \phi_i^{GB}(w)$ for $i = 1, \ldots, k$, where

$$\phi_i^{GB}(w) = \int_{-\infty}^{w} z f_i(z)\mathrm{d}z \Big/ \int_{-\infty}^{w} f_i(z)\mathrm{d}z. \tag{2.6}$$

The conditions (A1) and (A2) imply that $D_i(\mu_i) \ge 0$ and $B_i(\mu_i) \ge 0$ for any $\mu_i > 0$ and $i = 1, \ldots, k$, so that the condition (SC2) is sufficient for the minimaxity.

[2] The uniform prior generalized Bayes estimator and the necessary and sufficient condition. In general, the sufficient conditions (SC1) and (SC2) are not necessary for minimaxity. However, it is interesting to note that the condition (SC2) is necessary and sufficient for minimaxity of the uniform prior generalized Bayes estimator

$$\widehat{\boldsymbol{\theta}}^{GB} = \boldsymbol{A}^t \widehat{\boldsymbol{\mu}}^{GB}, \quad \text{for} \quad \widehat{\boldsymbol{\mu}}^{GB} = (\widehat{\mu}_1^{GB}, \ldots, \widehat{\mu}_k^{GB})^t,$$

where $\widehat{\mu}_i^{GB}$ is given in (1.1). In fact, note that $\widehat{\mu}_i^{GB}$ may be expressed as

$$\widehat{\mu}_i^{GB} = X_i - \phi_i^{GB}(X_i),$$

and that $\phi_i^{GB}(w)$ satisfies the conditions (A1) and (A2). Also, note that the risk function of the uniform prior generalized Bayes estimator $\widehat{\mu}_i^{GB}$ attains the constant minimax risk at $\mu_i = 0$ as verified in Kubokawa (2010), namely, in (2.2) and (2.3),

$$R_i(0) = E[(X_i - c_i - \mu_i)^2], \quad \text{or} \quad D_i(0) = 0,$$

for $\phi_i(w) = \phi_i^{GB}(w)$. Thus, the necessary condition (2.5) becomes that

$$\sum_{i \in C} \sum_{j \in C, j \neq i} (\boldsymbol{A}\boldsymbol{A}^t)_{ij} B_i(0) B_j(0) \leq 0,$$

for all subset C of $\{1, \ldots, k\}$. Since $B_i(0) > 0$ for $i = 1, \ldots, k$, it can be seen that this necessary condition is reduced to the condition (SC2).

Proposition 2.3 *A necessary and sufficient condition for the uniform prior generalized Bayes estimator* $\widehat{\boldsymbol{\theta}}^{GB}$ *to be minimax is that all off-diagonal elements of* $\boldsymbol{A}\boldsymbol{A}^t$ *satisfy the condition* (SC2).

[3] Examples of matrix $\boldsymbol{A}$ satisfying (SC2). We here investigate when the condition (SC2) is satisfied through some examples.

Example 2.1 (Case of $k = 2$) Let $\boldsymbol{a}_i = (a_{i1}, a_{i2})^t$ for $i = 1, 2$. In the case of $\ell = 1$, we have $(\boldsymbol{A}\boldsymbol{A}^t)_{12} = (\boldsymbol{a}_1\boldsymbol{a}_1^t)_{12} = a_{11}a_{12}$. Then from Proposition 2.3, it follows that the uniform prior generalized Bayes estimator $\widehat{\boldsymbol{\theta}}^{GB}$ is minimax if and only if $a_{11}a_{12} \leq 0$. For example, it is minimax for $\boldsymbol{a}_1^t = (1, -1)$, but not for $\boldsymbol{a}_1^t = (1, 1)$. This corresponds to the result in Kubokawa (2010).

In the case of $\ell = 2$, we have $(\boldsymbol{A}\boldsymbol{A}^t)_{12} = (\boldsymbol{a}_1\boldsymbol{a}_1^t + \boldsymbol{a}_1\boldsymbol{a}_1^t)_{12} = a_{11}a_{12} + a_{21}a_{22}$, so that $\widehat{\boldsymbol{\theta}}^{GB}$ is minimax if and only if $a_{11}a_{12} + a_{21}a_{22} \leq 0$. Since $a_{11}a_{12} + a_{21}a_{22} = (a_{11}, a_{21})(a_{12}, a_{22})^t$, it is convenient to express $\boldsymbol{A}$ as

$$\boldsymbol{A}^t = \begin{pmatrix} \boldsymbol{a}_1^t \\ \boldsymbol{a}_2^t \end{pmatrix} = \left(\boldsymbol{a}_{(1)}, \boldsymbol{a}_{(2)}\right)$$

for $\boldsymbol{a}_{(1)} = (a_{11}, a_{21})^t$ and $\boldsymbol{a}_{(2)} = (a_{12}, a_{22})^t$. Then, the necessary and sufficient condition is equivalent to $\boldsymbol{a}_{(1)}^t \boldsymbol{a}_{(2)} \leq 0$. For example, let $\boldsymbol{a}_1^t = (1, 1)$. Then, $\widehat{\boldsymbol{\theta}}^{GB}$ is minimax for $\boldsymbol{a}_2^t = (1, -1)$, $(1, -2)$, but not minimax for $\boldsymbol{a}_2^t = (1, -1/2)$.

Example 2.2 (Case of $k = 3$) Let $\boldsymbol{a}_i = (a_{i1}, a_{i2}, a_{i3})^t$ for $i = 1, 2, 3$. In the case of $\ell = 1$, $\widehat{\boldsymbol{\theta}}^{GB}$ is minimax if and only if $(\boldsymbol{A}\boldsymbol{A}^t)_{12} = (\boldsymbol{a}_1\boldsymbol{a}_1^t)_{12} = a_{11}a_{12} \leq 0$, $(\boldsymbol{A}\boldsymbol{A}^t)_{13} = a_{11}a_{13} \leq 0$ and $(\boldsymbol{A}\boldsymbol{A}^t)_{23} = a_{12}a_{13} \leq 0$. There is no solution of non-zero a_{i1}, a_{i2} and a_{i3} satisfying these inequalities. This implies that $\widehat{\boldsymbol{\theta}}^{GB}$ is not minimax in the case of $\ell = 1$ and $k = 3$. This corresponds to the result of Kubokawa (2010).

In the case of $\ell = 2$, we have

$$\boldsymbol{A}^t = \begin{pmatrix} \boldsymbol{a}_1^t \\ \boldsymbol{a}_2^t \end{pmatrix} = \big(\boldsymbol{a}_{(1)}, \boldsymbol{a}_{(2)}, \boldsymbol{a}_{(3)})\big)$$

for $\boldsymbol{a}_{(1)} = (a_{11}, a_{21})^t$, $\boldsymbol{a}_{(2)} = (a_{12}, a_{22})^t$ and $\boldsymbol{a}_{(3)} = (a_{13}, a_{23})^t$. Then, the necessary and sufficient condition is $\boldsymbol{a}_{(1)}^t\boldsymbol{a}_{(2)} \leq 0$, $\boldsymbol{a}_{(2)}^t\boldsymbol{a}_{(3)} \leq 0$ and $\boldsymbol{a}_{(3)}^t\boldsymbol{a}_{(1)} \leq 0$. For example, $\widehat{\boldsymbol{\theta}}^{GB}$ is minimax for

$$\boldsymbol{A}^t = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix},$$

but not minimax for

$$\boldsymbol{A}^t = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}.$$

In the case of $\ell = 3$, we have

$$\boldsymbol{A}^t = \begin{pmatrix} \boldsymbol{a}_1^t \\ \boldsymbol{a}_2^t \\ \boldsymbol{a}_3^t \end{pmatrix} = \big(\boldsymbol{a}_{(1)}, \boldsymbol{a}_{(2)}, \boldsymbol{a}_{(3)})\big)$$

for $\boldsymbol{a}_{(1)} = (a_{11}, a_{21}, a_{31})^t$, $\boldsymbol{a}_{(2)} = (a_{12}, a_{22}, a_{32})^t$ and $\boldsymbol{a}_{(3)} = (a_{13}, a_{23}, a_{33})^t$. Then, the necessary and sufficient condition is $\boldsymbol{a}_{(1)}^t\boldsymbol{a}_{(2)} \leq 0$, $\boldsymbol{a}_{(2)}^t\boldsymbol{a}_{(3)} \leq 0$ and $\boldsymbol{a}_{(3)}^t\boldsymbol{a}_{(1)} \leq 0$. For example, let $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ be orthonormal vectors. Then, $\boldsymbol{A}$ becomes an orthogonal matrix, which means that $\{\boldsymbol{a}_{(1)}, \boldsymbol{a}_{(2)}, \boldsymbol{a}_{(3)}\}$ are orthonormal vectors and the mutual inner products are zero. Thus, the condition (SC2) is satisfied and the minimaxity of $\widehat{\boldsymbol{\theta}}^{GB}$ is established.

Example 2.3 (General cases) For $k \geq 4$ and $\ell \geq 1$, let

$$\boldsymbol{A}^t = \begin{pmatrix} \boldsymbol{a}_1^t \\ \vdots \\ \boldsymbol{a}_\ell^t \end{pmatrix} = \big(\boldsymbol{a}_{(1)}, \ldots, \boldsymbol{a}_{(k)})\big).$$

Then, $\widehat{\boldsymbol{\theta}}^{GB}$ is minimax if and only if $\boldsymbol{a}_{(i)}^t\boldsymbol{a}_{(j)} \leq 0$ for all i, j $(i \neq j)$. In the case of $\ell = 1$, such vectors $\boldsymbol{a}_{(i)}$'s do not exist, and it is not minimax. In the case that $\ell = k$ and $\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell)$ is an orthogonal matrix, the condition (SC2) is satisfied, and $\widehat{\boldsymbol{\theta}}^{GB}$ is minimax.

3 Necessary and Sufficient Conditions for Minimaxity of General Estimators

3.1 Sufficiency of the necessary condition (NC)

We here consider the interesting question of whether the necessary condition (NC) given in (2.5) is sufficient or not. To answer the question, we need to show that the risk function attains its maximum on the boundary of the parameter space. Differentiating $\Delta(\boldsymbol{\mu})$ with respect to μ_i, we see from (2.4) that

$$\frac{\partial}{\partial \mu_i}\Delta(\boldsymbol{\mu}) = -(\boldsymbol{A}\boldsymbol{A}^t)_{ii}\frac{\partial D_i(\mu_i)}{\partial \mu_i} + 2\frac{\partial B_i(\mu_i)}{\partial \mu_i}\sum_{j=1,j\neq i}^{k}(\boldsymbol{A}\boldsymbol{A}^t)_{ij}B_j(\mu_j). \tag{3.1}$$

It is here noted that $B_i(\mu_i)$ and $D_i(\mu_i)$ can rewritted as

$$\begin{aligned} B_i(\mu_i) =& E[c_i - \phi_i(Z_i+\mu_i)], \\ D_i(\mu_i) =& E[\{c_i - \phi_i(Z_i+\mu_i)\}\{c_i + \phi_i(Z_i+\mu_i) - 2Z_i\}], \end{aligned}$$

for $Z_i = X_i - \mu_i$. Differentiating these functions with respect to μ_i gives the expressions

$$\begin{aligned} \frac{\partial B_i(\mu_i)}{\partial \mu_i} =& -E[\phi_i'(X_i)], \\ \frac{\partial D_i(\mu_i)}{\partial \mu_i} =& 2E[\phi_i'(X_i)\{X_i - \mu_i - \phi_i(X_i)\}], \end{aligned}$$

where $\phi_i'(w) = (\mathrm{d}/\mathrm{d}w)\phi_i(w)$. Thus, the partial derivative of $\Delta(\boldsymbol{\mu})$ is expressed as

$$\begin{aligned} \frac{\partial}{\partial \mu_i}\Delta(\boldsymbol{\mu}) =& 2(\boldsymbol{A}\boldsymbol{A}^t)_{ii}E[\phi_i'(X_i)\{\phi_i(X_i) - (X_i-\mu_i)\}] - 2E[\phi_i'(X_i)]\sum_{j=1,j\neq i}^{k}(\boldsymbol{A}\boldsymbol{A}^t)_{ij}B_j(\mu_j) \\ =& 2E[\phi_i'(X_i)]\Big\{(\boldsymbol{A}\boldsymbol{A}^t)_{ii}H_i(\mu_i) - \sum_{j=1,j\neq i}^{k}(\boldsymbol{A}\boldsymbol{A}^t)_{ij}B_j(\mu_j)\Big\}, \end{aligned} \tag{3.2}$$

where $H_i(\mu_i)$ is

$$H_i(\mu_i) = \frac{E[\phi_i'(X_i)\{\phi_i(X_i) - (X_i-\mu_i)\}]}{E[\phi_i'(X_i)]}. \tag{3.3}$$

It is noted that only $E[\phi_i'(X_i)]$ and $H_i(\mu_i)$ depend on μ_i while all other terms are independent of μ_i in (3.2). Assuming that $E[\phi_i'(X_i)] > 0$, we see that if $H_i(\mu_i)$ is nondecreasing in μ_i, then we can consider the three cases: (1) $(\partial/\partial\mu_i)\Delta(\boldsymbol{\mu}) \geq 0$ for all $\mu_i > 0$, (2) $(\partial/\partial\mu_i)\Delta(\boldsymbol{\mu}) \leq 0$ for all $\mu_i > 0$, or (3) there is a positive point $\mu_{i,0}$ such that $(\partial/\partial\mu_i)\Delta(\boldsymbol{\mu}) < 0$ for $0 < \mu_i < \mu_{i,0}$, and $(\partial/\partial\mu_i)\Delta(\boldsymbol{\mu}) \geq 0$ for all $\mu_i \geq \mu_{i,0}$. This implies that

$$\Delta(\boldsymbol{\mu}) \leq \max\Big\{\lim_{\mu_i\to 0}\Delta(\boldsymbol{\mu}),\ \lim_{\mu_i\to\infty}\Delta(\boldsymbol{\mu})\Big\}.$$

Applying this argument for all $i \in \{1,\dots,k\}$ and assuming that $\lim_{\mu_i\to\infty} B_i(\mu_i) = 0$ and $\lim_{\mu_i\to\infty} D_i(\mu_i) = 0$, we see that

$$\Delta(\boldsymbol{\mu}) \le \max_C \Big\{ -\sum_{i\in C} (\boldsymbol{A}\boldsymbol{A}^t)_{ii} D_i(0) + \sum_{i\in C} \sum_{j\in C, j\ne i} (\boldsymbol{A}\boldsymbol{A}^t)_{ij} B_i(0) B_j(0) \Big\}$$

where C is all subsets of $\{1,\dots,k\}$. This implies that the sufficient condition leads to the necessary condition (2.5). Hence, we have the following result.

Proposition 3.1 *Assume that* $E[\phi_i'(X_i)] > 0$ *for* $i = 1,\dots,k$ *as well as* $\lim_{\mu_i\to\infty} B_i(\mu_i) = 0$ *and* $\lim_{\mu_i\to\infty} D_i(\mu_i) = 0$*. If* $H_i(\mu_i)$ *given in* (3.3) *is nondecreasing in* μ_i *for* $i = 1,\dots,k$*, then the condition* (2.5) *is a necessary and sufficient condition for the estimator* $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}}$ *to be minimax.*

As verified below Proposition 2.1, the condition (A1) implies that $E[\phi_i'(X_i)] > 0$ for $i = 1,\dots,k$ and that $\lim_{\mu_i\to\infty} B_i(\mu_i) = 0$ and $\lim_{\mu_i\to\infty} D_i(\mu_i) = 0$, namely, the assumption of Proposition 3.1 is satisfied.

Although this proposition provides a nice necessary and sufficient condition, it may be hard to check the monotonicity of the function $H_i(\mu_i)$. For specific cases of estimator and distribution, however, we can verify this monotonicity. The following proposition give us a condition on $\phi_i(\cdot)$ which implies the monotonicity of $H_i(\mu_i)$.

Proposition 3.2 *Assume the condition* (A1)*. If* $\phi_i'(z+u_2)/\phi_i'(z+u_1)$ *and* $\phi_i(z) - z$ *are nonincreasing in* z *on* $\{z|\phi_i'(z+u_1) > 0\}$ *for* $0 < u_1 < u_2$*, then* $H_i(\mu_i)$ *is nondecreasing in* μ_i*, and the condition* (2.5) *is necessary and sufficient.*

Proof. We omit the index i in this proof. For $0 < u_1 < u_2$, we need to show that

$$\frac{E[\phi'(Z+u_1)\{\phi(Z+u_1) - Z\}]}{E[\phi'(Z+u_1)]} \le \frac{E[\phi'(Z+u_2)\{\phi(Z+u_2) - Z\}]}{E[\phi'(Z+u_2)]},$$

which holds if

$$\frac{E[\phi'(Z+u_1)\{\phi(Z+u_1) - Z\}]}{E[\phi'(Z+u_1)]} \le \frac{E[\phi'(Z+u_2)\{\phi(Z+u_1) - Z\}]}{E[\phi'(Z+u_2)]}.$$

Since $\phi_i'(Z+u_2)/\phi_i'(Z+u_1)$ and $\phi_i(Z+u_1) - Z$ are nonincreasing in Z, this inequality holds. ∎

When $\phi_i'(z)$ is differentiable, it is noted that $\phi_i'(z+u_2)/\phi_i'(z+u_1)$ is nonincreasing in z for $0 < u_1 < u_2$ if $\phi_i''(z)/\phi_i(z)$ is nonincreasing in z.

Remark 3.1 In the general setup, it would be interesting if we could show that $H_i(\mu_i)$ is nondecreasing in μ_i for the uniform prior generalized Bayes estimators. Unfortunately we have not been able to show this. However, when the distributions are normal, it will be verified in Section 4.1 that the uniform prior generalized Bayes estimator as well as the MLE lead to monotonicity of $H_i(\mu_i)$.

3.2 Truncated estimators and the necessary and sufficient condition

In this subsection we study minimaxity of truncated estimators. The truncated estimator of μ_i is

$$\widehat{\mu}_i^{TR} = \max\{X_i - c_i, 0\} = X_i - \phi_i^{TR}(X_i),$$

for $\phi_i^{TR}(X_i) = \min\{X_i, c_i\}$. This is the MLE of μ_i in a symmetric unimodal distribution. More generally, we can consider the truncated estimators

$$\widehat{\mu}_i^{TR\gamma} = \max\{X_i - c_i, (1-\gamma)(X_i - c_i)\} = X_i - \phi_i^{TR\gamma}(X_i),$$

where $\gamma > 0$ and

$$\phi_i^{TR\gamma}(X_i) = \min\{\gamma(X_i - c_i) + c_i, c_i\} = \begin{cases} \gamma(X_i - c_i) + c_i & \text{for } X_i < c_i, \\ c_i & \text{for } X_i \geq c_i. \end{cases}$$

For the condition (A2), it is noted that the function $\phi_i^{GB}(w)$ given in (2.6) is increasing in w and $\lim_{w\to\infty} \phi^{GB}(w) = c_i$, and that $\phi_i^{GB}(w) \leq \min\{w, c_i\} \leq \min\{\gamma(w - c_i) + c_i, c_i\}$ for $0 < \gamma \leq 1$. Thus, the function $\phi_i^{TR\gamma}(w)$ satisfies the conditions (A1) and (A2) for $0 < \gamma \leq 1$, so that from Kubokawa (2010), the resulting truncated estimator $\widehat{\mu}_i^{TR\gamma}$ dominates X_i, namely, it is minimax for $0 < \gamma \leq 1$ in the context of estimation of μ_i. However, Kubokawa (2010)'s result can not be used to extend this dominance result to the case of $0 < \gamma \leq 2$.

To show directly the dominance result for $0 < \gamma \leq 2$, it is noted that $\widehat{\mu}_i^{TR\gamma} = X_i - c_i - \gamma(X_i - c_i)I(X_i < c_i)$ for the indicator function $I(\cdot)$, so that the bias and the risk of $\widehat{\mu}_i^{TR\gamma}$ are written as

$$\begin{aligned} B_i(\mu_i) =& E[X_i - c_i - \mu_i - \gamma(X_i - c_i)I(X_i < c_i)] \\ =& -\gamma E[(X_i - c_i)I(X_i < c_i)], \\ R_i(\mu_i) =& E[\{X_i - c_i - \mu_i - \gamma(X_i - c_i)I(X_i < c_i)\}^2] \\ =& E[(X_i - c_i - \mu_i)^2] - D_i(\mu_i), \end{aligned} \tag{3.4}$$

where $D_i(\mu_i) = \gamma E[\{2(X_i - c_i - \mu_i) - \gamma(X_i - c_i)\}(X_i - c_i)I(X_i < c_i)]$. Letting $F_i(w) = \int_{-\infty}^{w} f_i(z)\mathrm{d}z$ and $z = x_i - \mu_i$, we can rewrite them as

$$\begin{aligned} B_i(\mu_i) =& -\gamma \int_{-\infty}^{c_i-\mu_i} (z - c_i + \mu_i) f_i(z)\mathrm{d}z = \gamma \int_{-\infty}^{c_i-\mu_i} F_i(z)\mathrm{d}z, \\ D_i(\mu_i) =& \gamma(2-\gamma) \int_{-\infty}^{c_i-\mu_i} (z - c_i)(z - c_i + \mu_i) f_i(z)\mathrm{d}z + \gamma\mu_i B_i(\mu_i) \\ =& 2\gamma\Big\{-(2-\gamma) \int_{-\infty}^{c_i-\mu_i} (z - c_i) F_i(z)\mathrm{d}z + (\gamma - 1)\mu_i \int_{-\infty}^{c_i-\mu_i} F_i(z)\mathrm{d}z\Big\}, \end{aligned}$$

since $\int_{-\infty}^{c_i-\mu_i} (z - c_i)(z - c_i + \mu_i) f_i(z)\mathrm{d}z = -2\int_{-\infty}^{c_i-\mu_i} (z - c_i) F_i(z)\mathrm{d}z - \mu_i \int_{-\infty}^{c_i-\mu_i} F_i(z)\mathrm{d}z$ as shown in Kubokawa and Strawderman (2011). Then, for $\mu_i = 0$, we have

$$
\begin{aligned}
B_i(0) =& -\gamma \int_{-\infty}^{c_i} (z - c_i) f_i(z) \mathrm{d}z = -\gamma \int_{-\infty}^{c_i} F_i(z) \mathrm{d}z, \\
D_i(0) =& \gamma(2-\gamma) \int_{-\infty}^{c_i} (z - c_i)^2 f_i(z) \mathrm{d}z \\
=& -2\gamma(2-\gamma) \int_{-\infty}^{c_i} (z - c_i) F_i(z) \mathrm{d}z.
\end{aligned} \tag{3.5}
$$

Proposition 3.3 *The truncated estimator* $\widehat{\mu}_i^{TR\gamma}$ *always dominates* $X_i - c_i$ *for* $0 < \gamma \leq 1$, *while this dominance result holds for* $1 < \gamma \leq 2$ *if* $f'(z)/f(z)$ *is non-increasing in* z. *When* $\gamma > 2$, *however,* $\widehat{\mu}_i^{TR\gamma}$ *does not dominate* $X_i - c_i$.

Proof. For simplicity, we omit the index i in X_i, μ_i, $D_i(\cdot)$, c_i and others. It is noted that the dominance of $\widehat{\mu}^{TR\gamma}$ over $X - c$ is equivalent to the inequality $D(\mu) \geq 0$ for any $\mu > 0$. Also, note that $\lim_{\mu\to\infty} B(\mu) = \lim_{\mu\to\infty} D(\mu) = 0$. When $\gamma > 2$, it follows from (3.5) that $D(0) < 0$, which means that $\widehat{\mu}^{TR\gamma}$ does not dominate $X - c$. Then, we shall establish the dominance property when $0 < \gamma \leq 2$. Differentiating $B(\mu)$ and $D(\mu)$ given in (3.4) with respect to μ gives $B'(\mu) = -\gamma \int_{-\infty}^{c-\mu} f(z)\mathrm{d}z$ and

$$
\begin{aligned}
D'(\mu) =& \gamma(2-\gamma) \int_{-\infty}^{c-\mu} (z - c) f(z) \mathrm{d}z + \gamma B(\mu) + \gamma\mu B'(\mu) \\
=& 2\gamma \Big\{ (1-\gamma) \int_{-\infty}^{c-\mu} (z - c + \mu) f(z) \mathrm{d}z - \mu \int_{-\infty}^{c-\mu} f(z) \mathrm{d}z \Big\}.
\end{aligned} \tag{3.6}
$$

Since $\int_{-\infty}^{c-\mu} (z - c + \mu) f(z)\mathrm{d}z \leq 0$, it is seen that $D'(\mu) \leq 0$ for $\mu > 0$ when $0 < \gamma \leq 1$, so that $D(\mu)$ is non-increasing in μ. Thus, $D(\mu) \geq 0$ for all $\mu > 0$ since $D(0) > 0$ and $\lim_{\mu\to\infty} D(\mu) = 0$. When $\gamma > 1$, making the transformation $x = z + \mu$, we rewrite $D'(\mu)$ as

$$
D'(\mu) = 2\gamma \int_{-\infty}^{c} f(x - \mu)\mathrm{d}x \Big\{ (1-\gamma) \frac{\int_{-\infty}^{c} x f(x-\mu)\mathrm{d}x}{\int_{-\infty}^{c} f(x-\mu)\mathrm{d}x} - c(1-\gamma) - \mu \Big\}.
$$

We here show that $\int_{-\infty}^{c} x f(x-\mu)\mathrm{d}x / \int_{-\infty}^{c} f(x-\mu)\mathrm{d}x$ is non-decreasing in μ if $f'(z)/f(z)$ is non-increasing in z. In fact, differentiating the function with respect to μ, we observe that the derivative is proportional to

$$
-\frac{\int_{-\infty}^{c} x f'(x-\mu)\mathrm{d}x}{\int_{-\infty}^{c} f(x-\mu)\mathrm{d}x} + \frac{\int_{-\infty}^{c} x f(x-\mu)\mathrm{d}x}{\int_{-\infty}^{c} f(x-\mu)\mathrm{d}x} \frac{\int_{-\infty}^{c} f'(x-\mu)\mathrm{d}x}{\int_{-\infty}^{c} f(x-\mu)\mathrm{d}x}
$$

which is non-negative if $f'(x-\mu)/f(x-\mu)$ is non-increasing in x. Since $(1-\gamma)\int_{-\infty}^{c} x f(x-\mu)\mathrm{d}x / \int_{-\infty}^{c} f(x-\mu)\mathrm{d}x$ is non-increasing in μ for $\gamma > 1$, it can be seen that $D'(\mu) \leq 0$ for all $\mu > 0$, or there exists a point μ_0 such that $D'(\mu) \geq 0$ for $0 < \mu \leq \mu_0$, and $D'(\mu) < 0$ for $\mu > \mu_0$. This implies that $D(\mu) \geq \min\{D((0), 0\}$ for all $\mu > 0$. Hence, $D(\mu) \geq 0$ for all $\mu > 0$ when $1 < \gamma \leq 2$. ■

Proposition 3.4 *For the function $\phi_i^{TR\gamma}(w)$, the function $H_i(\mu_i)$ is increasing in $\mu_i > 0$ when $0 < \gamma \leq 1$. When $\gamma > 1$, the function $H_i(\mu_i)$ is also increasing in μ_i if $f'(z)/f(z)$ is non-increasing in z. Thus, for both cases of γ, (2.5) is a necessary and sufficient condition for minimaxity of $\widehat{\boldsymbol{\theta}}^{TR\gamma} = \boldsymbol{A}^t \widehat{\boldsymbol{\mu}}^{TR\gamma}$ for $\widehat{\boldsymbol{\mu}}^{TR\gamma} = (\widehat{\mu}_1^{TR\gamma}, \ldots, \widehat{\mu}_k^{TR\gamma})^t$.*

Proof. For simplicity, we omit the index i in X_i, μ_i, $H_i(\cdot)$, c_i and others. Also, we here write $\phi_i^{TR\gamma}(X)$ as $\phi(X)$ for notatonal convenience. Since $\phi'(X) = \gamma I(X < c)$, it is seen that $E[\phi'(X)] = \gamma \int I(x < c) f(x - \mu) \mathrm{d}x = \gamma \int_{-\infty}^{c-\mu} f(z) \mathrm{d}z$. On the other hand,

$$E[\phi'(X)\{\phi(X) - (X - \mu)\}] = \int \gamma I(x < c)\{\gamma(x - c) + c - (x - \mu)\} f(x - \mu) \mathrm{d}x$$
$$= \gamma \int_{-\infty}^{c-\mu} \{(\gamma - 1)(z - c + \mu) + \mu\} f(z) \mathrm{d}u,$$

so that the function $H(\mu)$ can be written as

$$\begin{aligned} H(\mu) =& (\gamma - 1) \frac{\int_{-\infty}^{c-\mu} (z - c + \mu) f(z) \mathrm{d}z}{\int_{-\infty}^{c-\mu} f(z) \mathrm{d}z} + \mu \\ =& (\gamma - 1)\phi^{GB}(c - \mu) - c(\gamma - 1) + \gamma\mu, \end{aligned} \tag{3.7}$$

for $\phi^{GB}(w)$ defined in (2.6). Since $\phi^{GB}(w)$ is increasing in w, it is seen that $\phi^{GB}(c - \mu)$ is decreasing in μ. Thus, $H(\mu)$ is increasing in μ when $0 < \gamma \leq 1$. When $\gamma > 1$, from (3.7), $H(\mu)$ is expressed as

$$H(\mu) = (\gamma - 1) \frac{\int_{-\infty}^{c} x f(x - \mu) \mathrm{d}x}{\int_{-\infty}^{c} f(x - \mu) \mathrm{d}x} + \mu.$$

As verified in the proof of Proposition 3.3, $\int_{-\infty}^{c} x f(x - \mu) \mathrm{d}x / \int_{-\infty}^{c} f(x - \mu) \mathrm{d}x$ is non-decreasing in μ if $f'(z)/f(z)$ is non-increasing in z. Thus, $H(\mu)$ is increasing in μ when $\gamma > 1$, and the proof is complete. ■

It is noted that the assumption that $f'(z)/f(z)$ is nonincreasing in z is satisfied by the normal distribution and, more generally, all other location density functions with monotone likelihood ratio.

3.3 Case of the same distribution

Consider a special case that $f_1(z) = \cdots = f_k(z) = f(z)$. Then, condition (2.5) is expressed as

$$\sum_{i \in C} \sum_{j \in C, j \neq i} (\boldsymbol{A}\boldsymbol{A}^t)_{ij} \{B(0)\}^2 \leq \sum_{i \in C} (\boldsymbol{A}\boldsymbol{A}^t)_{ii} D(0),$$

for all subsets C of $\{1, \ldots, k\}$. This condition is simplified as

$$\frac{\sum_{i \in C} \sum_{j \in C} (\boldsymbol{A}\boldsymbol{A}^t)_{ij}}{\sum_{i \in C} (\boldsymbol{A}\boldsymbol{A}^t)_{ii}} \leq K_{f,\phi}, \tag{3.8}$$

where

$$K_{f,\phi} = \frac{D(0)}{\{B(0)\}^2} + 1. \tag{3.9}$$

For example, consider the case of $\boldsymbol{a}_1^t = (1, \ldots, 1)$, $\boldsymbol{a}_2^t = (1, -1, 0, \ldots, 0)$, $\boldsymbol{a}_3^t = (0, 1, -1, 0, \ldots, 0)$, $\ldots$, $\boldsymbol{a}_\ell^t = (0, \ldots, 0, 1, -1, 0, \ldots, 0)$ for $1 \le \ell \le k+1$, namely,

$$\boldsymbol{A}^t = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & -1 & 0 & \cdots & 0 \end{pmatrix},$$

where for $\ell = k+1$, the bottom row vector of $\boldsymbol{A}^t$ is given by $\boldsymbol{a}_{k+1}^t = (-1, 0, \ldots, 0, 1)$. This setup means that $\theta_1 = \sum_{j=1}^k \mu_j$ is the sum of means and the other parameters are contrasts $\theta_2 = \mu_1 - \mu_2$, $\ldots$, $\theta_k = \mu_{k-1} - \mu_k$ and $\theta_{k+1} = \mu_k - \mu_{k+1}$. In this case, the condition (3.8) is simplified as

$$\frac{k^2}{k + 2(\ell - 1)} \le K_{f,\phi}. \tag{3.10}$$

For $\ell = 1$, this is just $k \le K_{f,\phi}$, which corresponds to the result of Kubokawa and Strawderman (2011). It is interesting to note that larger ℓ eases more the condition for the minimaxity, and for $\ell = k+1$, it is $k/3 \le K_{f,\phi}$.

We may summarize much of the discussion for the case of $f_1(z) = \cdots = f_k(z) = f(z)$ as follows:

(I) A necessary and sufficient condition for the uniform prior generalized Bayes estimator to be minimax is that condition (SC2) holds, namely, $(\boldsymbol{A}\boldsymbol{A}^t)_{ij} \le 0$ for any i and j, $(i \ne j)$.

(II) A necessary and sufficient condition for the truncated estimator $\widehat{\boldsymbol{\theta}}^{TR\gamma}$ to be minimax is that condition (3.8) hold. In this case, we can obtain $K_{f,TR\gamma} = K_{f,\phi^{TR\gamma}} = D(0)/\{B(0)\}^2 + 1$ for $\phi^{TR\gamma}(w) = \min\{\gamma(w-c)+c, c\}$, where $B(0)$ and $D(0)$ are given in (3.5), and $\gamma > 0$. The following values of $K_{f,TR\gamma}$ have been found in Kubokawa and Strawderman (2011).

(a) Normal distribution, $X \sim \mathcal{N}(0,1)$: $K_{f,TR\gamma} = (2/\gamma - 1)\pi + 1$.

(b) Variance mixtures of normal distributions, namely, $X|V \sim \mathcal{N}(0, V)$ and $V \sim G$: $K_{f,TR\gamma} = (2/\gamma - 1)\pi E[V]/\{E[V^{1/2}]\}^2 + 1$.

In particular, t- and double exponential distributions give the following values.

(b1) t-distribution, $X \sim t_\nu$,

$$K_{f,TR\gamma} = (2/\gamma - 1)\pi \frac{2}{\nu - 2}\Big(\frac{\Gamma(\nu/2)}{\Gamma((\nu-1)/2)}\Big)^2 + 1.$$

(b2) Double exponential distribution, $X \sim DE(0)$: $K_{f,TR\gamma} = 4(2/\gamma - 1) + 1$.

(b3) Logistic distribution, $K_{f,TR\gamma} = (2/\gamma - 1)\pi^2/\{6[\log(2)]^2\} + 1 = 3.424(2/\gamma - 1) + 1$.

(c) $X \sim$ Symmetric unimodal distributions: $K_{f,TR\gamma} = 2(2/\gamma - 1)E[X^2]/\{E[|X|]\}^2 + 1 \geq (8/3)(2/\gamma - 1) + 1$.

Of course, for the case

$$\boldsymbol{A}^t = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}, \quad \boldsymbol{A}\boldsymbol{A}^t = \begin{pmatrix} \boldsymbol{I}_\ell & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

the condition (3.8) holds for all distributions. This implies that any subset of estimators $(\widehat{\mu}_{i_1}, \ldots, \widehat{\mu}_{i_\ell})$ is minimax for $(\mu_{i_1}, \ldots, \mu_{i_\ell})$. However, as we have shown, for other linear combinations, the trucated estimator (often the MLE) may or may not be minimax, but the condition (3.8) gives a reasonably straightforward necessary and sufficient condition for minimaxity for a set of linear combinations.

Remark 3.2 It is interesting to note that it may be possible to construct an example such that each estimator is not minimax in the context of estimating a single location parameter, but such that their combination is minimax for estimating the corresponding linear combination. For example, consider the estimation of $\theta = \mu_1 - \mu_2$ in normal distributions. Since $\boldsymbol{A}^t = (1, -1)$, the l.h.s. of the inequality (3.8) is zero, and the minimaxity of the truncated estimator $\hat{\theta}^{TR\gamma}$ is given by $0 \leq (2/\gamma - 1)\pi + 1$, or $0 < \gamma \leq 2\pi/(\pi - 1)$. Taking $\gamma = 2\pi/(\pi - 1)$, from Proposition 3.3, we can see that each estimator $\widehat{\mu}_i^{TR\gamma}$ is not minimax for estimating μ_i since $\gamma > 2$, while $\hat{\theta}^{TR\gamma}$ is minimax in estimation of $\theta = \mu_1 - \mu_2$.

4 Minimaxity and Non-minimaxity in Normal Distributions

In this section, we consider the cases of normal distributions with known or unknown variance. In Section 4.1, we provide a unified approach to necessary and sufficient conditions for minimaxity of the uniform prior generalized Bayes estimator and the MLE. In Section 4.2, we extend the dominance results to the unknown variance case.

4.1 A unified condition for minimaxity

Let $X_1, \ldots, X_k$ be mutually independent random variables such that X_i has a normal distribution with mean μ_i and unit variance, namely, $X_i \sim \mathcal{N}(\mu_i, 1)$ for $\mu_i > 0$. In the estimation of μ_i, an unbiased estimator of μ_i is $\widehat{\mu}_i^U = X_i$, which is minimax relative to the mean squared error. The maximum likelihood estimator of μ_i is $\widehat{\mu}_i^{ML} = \max\{X_i, 0\} =$

$X_i - \phi^{TR}(X_i)$ for $\phi^{TR}(w) = \min\{w, 0\}$. The uniform prior generalized Bayes estimator is written as $\widehat{\mu}_i^{GB} = X_i - \phi^{GB}(X_i)$ for

$$\phi^{GB}(w) = \int_{-\infty}^{w} z \exp\{-z^2/2\}\mathrm{d}z \Big/ \int_{-\infty}^{w} \exp\{-z^2/2\}\mathrm{d}z.$$

Both estimators are minimax, namely, they dominate $\widehat{\mu}_i^U$. As mentioned before, in general, an estimator of the form $X_i - \phi(X_i)$ is minimax if

(A1′) $\phi(w)$ is nondecreasing and $\lim_{w\to\infty} \phi(w) = 0$,

(A2′) $\phi(w) \geq \phi^{GB}(w)$.

We now treat the estimation of the vector of the linear combinations $\boldsymbol{\theta} = \boldsymbol{A}^t\boldsymbol{\mu}$ and consider the estimators $\widehat{\boldsymbol{\theta}}_\phi = \boldsymbol{A}^t\widehat{\boldsymbol{\mu}}_\phi$ where $\widehat{\boldsymbol{\mu}}_\phi = (\widehat{\mu}_{\phi,1}, \ldots, \widehat{\mu}_{\phi,k})^t$ for $\widehat{\mu}_{\phi,i} = X_i - \phi(X_i)$. Necessary and sufficient conditions for minimaxity of $\widehat{\boldsymbol{\theta}}_\phi$ are summarized in the following proposition which follows from Propositions 2.1, 2.2 and 3.1. Let $B(0) = -E[\phi(Z)]$, $D(0) = 2E[\phi(Z)\{Z - \phi(Z)/2\}]$ and $H(\mu) = E[\phi'(Z+\mu)\{\phi(Z+\mu) - Z\}]/E[\phi'(Z+\mu)]$ for $Z \sim \mathcal{N}(0,1)$.

Proposition 4.1 (1) *Assume the condition* (A1′). *If the estimator* $\widehat{\boldsymbol{\theta}}_\phi$ *is minimax, then for all subset* C *of* $\{1, \ldots, k\}$,

$$\frac{\sum_{i\in C}\sum_{j\in C}(\boldsymbol{A}\boldsymbol{A}^t)_{ij}}{\sum_{i\in C}(\boldsymbol{A}\boldsymbol{A}^t)_{ii}} \leq K_\phi, \tag{4.1}$$

where $K_\phi = D(0)/\{B(0)\}^2 + 1$ *for the standard normal distribution.*

(2) *Assume the conditions* (A1′) *and* (A2′). *Then, the estimator* $\widehat{\boldsymbol{\theta}}_\phi$ *is minimax if*
(SC2) $(\boldsymbol{A}\boldsymbol{A}^t)_{ij} \leq 0$ *for all* i, j $(i \neq j)$.

(3) *Assume the condition* (A1′) *and that* $H(\mu)$ *is nondecreasing in* μ. *Then, the condition* (4.1) *is a necessary and sufficient condition for the estimator* $\widehat{\boldsymbol{\theta}}_\phi$ *to be minimax.*

In general, it may be hard to check the monotonicity of $H(\mu)$ in part (3) of Proposition 4.1. When $\phi(w)$ is twice differentiable, however, this condition can be simplified in normal distributions as

(A3) The derivative $\phi'(w)$ is absolutely continuous and $\phi(w) - \phi''(w)/\phi'(w)$ is nondecreasing in w.

Proposition 4.2 *Under the normality assumption, if the condition* (A3) *is satisfied, then* $H(\mu)$ *is nondecreasing in* μ, *and the condition* (4.1) *is a necessary and sufficient condition for the estimator* $\widehat{\boldsymbol{\theta}}_\phi$ *to be minimax.*

Proof. The function $H(\mu)$ is expressed as

$$H(\mu) = \frac{\int_{-\infty}^{\infty} \phi'(x)\{\phi(x) - (x-\mu)\}f(x-\mu)\mathrm{d}x}{\int_{-\infty}^{\infty} \phi'(x)f(x-\mu)\mathrm{d}x}.$$

Note that $-\int_{-\infty}^{\infty} \phi'(x)(x-\mu)f(x-\mu)\mathrm{d}x = \int_{-\infty}^{\infty} \phi'(x)f'(x-\mu)\mathrm{d}x$. Since $\phi(x)$ is absolutely continuous, by integration by parts, it can be seen that

$$\int_{-\infty}^{\infty} \phi'(x)f'(x-\mu)\mathrm{d}x = -\int_{-\infty}^{\infty} \phi''(x)f(x-\mu)\mathrm{d}x,$$

so that $H(\mu)$ is rewritten as

$$H(\mu) = \frac{\int_{-\infty}^{\infty} \{\phi'(x)\phi(x) - \phi''(x)\}f(x-\mu)\mathrm{d}x}{\int_{-\infty}^{\infty} \phi'(x)f(x-\mu)\mathrm{d}x}.$$

Differentiating $H(\mu)$ with respect to μ shows that the derivative is proportional to

$$\begin{aligned} -&\frac{\int_{-\infty}^{\infty} \{\phi'(x)\phi(x) - \phi''(x)\}f'(x-\mu)\mathrm{d}x}{\int_{-\infty}^{\infty} \phi'(x)f(x-\mu)\mathrm{d}x} \\ &+ \frac{\int_{-\infty}^{\infty} \{\phi'(x)\phi(x) - \phi''(x)\}f(x-\mu)\mathrm{d}x}{\int_{-\infty}^{\infty} \phi'(x)f(x-\mu)\mathrm{d}x} \frac{\int_{-\infty}^{\infty} \phi'(x)f'(x-\mu)\mathrm{d}x}{\int_{-\infty}^{\infty} \phi'(x)f(x-\mu)\mathrm{d}x}. \end{aligned} \tag{4.2}$$

Since $f'(x-\mu)/f(x-\mu)$ is decreasing in x, and $\{\phi'(x)\phi(x)-\phi''(x)\}/\phi'(x)$ is nondecreasing in x, it can be seen that the derivative in (4.2) is positive, so that $H(\mu)$ is increasing in μ. ■

As shown below, the uniform prior generalized Bayes estimator satisfies the condition (A3). Taking into account this fact and Propositions 4.2 and 4.2, in normal distributions, we can provide a unified necessary and sufficient condition for minimaxity of the MLE and the uniform prior generalized Bayes estimator.

As an application of Propositions 4.1 and 4.2, we deal with estimators of the form $\widehat{\mu}_{i,\lambda} = X_i - \phi_\lambda(X_i)$ where

$$\phi_\lambda(w) = \int_{-\infty}^{w} z \exp\{-\lambda z^2/2\}\mathrm{d}z \Big/ \int_{-\infty}^{w} \exp\{-\lambda z^2/2\}\mathrm{d}z$$

for $\lambda > 0$. This estimator was studied by Maruyama and Iwasaki (2005), who showed that $\phi_\lambda(w)$ is nondecreasing in w and also in λ, which implies that $\phi_\lambda(w) \geq \phi_1(w) = \phi^{GB}(w)$ for $\lambda \geq 1$. Maruyama and Iwasaki (2005) proved that $\widehat{\mu}_{i,\lambda}$ is minimax if and only if $\lambda \geq 1$. Considering the estimation of $\boldsymbol{\theta}$, we can get sufficient conditions for minimaxity of the corresponding estimator $\widehat{\boldsymbol{\theta}}_\lambda = \boldsymbol{A}^t \widehat{\boldsymbol{\mu}}_\lambda$ for $\widehat{\boldsymbol{\mu}}_\lambda = (\widehat{\mu}_{1,\lambda}, \ldots, \widehat{\mu}_{k,\lambda})^t$.

Proposition 4.3 (1) *For $\lambda > 1$, the estimator $\widehat{\boldsymbol{\theta}}_\lambda$ is minimax if the condition* (SC2) *holds.*

(2) *For $0 < \lambda \leq 1$, the estimator $\widehat{\boldsymbol{\theta}}_\lambda$ is minimax if and only if the condition* (4.1) *holds.*

Proof. For part (1), the function $\phi_\lambda(w)$ satisfies the conditions (A1$'$) and (A2$'$) as demonstrated in Maruyama and Iwasaki (2005), so that the estimators $\widehat{\mu}_{i,\lambda}$ are minimax. Thus, the minimaxity result for $\lambda > 1$ follows from Proposition 4.1 (2).

For part (2), we shall show that the function $\phi_\lambda(w)$ satisfies the condition (A3). Since $\int_{-\infty}^{w} z\exp\{-\lambda z^2/2\}\mathrm{d}z = -\lambda^{-1}\exp\{-\lambda w^2/2\}$, the function $\phi_\lambda(w)$ is written as

$$\phi_\lambda(w) = -\frac{1}{\lambda}\frac{\exp\{-\lambda w^2/2\}}{\int_{-\infty}^{w}\exp\{-\lambda z^2/2\}\mathrm{d}z}.$$

Then,

$$\begin{aligned}\phi'_\lambda(w) =& -\lambda(w-\phi_\lambda(w))\phi_\lambda(w),\\ \phi''_\lambda(w) =& -\lambda(1-\phi'_\lambda(w))\phi_\lambda(w) - \lambda(w-\phi_\lambda(w))\phi'_\lambda(w),\end{aligned}$$

so that

$$\frac{\phi''_\lambda(w)}{\phi'_\lambda(w)} = \lambda\phi_\lambda(w) - \lambda(w-\phi_\lambda(w)) + \frac{1}{w-\phi_\lambda(w)}.$$

Thus,

$$\phi_\lambda(w) - \frac{\phi''_\lambda(w)}{\phi'_\lambda(w)} = (1-\lambda)\phi_\lambda(w) + \lambda(w-\phi_\lambda(w)) - \frac{1}{w-\phi_\lambda(w)}.$$

Note that $w-\phi_\lambda(w) = \int_0^\infty u\exp\{-\lambda(w-u)^2/2\}\mathrm{d}u/\int_0^\infty \exp\{-\lambda(w-u)^2/2\}\mathrm{d}u$. Differentiating $w-\phi_\lambda(w)$ with respect to w, we can see that $w-\phi_\lambda(w)$ is increasing in w. Hence, $\phi_\lambda(w)-\phi''_\lambda(w)/\phi'_\lambda(w)$ is increasing in w for $0<\lambda\le 1$. The result (2) follows from Proposition 4.2. ■

As shown in Maruyama and Iwasaki (2005), the estimator $\widehat{\mu}_{i,\lambda}$ is not minimax for $0<\lambda<1$. It is, however, interesting to note that Proposition 4.3 (2) implies that the estimator $\widehat{\boldsymbol{\theta}}_\lambda$ is minimax for certain conditions on $(\boldsymbol{A}\boldsymbol{A}^t)_{ij}$ even for $0<\lambda<1$. It is also worth noting that the case of $\lambda=1$ corresponds to the uniform prior generalized Bayes estimator and that Proposition 4.3 means the function $\phi^{GB}(w)$ satisfies the condition (A3). That is, the condition (4.1) is a necessary and sufficient condition for the uniform prior generalized Bayes estimator to be minimax. Since $D(0)=0$ for the uniform prior generalized Bayes estimator $\widehat{\mu}_i^{GB}$, the condition (4.1) is identical to the condition (SC2) as shown in Proposition 2.3.

4.2 An extension to the case of unknown variance

It is quite interesting to consider the extension of the previous results to location-scale families. In general, however, this extension may be difficult, because estimators of location parameters are not necessarily independent of estimators of scale parameters. Extension for a specific distribution may be feasible however. We here treat normal distributions with common unknown variance.

Let $X_1,\ldots,X_k$ and S be mutually independent random variables distributed as

$$\begin{aligned}X_i \sim& \mathcal{N}(\mu_i,\sigma^2), \quad \text{for} \quad i=1,\ldots,k,\\ S \sim& \sigma^2\chi^2_m,\end{aligned}$$

where μ_i's are restricted as $\mu_i>0$ and χ^2_m denoted a chi-square distribution with m degrees of freedom. This is a canonical form of a random sample from k normal populations with unknown common variance. As studied in the previous sections, we deal with estimation of a set of the linear combinations $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_\ell)^t = \boldsymbol{A}^t\boldsymbol{\mu}$ for the restricted parameter space $\Omega = \{(\boldsymbol{\mu}, \sigma^2) | \mu_i > 0, i = 1, \ldots, k, \sigma^2 > 0\}$ for $\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell)$, a $k \times \ell$ matrix, and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)^t$. An estimator $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is evaluated relative to the quadratic loss $\|\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|^2/\sigma^2 = \sum_{i=1}^k (\widehat{\mu}_i - \mu_i)^2/\sigma^2$.

Proposition 4.4 *The unbiased estimator $\widehat{\boldsymbol{\theta}}^U = \boldsymbol{A}^t\boldsymbol{X}$ for $\boldsymbol{X} = (X_1, \ldots, X_k)^t$ is minimax under the restricted space Ω with unknown variance σ^2.*

The proof of Proposition 4.4 is given in the appendix. To construct a class of minimax estimators improving on $\widehat{\boldsymbol{\theta}}^U$, consider estimators of the form

$$\widehat{\boldsymbol{\theta}}_\phi = \boldsymbol{A}^t\widehat{\boldsymbol{\mu}}_\phi \quad \text{for} \quad \widehat{\boldsymbol{\mu}}_\phi = (\widehat{\mu}_{\phi,1}, \ldots, \widehat{\mu}_{\phi,k})^t, \tag{4.3}$$

where

$$\widehat{\mu}_{\phi,i} = X_i - \sqrt{S}\phi\left(X_i/\sqrt{S}\right),$$

for an absolutely continuous function ϕ. It can be seen that the expectations $E[(\widehat{\mu}_{\phi,i} - \mu_i)(\widehat{\mu}_{\phi,j} - \mu_j)/\sigma^2]$ depends on μ_i, μ_j and σ^2 through λ_i and λ_j for $\lambda_i = \mu_i/\sigma$ and $\lambda_j = \mu_j/\sigma$. Let $R(\lambda_i) = E[(\widehat{\mu}_{\phi,i} - \mu_i)^2/\sigma^2]$. The following lemmas due to Kubokawa (2004) are useful for deriving conditions for minimaxity of the estimator $\widehat{\boldsymbol{\theta}}_\phi$.

Lemma 4.1 *The risk difference of the two estimators X_i and $\widehat{\mu}_{\phi,i}$ is*

$$\begin{aligned} D(\lambda_i) =& E[(X_i - \mu_i)^2/\sigma^2] - R(\lambda_i) \\ =& 2c\int \{G_{\lambda_i}(w_i) - \phi(w_i)\}\,\phi'(w_i)F_{\lambda_i}(w_i)\mathrm{d}w_i, \end{aligned}$$

where c is the normalizing constant and

$$\begin{aligned} G_\lambda(w) =& \frac{\int\int_{-\infty}^w (y - \lambda/\sqrt{v})v^{(m+1)/2}e^{-v\{1+(y-\lambda/\sqrt{v})^2\}/2}\mathrm{d}y\mathrm{d}v}{\int\int_{-\infty}^w v^{(m+1)/2}e^{-v\{1+(y-\lambda/\sqrt{v})^2\}/2}\mathrm{d}y\mathrm{d}v} \\ =& -\int v^{(m-1)/2}e^{-v(1+w^2)/2+\sqrt{v}w\lambda}\mathrm{d}v/F_\lambda(w). \end{aligned}$$

for $F_\lambda(w) = \int\int_{-\infty}^w v^{(m+1)/2}e^{-v(1+y^2)/2+\sqrt{v}y\lambda}\mathrm{d}y\mathrm{d}v$.

Lemma 4.2 *When λ goes to zero, $G_\lambda(w)$ converges to $\lim_{\lambda\to 0} G_\lambda(w) = \phi^{GB}(w)$, where*

$$\begin{aligned} \phi^{GB}(w) =& \int_0^\infty \int_{-\infty}^w ye^{-v(1+y^2)/2}\mathrm{d}yv^{(m+1)/2}\mathrm{d}v\ /F_0(w) \\ =& -\frac{1}{m+1}\frac{(1+w^2)^{-(m+1)/2}}{\int_{-\infty}^w (1+y^2)^{-(m+1)/2-1}\mathrm{d}y}, \end{aligned} \tag{4.4}$$

for $F_0(w) = \lim_{\lambda\to 0} F_\lambda(w)$. Also, $G_\lambda(w) \leq \phi^{GB}(w)$.

The above lemmas imply the following proposition.

Proposition 4.5 *Assume that $\phi(w)$ satisfies the following conditions:*
(A1$''$) *$\phi(w)$ is nondecreasing in w, and $\lim_{w\to\infty}\phi(w)=0$,*
(A2$''$) *$\phi(w)\ge\phi^{GB}(w)$.*
Then $\widehat{\mu}_{\phi,i}$ dominates the unbiased estimator X_i.

Let $B(\lambda_i,\lambda_j)=E[(\widehat{\mu}_{\phi,i}-\mu_i)(\widehat{\mu}_{\phi,j}-\mu_j)/\sigma^2]$ for $i\neq j$. Then, $B(0,0)$ for $\lambda_i=\lambda_j=0$ is written as

$$B(0,0)=E[T\phi(Z_1/\sqrt{T})\phi(Z_2/\sqrt{T})], \tag{4.5}$$

where Z_1, Z_2 and T are mutually independent random variables such that $Z_1\sim\mathcal{N}(0,1)$, $Z_2\sim\mathcal{N}(0,1)$ and $T\sim\chi^2_m$.

Proposition 4.6 (1) *Assume the condition* (A1$''$). *If the estimator $\widehat{\boldsymbol{\theta}}_\phi$ is minimax, then for all subset C of $\{1,\ldots,k\}$,*

$$\frac{\sum_{i\in C}\sum_{j\in C}(\boldsymbol{A}\boldsymbol{A}^t)_{ij}}{\sum_{i\in C}(\boldsymbol{A}\boldsymbol{A}^t)_{ii}}\le K^*_\phi, \tag{4.6}$$

where $K^*_\phi=D(0)/B(0,0)+1$.

(2) *Assume the conditions* (A1$''$) *and* (A2$''$). *Then, the estimator $\widehat{\boldsymbol{\theta}}_\phi$ is minimax if*
(SC2) $(\boldsymbol{A}\boldsymbol{A}^t)_{ij}\le 0$ *for all i, j $(i\neq j)$.*

Proof. Let $\boldsymbol{\lambda}=(\lambda_1,\ldots,\lambda_k)^t$. The risk function of the estimator $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}}$ is written as

$$\begin{aligned}R(\boldsymbol{\lambda},\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}})=&E[\|\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}}-\boldsymbol{\theta}\|^2/\sigma^2]=E[(\widehat{\boldsymbol{\mu}}_{\boldsymbol{\phi}}-\boldsymbol{\mu})^t\boldsymbol{A}\boldsymbol{A}^t(\widehat{\boldsymbol{\mu}}_{\boldsymbol{\phi}}-\boldsymbol{\mu})/\sigma^2]\\ =&\sum_{i=1}^k(\boldsymbol{A}\boldsymbol{A}^t)_{ii}R(\lambda_i)+\sum_{i=1}^k\sum_{j=1,j\neq i}^k(\boldsymbol{A}\boldsymbol{A}^t)_{ij}B(\lambda_i,\lambda_j),\end{aligned}$$

so that the risk difference of $\widehat{\boldsymbol{\theta}}_\phi$ and $\widehat{\boldsymbol{\theta}}^U$ is

$$\begin{aligned}\Delta(\boldsymbol{\lambda};\phi)=&R(\boldsymbol{\lambda},\widehat{\boldsymbol{\theta}}_{\boldsymbol{\phi}})-R(\boldsymbol{\lambda},\widehat{\boldsymbol{\theta}}^U)\\ =&-\sum_{i=1}^k(\boldsymbol{A}\boldsymbol{A}^t)_{ii}D(\lambda_i)+\sum_{i=1}^k\sum_{j=1,j\neq i}^k(\boldsymbol{A}\boldsymbol{A}^t)_{ij}B(\lambda_i,\lambda_j).\end{aligned} \tag{4.7}$$

For part (i), note that $\lim_{\lambda_i\to\infty}D(\lambda_i)=0$ and $\lim_{\lambda_i\to\infty}B(\lambda_i,\lambda_j)=0$. Let C be any subset of $\Lambda=\{1,\ldots,k\}$. If $\mu_i\to 0$ for all $i\in C$, and if $\mu_j\to\infty$ for all $j\in\Lambda\backslash C$, then the risk difference $\Delta(\boldsymbol{\mu})$ converges to

$$-\sum_{i\in C}(\boldsymbol{A}\boldsymbol{A}^t)_{ii}D(0)+\sum_{i\in C}\sum_{j\in C,j\neq i}(\boldsymbol{A}\boldsymbol{A}^t)_{ij}B(0,0),$$

which yields the necessary condition (4.6).

For part (2), Proposition 4.5 shows that the conditions (A1$''$) and (A2$''$) imply that $D(\lambda_i) \geq 0$ and $B(\lambda_i, \lambda_j) \geq 0$ for $\lambda_i > 0$ and $\lambda_j > 0$. Thus, part (2) follows from (4.7). ■

Particular attention is paid to the truncated invariant prior generalized Bayes estimator and the maximum likelihood estimator. When we assume the truncated invariant prior distribution $(\prod_{i=1}^k \mathrm{d}\mu_i)\mathrm{d}\sigma^2/\sigma^2$ over $\mu_i > 0$, $\sigma^2 > 0$ for $i = 1, \ldots, k$, the generalized Bayes estimator of $\boldsymbol{\theta}$ is written as

$$\widehat{\boldsymbol{\theta}}^{GB} = \boldsymbol{A}^t \widehat{\boldsymbol{\mu}}^{GB} \quad \text{for} \quad \widehat{\boldsymbol{\mu}}^{GB} = (\widehat{\mu}_1^{GB}, \ldots, \widehat{\mu}_k^{GB})^t, \tag{4.8}$$

where

$$\begin{aligned}\widehat{\mu}_i^{GB} =& \frac{\int_0^\infty \int_0^\infty \mu_i (\sigma^2)^{-(m+1)/2-2} e^{-\{(X_i-\mu_i)^2+S\}/2\sigma^2} \mathrm{d}\mu_i \mathrm{d}\sigma^2}{\int_0^\infty \int_0^\infty (\sigma^2)^{-(m+1)/2-2} e^{-\{(X_i-\mu_i)^2+S\}/2\sigma^2} \mathrm{d}\mu_i \mathrm{d}\sigma^2} \\ =& X_i - \sqrt{S}\phi^{GB}(X_i/\sqrt{S}),\end{aligned}$$

where the function $\phi^{GB}(w)$ is defined by (4.8). It can be seen that $\phi^{GB}(w)$ satisfies all the conditions of Proposition 4.5, and the generalized Bayes estimator $\widehat{\mu}^{GB}$ dominates X_i. Also, note that $D(\lambda_i) = 1 - R(\lambda_i) \geq 0$ and $L(0) = 0$ for $\phi^{GB}(w)$. This fact implies that the necessary and sufficient condition for minimaxity is (SC2).

Concerning the MLE, it is given by $\widehat{\boldsymbol{\theta}}^{TR} = \boldsymbol{A}^t \widehat{\boldsymbol{\mu}}^{TR}$ for $\widehat{\boldsymbol{\mu}}^{TR} = (\widehat{\mu}_1^{TR}, \ldots, \widehat{\mu}_k^{TR})^t$, where

$$\widehat{\mu}_i^{TR} = \max\{X_i, 0\} = X_i - \sqrt{S}\phi^{TR}(X_i/\sqrt{S}),$$

for $\phi^{TR}(w) = \min\{w, 0\}$. It is noted that $\phi^{TR}(X_i/\sqrt{S}) = \min\{X_i, 0\}/\sqrt{S}$. It can be seen that the same arguments as in the case of known σ^2 can be used to derive the minimaxity of the MLE. Thus, the necessary and sufficient condition for minimaxity of $\widehat{\boldsymbol{\theta}}^{TR}$ is given by (4.6), where $K_{\phi^{TR}} = D(0)/\{B(0)\}^2 + 1 = \pi + 1$.

Proposition 4.7 (1) *The truncated invariant prior generalized Bayes estimator* $\widehat{\boldsymbol{\theta}}^{GB}$ *given in* (4.8) *is minimax if and only if the condition* (SC2) *is satisfied.*

(2) *The MLE* $\widehat{\boldsymbol{\theta}}^{TR}$ *is minimax if and only if for all subsets* C *of* $\{1, \ldots, k\}$,

$$\sum_{i \in C}\sum_{j \in C} (\boldsymbol{A}\boldsymbol{A}^t)_{ij} \Big/ \sum_{i \in C} (\boldsymbol{A}\boldsymbol{A}^t)_{ii} \leq \pi + 1.$$

Remark 4.1 It is noted that $B(\lambda_i, \lambda_j)$ defined around (4.5) can be written as $B(\lambda_i, \lambda_j) = E[T\phi((Z_i+\lambda_i)/\sqrt{T})\phi((Z_j+\lambda_j)/\sqrt{T})]$ for $i \neq j$. Thus, the partial derivative with respect to λ_i is

$$\begin{aligned}\frac{\partial}{\partial \lambda_i} B(\lambda_i, \lambda_j) =& E[\sqrt{T}\phi'((Z_i+\lambda_i)/\sqrt{T})\phi((Z_j+\lambda_j)/\sqrt{T})] \\ =& E^T\left[\sqrt{T}E[\phi'((Z_i+\lambda_i)/\sqrt{T})|T] \cdot E[\phi((Z_j+\lambda_j)/\sqrt{T})|T]\right],\end{aligned}$$

which cannot be separated into two expectations like (3.1). Hence, we cannot use the same arguments as in Section 3.1 to show the sufficiency of the necessary condition. It would be desirable to get a result corresponding to Propositions 3.1 and 4.1(3) in the case of unknown variance.

Remark 4.2 Proposition 4.7 (2) can be extended to symmetric unimodal location-scale distributions. Let $(X_1, V_1), \ldots, (X_k, V_k)$ be mutually independent random variables such that (X_i, V_i) has joint density function $\sigma^{-2} f_i((x_i - \mu_i)/\sigma, v_i/\sigma)$ with $\mu_i > 0$ and unknown scale σ, where $f_i(z, w)$ is symmetric and unimodal on $z = 0$ with respect to z. Then, the MLE of $\boldsymbol{\theta}$ is $\widehat{\boldsymbol{\theta}}^{TR} = \boldsymbol{A}^t \widehat{\boldsymbol{\mu}}^{TR}$ where $\widehat{\boldsymbol{\mu}}^{TR}$ is a vector of $\max\{X_i, 0\}$. Noting that the MLE $\widehat{\boldsymbol{\theta}}^{TR}$ does not depend on statistics V_i's, and that X_i's are mutually independent, we can apply the same arguments as in Section 3.1 to show that the necessary sufficient condition for minimaxity of $\widehat{\boldsymbol{\theta}}^{TR}$ is (2.5), where from (3.4) with $\gamma = 1$,

$$B_i(\lambda_i) = -\int_{-\infty}^{-\lambda_i} (z + \lambda_i) g_i(z) \mathrm{d}z,$$

$$D_i(\lambda_i) = \int_{-\infty}^{-\lambda_i} z(z + \lambda_i) g_i(z) \mathrm{d}z + \lambda_i B_i(\lambda_i).$$

where $g_i(z) = \int_0^\infty f_i(z, w) \mathrm{d}w$ and $\lambda_i = \mu_i/\sigma$.

5 Concluding Remarks

In this paper, we have derived necessary and/or sufficient conditions for minimaxity of general types of estimators in the simultaneous estimation of a set of linear combinations $\boldsymbol{\theta} = \boldsymbol{A}^t \boldsymbol{\mu}$ where the location parameters μ_i's are restricted to positive real numbers. When $\boldsymbol{\theta}$ is estimated by the uniform prior generalized Bayes estimator $\widehat{\boldsymbol{\theta}}^{GB}$, the necessary and sufficient codition for minimaxity is that all the off-diagonal elements $(\boldsymbol{A}\boldsymbol{A}^t)_{ij}$, $(i \neq j)$, are not positive. Hartigan (2004) proved that $\widehat{\boldsymbol{\theta}}^{GB}$ is always minimax in normal distributions when $\boldsymbol{A}$ is the identity matrix $\boldsymbol{I}_k$, where his result guarantees the minimaxity when $\boldsymbol{\mu}$ is restricted to a general convex set. When $\boldsymbol{A}^t = (a_1, \ldots, a_k)$, on the other hand, Kubokawa (2010) showed that $\widehat{\boldsymbol{\theta}}^{GB}$ is minimax if and only if $k = 1$, or $(k = 2, a_1 a_2 \leq 0)$. This means, in a sense, that the results given in this paper fill in gaps between the two results given by Hartigan (2004) and Kubokawa (2010).

The paper also gives conditions on estimators under which the condition (2.5) becomes necessary and sufficient for minimaxity, and this result has been applied to a class of truncated estimators. When the underlying distributions are normal, we have shown that this gives a unified condition which can be applied to both the uniform prior generalized Bayes estimator and the MLE.

Finally, we want to conclude this section by describing some interesting related issues to be (hopefully) resolved in the future.

(1) Is it possible to construct a prior distribution, other than the uniform prior, such that the resulting generalized Bayes estimator is minimax ? No such prior has been found even for $k = 1$.

(2) An admissible and minimax estimator of $\boldsymbol{\theta} = \boldsymbol{A}^t \boldsymbol{\mu}$ was derived by Kubokawa (2010) when $\boldsymbol{A}^t = (a_1, \ldots, a_k)$, namely, $\ell = 1$. Can this result be extended to the case of $\ell \geq 2$?

(3) In this paper, the location parameters μ_i's are restricted to positive real numbers. Can the results given in this paper be extended to the case where $\boldsymbol{\mu}$ is restricted to a general convex cone?

(4) When $\boldsymbol{X} \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{I}_k)$ and $\boldsymbol{\mu}$ is restricted to a convex set, Hartigan (2004) proved that $\boldsymbol{X}$ is dominated by the uniform prior generalized Bayes estimator. Is it possible to extend his result to the case of $\boldsymbol{X} \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ for known $\boldsymbol{\Sigma}$ or more general location family?

Acknowledgments. This research was supported in part by Grant-in-Aid for Scientific Research Nos. 19200020 and 21540114 from Japan Society for the Promotion of Science.

References

[1] Casella, G. and Strawderman, W.E. (1981). Estimating a bounded mean. *Ann. Statist.*, **9**, 870-879.

[2] Farrell, R.H. (1964). Estimators of a location parameter in the absolutely continuous case. *Ann. Math. Statist.*, **35**, 949-998.

[3] Hartigan, J. (2004). Uniform priors on convex sets improve risk. *Statist. Prob. Letters*, **67**, 285-288.

[4] Katz, M. (1961). Admissible and minimax estimates of parameters in truncated spaces. *Ann. Math. Statist.*, **32**, 136-142..

[5] Kubokawa,T. (2004). Minimaxity in estimation of restricted parameters. *J. Japan Statist. Soc.*, **34**, 229-253.

[6] Kubokawa, T. (2010). Minimax estimation of linear combinations of restricted location parameters. Discussion Paper Series in Faculty of Economics, University of Tokyo, CIRJE-F-723. *IMS Monograph Series*, to appear.

[7] Kubokawa, T. and Strawderman, W. E. (2011). Non-minimaxity of Linear Combinations of Restricted Location Estimators and Related Problems. *J. Statist. Plan. Inf.*, to appear.

[8] Marchand, E. and Perron, F. (2001). Improving on the MLE of a bounded normal mean. *Ann. Statist.*, **29**, 1078-1093.

[9] Marchand, E. and Strawderman, W. E. (2004). Estimation in restricted parameter spaces: A review. *Festschrift for Herman Rubin, IMS Lecture Notes-Monograph Series*, **45**, 21-44.

[10] Maruyama, Y. and Iwasaki, K. (2005). Sensitivity of minimaxity and admissibility in the estimation of a positive normal mean. *Ann. Inst. Statist. Math.*, **57**, 145-156.

[11] Oono, Y. and Shinozaki, N. (2005). Estimation of two order restricted normal means with unknown and possibly unequal variances. *J. Statist. Plann. Inf.*, **131**, 349-363.

[12] Tsukuma, H. and Kubokawa, T. (2008). Stein's phenomenon in estimation of means restricted to a polyhedral convex cone. *J. Multivariate Analysis*, **99**, 141-164.

[13] van Eeden, C. (2006). *Restricted-parameter-space Estimation Problems - Admissibility and Minimaxity Properties.* Springer, Lecture Notes in Statistics.

A Minimaxity of Unbiased Estimators in Location-Scale Family

We here give the proof of Proposition 4.4, namely, minimaxity of the unbiased estimator $\boldsymbol{A}^t\boldsymbol{X}$ under the restriction $\Omega = \{(\boldsymbol{\mu},\sigma)|\mu_i > 0, i = 1,\dots,k, \sigma > 0\}$. In this section, we treat the following location-scale family: Let $\boldsymbol{X} = (X_1,\dots,X_k)^t$ and V be random variables whose joint density function is given by $\sigma^{-k-1}f((\boldsymbol{x}-\boldsymbol{\mu})/\sigma, v/\sigma)$ where $(\boldsymbol{\mu},\sigma)\in\Omega$ and $f((\boldsymbol{x}-\boldsymbol{\mu})/\sigma, v/\sigma)$ denotes $f((x_1-\mu_1)/\sigma,\dots,(x_k-\mu_k)/\sigma, v/\sigma)$. This includes the canonical form treated in Section 4.2. The random variable S in Proposition 4.4 corresponds to $S = V^2$. When the quadratic loss $\|\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\|^2/\sigma^2$ is used to evaluate estimators of $\boldsymbol{\theta} = \boldsymbol{A}^t\boldsymbol{\mu}$, the unrestricted generalized Bayes estimator against the uniform prior $\mathrm{d}(\boldsymbol{\mu},\sigma) = (\prod_{i=1}^k \mathrm{d}\mu_i)\mathrm{d}\sigma/\sigma$ is written as $\widehat{\boldsymbol{\theta}}^M = \boldsymbol{A}^t\widehat{\boldsymbol{\mu}}^M$, where

$$\begin{aligned}\widehat{\boldsymbol{\mu}}^M =& \int \boldsymbol{a}b^{-k-4}f((\boldsymbol{X}-\boldsymbol{a})/b, V/b)\mathrm{d}(\boldsymbol{a},b)/\int b^{-k-4}f((\boldsymbol{X}-\boldsymbol{a})/b, V/b)\mathrm{d}(\boldsymbol{a},b)\\ =&\boldsymbol{X}-\boldsymbol{c}V,\end{aligned} \tag{A.1}$$

where $\boldsymbol{c} = \int \boldsymbol{\xi}\tau f(\boldsymbol{\xi},\tau)\mathrm{d}(\boldsymbol{\xi},\tau)/\int \tau^2 f(\boldsymbol{\xi},\tau)\mathrm{d}(\boldsymbol{\xi},\tau)$ and the integrals are taken over $-\infty < \xi_i < \infty$, $i = 1,\dots,k$, $0 < \tau < \infty$ for $\boldsymbol{\xi} = (\xi_1,\dots,\xi_k)^t$. When $f(\boldsymbol{\xi},\tau)$ is symmetric on $\xi_i = 0$ for $i = 1,\dots,k$, we have $\widehat{\boldsymbol{\mu}} = \boldsymbol{X}$ since $\boldsymbol{c} = \boldsymbol{0}$ in this case.

Proposition A.1 *The unrestricted generalized Bayes estimator* $\widehat{\boldsymbol{\theta}}^M$ *is minimax under the restriction* $(\boldsymbol{\mu},\sigma)\in\Omega$.

Proof. Let $\Omega_n = \{(\boldsymbol{\mu},\sigma)|0 < \mu_i < n, i = 1,\dots,k, n^{-1} < \sigma < n\}$ for $n \geq 2$. Consider the sequence of prior distributions given by

$$\pi_n(\boldsymbol{\mu},\sigma)\mathrm{d}(\boldsymbol{\mu},\sigma) = \begin{cases} \{2n^k\log n\}^{-1}\sigma^{-1}\mathrm{d}\boldsymbol{\mu}\mathrm{d}\sigma & \text{if } (\boldsymbol{\mu},\sigma)\in\Omega_n\\ 0 & \text{otherwise,}\end{cases}$$

where $\mathrm{d}(\boldsymbol{\mu},\sigma)$ means $\mathrm{d}\boldsymbol{\mu}\mathrm{d}\sigma$ for $\mathrm{d}\boldsymbol{\mu} = \prod_{i=1}^k \mathrm{d}\mu_i$. Then the Bayes estimators $\boldsymbol{\theta}$ are given by $\widehat{\boldsymbol{\theta}}_n^\pi = \boldsymbol{A}^t\widehat{\boldsymbol{\mu}}_n^\pi$ where

$$\widehat{\boldsymbol{\mu}}_n^\pi = \int_{\Omega_n}\boldsymbol{a}b^{-k-4}f((\boldsymbol{x}-\boldsymbol{a})/b, v/b)\mathrm{d}(\boldsymbol{a},b)\int_{\Omega_n} b^{-k-4}f((\boldsymbol{x}-\boldsymbol{a})/b, v/b)\mathrm{d}(\boldsymbol{a},b)$$

with the Bayes risk function

$$
\begin{aligned}
r_n(\pi_n, \widehat{\boldsymbol{\theta}}_n^{\pi}) =& \{2n^k \log n\}^{-1} \int_{\Omega_n} \int \frac{\|\boldsymbol{A}^t(\widehat{\boldsymbol{\mu}}_n^{\pi}(\boldsymbol{x}) - \boldsymbol{\mu})\|^2}{\sigma^2} \frac{1}{\sigma^{k+2}} f\left(\frac{\boldsymbol{x}-\boldsymbol{\mu}}{\sigma}, \frac{v}{\sigma}\right) \mathrm{d}\boldsymbol{x} \mathrm{d}(\boldsymbol{\mu}, \sigma) \\
=& \{2n^k \log n\}^{-1} \int_{\Omega_n} \int \frac{\|\boldsymbol{A}^t(\widehat{\boldsymbol{\mu}}_n^{\pi}(\sigma\boldsymbol{z}+\boldsymbol{\mu}) - \boldsymbol{\mu})\|^2}{\sigma^2} f(\boldsymbol{z}) \, \mathrm{d}\boldsymbol{z} \frac{1}{\sigma} \mathrm{d}(\boldsymbol{\mu}, \sigma) \qquad \text{(A.2)}
\end{aligned}
$$

where $\boldsymbol{z} = (\boldsymbol{x}-\boldsymbol{\mu})/\sigma$. Letting $\boldsymbol{t} = (t_1, \ldots, t_k)^t = (\boldsymbol{a}-\boldsymbol{\mu})/\sigma$ and $s = b/\sigma$, we see that

$$
\begin{aligned}
\frac{\widehat{\boldsymbol{\mu}}^{\pi}(\sigma\boldsymbol{z}+\boldsymbol{\mu}) - \boldsymbol{\mu}}{\sigma} =& \frac{\int_{\Omega_n} [(\boldsymbol{a}-\boldsymbol{\mu})/\sigma](\sigma/b)^{k+4} f([\boldsymbol{z} - (\boldsymbol{a}-\boldsymbol{\mu})/\sigma]\sigma/b) \mathrm{d}(\boldsymbol{a}, b)}{\int_{\Omega_n} (\sigma/b)^{k+4} f([\boldsymbol{z} - (\boldsymbol{a}-\boldsymbol{\mu})/\sigma]\sigma/b) \mathrm{d}(\boldsymbol{a}, b)} \\
=& \frac{\int_{\Omega_n^*} \boldsymbol{t} s^{-k-4} f((\boldsymbol{z}-\boldsymbol{t})s) \mathrm{d}(\boldsymbol{t}, s)}{\int_{\Omega_n^*} s^{-k-4} f((\boldsymbol{z}-\boldsymbol{t})s) \mathrm{d}(\boldsymbol{t}, s)}, \qquad \text{(A.3)}
\end{aligned}
$$

where $\Omega_n^* = \{(\boldsymbol{t}, s) \,\big|\, -\mu_i < \sigma t_i < n - \mu_i, i = 1, \ldots, k, n^{-1} < \sigma s < n\}$. Let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_k)^t$ for $\xi_i = (2/n)\mu_i - 1$ and $\eta = \log \sigma / \log n$. Then, Ω_n^* is rewritten as

$$
\begin{aligned}
\Omega_n^* =& \Big\{(\boldsymbol{t}, s)| - \frac{1}{2} n^{1-\eta}(1+\xi_i) < t_i < \frac{1}{2} n^{1-\eta}(1-\xi_i), i = 1, \ldots, k, \\
& -(1+\eta)\log n < \log s < (1-\eta)\log n\Big\} \qquad \text{(A.4)}
\end{aligned}
$$

and we denote the quantity (A.3) by $\widehat{\boldsymbol{\mu}}_n^*(\boldsymbol{z}|\boldsymbol{\xi}, \eta)$. Since the condition that $(\boldsymbol{\mu}, \sigma) \in \Omega_n$ is equivalently expressed by

$$
(\boldsymbol{\xi}, \eta) \in U_0 = \{(\boldsymbol{\xi}, \eta) \mid |\xi_i| < 1, i = 1, \ldots, \; |\eta| < 1\},
$$

the Bayes risk (A.2) is rewritten as

$$
\begin{aligned}
r_n(\pi_n, \widehat{\mu}_n^{\pi}) =& \frac{(n/2)^k \log n}{2n^k \log n} \int_{(\boldsymbol{\xi},\eta)\in U_0} \int \left\|\boldsymbol{A}^t \widehat{\boldsymbol{\mu}}_n^*(\boldsymbol{z}|\boldsymbol{\xi}, \eta)\right\|^2 f(\boldsymbol{z}) \mathrm{d}\boldsymbol{z} \mathrm{d}(\boldsymbol{\xi}, \eta) \\
\geq& \frac{1}{2^{k+1}} \int_{(\boldsymbol{\xi},\eta)\in U_\varepsilon} \int \left\|\boldsymbol{A}^t \widehat{\boldsymbol{\mu}}_n^*(\boldsymbol{z}|\boldsymbol{\xi}, \eta)\right\|^2 f(\boldsymbol{z}) \mathrm{d}\boldsymbol{z} \mathrm{d}(\boldsymbol{\xi}, \eta),
\end{aligned}
$$

where $U_\varepsilon = \{(\boldsymbol{\xi}, \eta) \mid |\xi_i| < 1-\varepsilon, i = 1, \ldots, \; |\eta| < 1-\varepsilon\}$ for $\varepsilon > 0$. Noting that $1-\eta > \varepsilon$, $1+\eta > \varepsilon$, $1-\xi > \varepsilon$ and $1+\xi > \varepsilon$, we see that the set Ω_n^* given in (A.4) contains the subset

$$
(\boldsymbol{t}, s) \in \left\{-\frac{\varepsilon}{2} n^{\varepsilon} < t_i < \frac{\varepsilon}{2} n^{\varepsilon}, \; i = 1, \ldots, k, -\varepsilon \log n < \log s < \varepsilon \log n\right\},
$$

which implies that all the end points of t_i and $\log s$ go to infinity or minus infinity as n tends to infinity, so that

$$
\lim_{n\to\infty} \widehat{\boldsymbol{\mu}}_n^*(\boldsymbol{z}|\boldsymbol{\xi}, \eta) = \widehat{\boldsymbol{\mu}}^M.
$$

Hence, Fatou's lemma is used to evaluate the Bayes risk as

$$
\begin{aligned}
\liminf_{n\to\infty} r_n(\pi_n, \widehat{\boldsymbol{\theta}}_n^{\pi}) \geq& \frac{1}{2^{k+1}} \int_{(\boldsymbol{\xi},\eta)\in U_\varepsilon} \mathrm{d}(\boldsymbol{\xi}, \eta) \int \|\boldsymbol{A}^t \widehat{\boldsymbol{\mu}}^M(\boldsymbol{z})\|^2 f(\boldsymbol{z}) \mathrm{d}\boldsymbol{z} \\
=& (1-\varepsilon)^{k+1} \int \|\boldsymbol{A}^t \widehat{\boldsymbol{\mu}}^M(\boldsymbol{z})\|^2 f(\boldsymbol{z}) \mathrm{d}\boldsymbol{z},
\end{aligned}
$$

which establishes the minimaxity of the estimator $\widehat{\boldsymbol{\theta}}^M$. ■

It is noted that the same arguments as in the above proof can be applied to show minimaxity of the unrestricted uniform prior generalized Bayes estimator in various location-scale distributions. For exmple, consider joint density function $\prod_{i=1}^k \{\sigma^{-2} f_i((x_i - \mu_i)/\sigma, v_i/\sigma)\}$ for random variables $(X_1, V_1), \ldots, (X_k, V_k)$. Then, the unrestricted generalized Bayes estimator against the uniform prior $\mathrm{d}(\boldsymbol{\mu}, \sigma) = (\prod_{i=1}^k \mathrm{d}\mu_i)\mathrm{d}\sigma/\sigma$ is written as $\widehat{\boldsymbol{\theta}}^{M*} = \boldsymbol{A}^t \widehat{\boldsymbol{\mu}}^{M*}$, where

$$
\begin{aligned}
\widehat{\boldsymbol{\mu}}^{M*} =& \frac{\int \boldsymbol{a} b^{-2k-3} \prod_{i=1}^k f_i((X_i - a_i)/b, V_i/b)\mathrm{d}(\boldsymbol{a}, b)}{\int b^{-2k-3} \prod_{i=1}^k f_i((X_i - a_i)/b, V_i/b)\mathrm{d}(\boldsymbol{a}, b)} \\
=& \boldsymbol{X} - \frac{\int \boldsymbol{\xi} \tau^k \prod_{i=1}^k f_i(\xi_i, V_i \tau)\mathrm{d}(\boldsymbol{\xi}, \tau)}{\int \tau^{k+1} \prod_{i=1}^k f_i(\xi_i, V_i \tau)\mathrm{d}(\boldsymbol{\xi}, \tau)}.
\end{aligned}
$$

The minimaxity of $\widehat{\boldsymbol{\theta}}^{M*}$ can be verified based on the same arguments.