CIRJE-F-578

Minimaxity of the Stein Risk-Minimization Estimator for a Normal Mean Matrix

Tatsuya Kubokawa.
University of Tokyo

Hisayuki Tsukuma
Toho University

July 2008

Minimaxity of the Stein Risk-Minimization Estimator for a Normal Mean Matrix

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

July 16, 2008

Abstract

This paper addresses the Stein conjecture in the simultaneous estimation of a matrix mean of a multivariate normal distribution with a known covariance matrix. Stein (1973) derived an unbiased estimator of a risk function for orthogonally equivariant estimators and considered to isotonize the estimator which minimizes the main part of the unbiased risk-estimator. We call it the Stein risk-minimization estimator (RM) in this paper. Although the Stein RM estimator has been recognized as an excellent procedure with a nice risk-performance, it has a complicated form based on the isotonizing algorithm, and no analytical properties such as minimaxity have been shown. The aim of this paper is to fix this conjecture in lower dimensional cases, that is, the minimaxity of the Stein RM estimator is established for the two and three dimensions.

Key words and phrases: Decision theory, isotonic regression, Stein estimator, minimaxity, quadratic loss, simultaneous estimation, unbiased estimate of risk.

1 Introduction

For $i = 1, \ldots, p$ and $j = 1, \ldots, m$, let x_{ij} be mutually independent random variables. Suppose that $p \leq m$ and that x_{ij}'s are distributed as the normal distribution with mean θ_{ij} and variance one, respectively. The simultaneous estimation of θ_{ij}'s is then considered under sum of the quadratic loss functions, $\sum_{i=1}^{p}\sum_{j=1}^{m}(\delta_{ij}-\theta_{ij})^2$, where δ_{ij}'s are, respectively, certain estimators of θ_{ij}'s. The estimation problem is written in the matrix form as

$$\boldsymbol{X} \sim \mathcal{N}_{p\times m}(\boldsymbol{\Theta}, \boldsymbol{I}_p \otimes \boldsymbol{I}_m), \tag{1.1}$$

where $\boldsymbol{X} = (x_{ij})$, $\boldsymbol{\Theta} = (\theta_{ij})$, $\boldsymbol{I}_k$ is the identity matrix of order k, and $\otimes$ means the Kronecker product. The notation $\mathcal{N}_{p\times m}(\boldsymbol{\Theta}, \boldsymbol{I}_p \otimes \boldsymbol{I}_m)$ denotes the matrix-variate normal

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

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

distribution with mean matrix $\boldsymbol{\Theta}$ and the identity covariance matrix $\boldsymbol{I}_p \otimes \boldsymbol{I}_m$. Then the above estimation is expressed as the problem of estimating the mean matrix $\boldsymbol{\Theta}$ under the quadratic loss function (the Frobenius norm)

$$\|\boldsymbol{\delta} - \boldsymbol{\Theta}\|^2 = \mathrm{tr}\,(\boldsymbol{\delta} - \boldsymbol{\Theta})(\boldsymbol{\delta} - \boldsymbol{\Theta})^t, \tag{1.2}$$

where $\boldsymbol{\delta} = (\delta_{ij})$, and $\mathrm{tr}\,\boldsymbol{A}$ and $\boldsymbol{A}^t$ denote the trace and the transpose of a square matrix $\boldsymbol{A}$, respectively.

A natural class of estimators of $\boldsymbol{\Theta}$ is one of orthogonally equivariant estimators. Denote by $\mathcal{O}_k$ the group of k-dimensionally orthogonal matrices and by $\mathcal{V}_{m,p}$ the Stiefel manifold, namely, $\mathcal{V}_{m,p} = \{\boldsymbol{V} \in \boldsymbol{R}^{m\times p} | \boldsymbol{V}^t\boldsymbol{V} = \boldsymbol{I}_p\}$. Then, $\boldsymbol{X}$ is expressed as $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{L}\boldsymbol{V}^t$ by the singular value decomposition, where $\boldsymbol{U} \in \mathcal{O}_p$, $\boldsymbol{V} \in \mathcal{V}_{m,p}$ and $\boldsymbol{L} = \mathrm{diag}\,(l_1, \ldots, l_p)$ with $l_1 > \cdots > l_p > 0$. When the orthogonal transformation $\boldsymbol{X} \to \boldsymbol{P}\boldsymbol{X}\boldsymbol{Q}$ is considered for any $\boldsymbol{P} \in \mathcal{O}_p$ and any $\boldsymbol{Q} \in \mathcal{O}_m$, the estimator equivariant under the orthogonal transformations can be represented as

$$\boldsymbol{\delta} = \boldsymbol{U}\boldsymbol{\Psi}(\boldsymbol{L})\boldsymbol{V}^t, \tag{1.3}$$

where $\boldsymbol{\Psi}(\boldsymbol{L}) = \mathrm{diag}\,(\psi_1(\boldsymbol{L}), \ldots, \psi_p(\boldsymbol{L}))$, a diagonal matrix whose i-th diagonal element $\psi_i(\boldsymbol{L})$ is a function of $\boldsymbol{L}$. Let

$$\boldsymbol{\Lambda} = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_p), \quad \lambda_1 > \ldots, \lambda_p, \quad \text{for} \quad \lambda_i = l_i^2,$$

and let $\boldsymbol{\Psi}(\boldsymbol{L}) = \boldsymbol{L}\{(\boldsymbol{I}_p - \boldsymbol{\Phi}(\boldsymbol{\Lambda})\}$. Then, a conventional form of shrinkage estimators is given by

$$\boldsymbol{\delta}(\boldsymbol{\Phi}) = \boldsymbol{X} - \boldsymbol{U}\boldsymbol{L}\boldsymbol{\Phi}(\boldsymbol{\Lambda})\boldsymbol{V}^t = \boldsymbol{U}\boldsymbol{L}\{\boldsymbol{I}_p - \boldsymbol{\Phi}(\boldsymbol{\Lambda})\}\boldsymbol{V}^t, \tag{1.4}$$

where $\boldsymbol{\Phi}(\boldsymbol{\Lambda}) = \mathrm{diag}\,(\phi_1(\boldsymbol{\Lambda}), \ldots, \phi_p(\boldsymbol{\Lambda}))$.

A powerful tool for finding a minimax estimator is the use of an unbiased estimator of risk function of the estimator $\boldsymbol{\delta}(\boldsymbol{\Phi})$. Using the so-called Stein identity of a normal distribution, Stein (1973) showed that the risk function of $\boldsymbol{\delta}(\boldsymbol{\Phi})$ can be expressed as $R(\boldsymbol{\delta}(\boldsymbol{\Phi}), \boldsymbol{\Theta}) = mp + E[\widehat{\Delta}]$, where

$$\widehat{\Delta} = \sum_{i=1}^p \Big\{\lambda_i\phi_i^2 - 2c_0\phi_i - 4\sum_{j\neq i}\frac{\lambda_i}{\lambda_i - \lambda_j}\phi_i - 4\frac{\partial}{\partial\lambda_i}(\lambda_i\phi_i)\Big\}, \tag{1.5}$$

for $c_0 = m - p - 1$. This means that $mp + \widehat{\Delta}$ is an unbiased estimator of the risk. Since $\boldsymbol{X}$ is a minimax estimator with the constant risk mp, the orthogonally equivariant estimator $\boldsymbol{\delta}(\boldsymbol{\Phi})$ is minimax if $\boldsymbol{\Phi}$ satisfies that $\widehat{\Delta} \leq 0$. Two representative examples of the orthogonally equivariant and minimax estimators are the Efron-Morris (1972) estimator $\boldsymbol{\delta}^{EM} = \boldsymbol{\delta}(\boldsymbol{\Phi}^{EM})$ for $\boldsymbol{\Phi}^{EM} = \mathrm{diag}\,(\phi_1^{EM}, \ldots, \phi_p^{EM})$ with $\phi_i^{EM} = c_0/\lambda_i$, and the Stein (1973) estimator $\boldsymbol{\delta}^{ST} = \boldsymbol{\delta}(\boldsymbol{\Phi}^{ST})$ for $\boldsymbol{\Phi}^{ST} = \mathrm{diag}\,(\phi_1^{ST}, \ldots, \phi_p^{ST})$ with $\phi_i^{ST} = (m + p - 2i - 1)/\lambda_i$. When we ignore the term $-4\phi_i\sum_{j\neq i}\lambda_i/(\lambda_i - \lambda_j) - 4(\partial/\partial\lambda_i)(\lambda_i\phi_i)$ in $\widehat{\Delta}$, the optimal ϕ_i is given by $\phi_i = c_0/\lambda_i$, which yields the Efron-Morris estimator. The Stein estimator modifies the constant c_0 in ϕ_i^{EM} as $(m + p - 2i - 1)$, and this modification leads to the improvement of $\boldsymbol{\delta}^{ST}$ upon $\boldsymbol{\delta}^{EM}$ as stated by Stein (1973). It is noted that the Stein estimator can be derived by incorporating a part of the term $-4\phi_i \sum_{j\neq i} \lambda_i/(\lambda_i - \lambda_j)$, since $\widehat{\Delta}$ can be rewritten as

$$\widehat{\Delta} = \sum_{i=1}^{p} \Big\{ \lambda_i \phi_i^2 - 2(m+p-2i-1)\phi_i - 4\sum_{j>i} \frac{\lambda_j(\phi_i - \phi_j)}{\lambda_i - \lambda_j} - 4\frac{\partial}{\partial \lambda_i}(\lambda_i \phi_i) \Big\}.$$

Another interesting idea of Stein (1973) is to minimize the risk-unbiased estimator $\widehat{\Delta}$ with respect to ϕ_i with incorporating the whole information in $-4\phi_i \sum_{j\neq i} \lambda_i/(\lambda_i - \lambda_j)$. From (1.5), the minimizing function ϕ_i is provided by

$$\phi_i^{RM} = \frac{c_0}{\lambda_i} + 2\sum_{j\neq i} \frac{1}{\lambda_i - \lambda_j}, \tag{1.6}$$

for $c_0 = m - p - 1$. The resulting estimator is denoted by $\boldsymbol{\delta}^{RM} = \boldsymbol{\delta}(\boldsymbol{\Phi}^{RM})$ for $\boldsymbol{\Phi}^{RM} = \mathrm{diag}\,(\phi_1^{RM}(\boldsymbol{\Lambda}), \ldots, \phi_p^{RM}(\boldsymbol{\Lambda}))$. Although $\boldsymbol{\delta}^{RM}$ is expected to possess nice risk properties, it has two shortcomings: One is that the inequality $\widehat{\Delta} \leq 0$ for ϕ_i^{RM}'s does not hold. In fact, it can be shown that $\widehat{\Delta} > 0$ with a positive probability as stated in Proposition 2.1. This means that the minimaxity of $\boldsymbol{\delta}^{RM}$ cannot be guaranteed by the approach of the risk unbiased estimation. The other shortcoming is that although the shrinkage functions ϕ_i's should possess the natural ordering $\phi_1 \leq \ldots \leq \phi_p$ as a desirable property, this ordering is not always satisfied for the minimizing functions ϕ_i^{RM}'s.

To fix the second shortcoming, we use the isotonizing method. When the natural ordering is violated, we perform the isotonizing algorithm by pooling the adjacent pairs $(\lambda_i, \lambda_i\phi_i^{RM})$ like

$$\phi_i^{RM*} = \phi_{i+1}^{RM*} = \cdots = \phi_{i+k}^{RM*} = \frac{\lambda_i\phi_i^{RM} + \lambda_{i+1}\phi_{i+1}^{RM} + \cdots + \lambda_{i+k}\phi_{i+k}^{RM}}{\lambda_i + \lambda_{i+1} + \cdots + \lambda_{i+k}} \equiv \overline{\phi}_{i,i+k}^{RM}.$$

See Stein (1977) and Robertson, Wright and Dykstra (1988) and see also Lin and Perlman (1985) for details of the algorithm. It is also noted that ϕ_i^{RM*} can be obtained from the solution g_i of minimizing $\sum_{i=1}^{p}(\phi_i^{RM} - g_i)^2\lambda_i$ subject to the restriction $0 \leq g_1 \leq \ldots \leq g_p$. This gives another expression of the isotonizing function ϕ_i^{RM*} as

$$\phi_i^{RM*} = \min_{b\geq i} \max_{a\leq i} \left\{ \frac{\sum_{a\leq j\leq b} \lambda_j \phi_j^{RM}}{\sum_{a\leq j\leq b} \lambda_j} \right\}. \tag{1.7}$$

The isotonizing functions ϕ_i^{RM*}'s satisfy the natural ordering $\phi_1^{RM*} \leq \ldots \leq \phi_p^{RM*}$. Then, the estimator modified by the isotonization is given by

$$\boldsymbol{\delta}^{RM*} = \boldsymbol{\delta}(\boldsymbol{\Phi}^{RM*}) = \boldsymbol{X} - \boldsymbol{UL}\boldsymbol{\Phi}^{RM*}(\boldsymbol{\Lambda})\boldsymbol{V}^t, \tag{1.8}$$

where $\boldsymbol{\Phi}^{RM*}(\boldsymbol{\Lambda}) = \mathrm{diag}\,(\phi_1^{RM*}(\boldsymbol{\Lambda}), \ldots, \phi_p^{RM*}(\boldsymbol{\Lambda}))$. We call, in this paper, $\boldsymbol{\delta}^{RM*}$ **the Stein risk-minimization (RM) estimator**. Then it is quite interesting to consider the problem of showing the minimaxity of the Stein RM estimator $\boldsymbol{\delta}^{RM*}$. This is a conjecture given by Stein (1973), and has been suggested from numerical investigations.

However, no analytical proof has been provided for the minimaxity, because the estimator is so complicated that it is very hard to evaluate the risk function.

The aim of this paper is to challenge this conjecture. We can establish the minimaxity of the Stein RM estimator $\boldsymbol{\delta}^{RM*}$ in the lower dimensional cases, namely, it is minimax for $c_0 \geq 1$ in the case $p = 2$ and for $c_0 \geq 2$ in the case $p = 3$. Although the general dimensional case p is too hard to handle, Proposition 2.2 given in Section 2 suggests that the condition $c_0 \geq p - 1$ may be imposed on the minimaxity. All the proofs are given in Section 3. The risk performances of the estimators are numerically investigated in Section 4, and it is shown that the positive-part Stein RM estimator is better than other competitive estimators in most cases.

2 Minimaxity of the Stein risk-minimization estimator

In this section, we state the main results concerning the minimaxity of the Stein risk-minimization estimator $\boldsymbol{\delta}^{RM*}$. Since the function $\widehat{\Delta}$ can be derived by using an integration by parts called the Stein-Haff identity, the functions ϕ_i's need to be absolutely continuous. It is noted that the non-order preserving functions ϕ_i^{RM}'s and the isotonizing functions ϕ_i^{RM*}'s are absolutely continuous. Thus, for the minimaxity of $\boldsymbol{\delta}^{RM*}$, it is sufficient to show that $\widehat{\Delta} \leq 0$ for all $\boldsymbol{\Lambda}$, where $\widehat{\Delta}$ is defined in (1.5).

We first show the following proposition which gives the expression of $\widehat{\Delta}$ for the non-order preserving functions ϕ_i^{RM}'s.

Proposition 2.1 *The estimator of the risk difference $\widehat{\Delta}$ for the estimator $\boldsymbol{\delta}^{RM}$ with the non-order preserving functions ϕ_i^{RM}'s is given by*

$$\widehat{\Delta} = -c_0^2 \sum_{i=1}^{p} \frac{1}{\lambda_i} + 4 \sum_{i=1}^{p} \lambda_i \left(\sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} \right)^2, \tag{2.1}$$

which can take positive values with a positive probability.

Proposition 2.1 shows that $\widehat{\Delta}$ for $\boldsymbol{\delta}^{RM}$ is not always negative, namely, the approach based on the risk unbiased estimator cannot guarantee the minimaxity of $\boldsymbol{\delta}^{RM}$. No analytical properties have been studied about whether $\boldsymbol{\delta}^{RM}$ is minimax or not. It may be difficult to resolve this problem because one needs to evaluate the expected value $E[\widehat{\Delta}]$ for $\boldsymbol{\delta}^{RM}$. When we consider the estimator $\boldsymbol{\delta}^{RM*}$ with the isotonizing functions ϕ_i^{RM*}'s instead of $\boldsymbol{\delta}^{RM}$, however, we can show the minimaxity, namely, $\widehat{\Delta} \leq 0$ for $\boldsymbol{\delta}^{RM*}$ in lower dimensional cases.

Our method for the proof of the minimaxity is to decompose the space of $\boldsymbol{\Lambda}$ into several subsets corresponding to the forms of $(\phi_1^{RM*}, \ldots, \phi_p^{RM*})$, and to show that $\widehat{\Delta} \leq 0$ on each decomposed subset. For example, in the case of $p = 2$, it is seen that $(\phi_1^{RM*}, \phi_2^{RM*}) = (\phi_1^{RM}, \phi_2^{RM})$ on the subset $\{\phi_1^{RM} \leq \phi_2^{RM}\}$, and $\phi_1^{RM*} = \phi_2^{RM*} = (\lambda_1\phi_1^{RM} + \lambda_2\phi_2^{RM})/(\lambda_1 + \lambda_2)$ on the subset $\{\phi_1^{RM} > \phi_2^{RM}\}$. Although this method works in lower dimensional cases, it is intractable for higher dimensions. However, it is possible to evaluate $\widehat{\Delta}$ on the subset of $\phi_1^{RM} \leq \ldots \leq \phi_p^{RM}$.

Proposition 2.2 *On the subset that ϕ_i^{RM}'s satisfy the order restriction $\phi_1^{RM} \leq \ldots \leq \phi_p^{RM}$, the estimator of the risk difference $\widehat{\Delta}$ for the Stein risk-minimization estimator $\boldsymbol{\delta}^{RM*}$ is evaluated above as*

$$\widehat{\Delta} \leq (p-1-c_0)\sum_{i=1}^{p} \phi_i^{RM} = (p-1-c_0)c_0 \sum_{i=1}^{p} \frac{1}{\lambda_i}.$$

Proposition 2.2 suggests that we need to assume the condition $c_0 \geq p-1$, namely $m \geq 2p$, for the minimaxity of the Stein minimization estimator. In the lower dimensional cases of $p=2$ and $p=3$, it can be verified that the condition $c_0 \geq p-1$ is sufficient for the minimaxity.

Proposition 2.3 *For $c_0 \geq 1$, the Stein risk-minimization estimator $\boldsymbol{\delta}^{RM*}$ is minimax for $p=2$.*

Proposition 2.4 *For $c_0 \geq 2$, the Stein risk-minimization estimator $\boldsymbol{\delta}^{RM*}$ is minimax for $p=3$.*

3 Proof of the minimaxity

In this section, we shall give proofs of Propositions 2.1-2.4. For simplicity, we use here the notations ϕ_i and ϕ_i^* instead of ϕ_i^{RM} and ϕ_i^{RM*}, respectively. Also let

$$F_i = \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j},$$

and ϕ_i and $\lambda_i\phi_i$ are expressed as $\phi_i = c_0/\lambda_i + 2F_i$ and $\lambda_i\phi_i = c_0 + 2\lambda_i F_i$.

3.1 Preliminary lemmas

To prove the main results, we prepare several lemmas.

Lemma 3.1 (1) *If $\phi_i \leq \phi_j$ and $\lambda_i > \lambda_j$, then*

$$\frac{4}{\lambda_i - \lambda_j}(\lambda_i F_i - \lambda_j F_j) = \frac{2}{\lambda_i - \lambda_j}(\lambda_i \phi_i - \lambda_j \phi_j) \leq \phi_i + \phi_j. \tag{3.1}$$

(2) *Let $\sum_{i,j,k}^*$ denote the summation over $\{(i,j,k) | i \neq j, j \neq k, k \neq i, i \in J_p, j \in J_p, k \in J_p\}$ for $J_p = \{1, 2, \ldots, p\}$. Then for $p \geq 3$,*

$$\sum\nolimits_{i,j,k}^* \frac{\lambda_i}{(\lambda_i - \lambda_j)(\lambda_i - \lambda_k)} = 0.$$

(3) $\sum_{i=1}^{p} F_i = 0$.

Proof. For (1), the inequality (3.1) is equivalent to

$$(\lambda_i + \lambda_j)(\phi_i - \phi_j) \le 0, \quad \lambda_i > \lambda_j,$$

both of which are satisfied by the conditions.

For the proof of (2), we first note that for constants D_{ij}'s,

$$\begin{aligned}\sum_{i=1}^{p}\sum_{j\neq i} D_{ij} =& \sum_{i=1}^{p}\sum_{j=1}^{i-1} D_{ij} + \sum_{i=1}^{p}\sum_{j=i+1}^{p} D_{ij}\\ =& \sum_{i=1}^{p}\sum_{j=1}^{i-1} \left(D_{ij} + D_{ji}\right),\end{aligned} \tag{3.2}$$

since $\sum_{i=1}^{p}\sum_{j=i+1}^{p} D_{ij} = \sum_{j=1}^{p}\sum_{i=1}^{j-1} D_{ij} = \sum_{i=1}^{p}\sum_{j=1}^{i-1} D_{ji}$. Let $I = \sum_{i,j,k}^{*} \lambda_i/\{(\lambda_i - \lambda_j)(\lambda_i - \lambda_k)\}$ and $D_{ij} = \sum_{k=1,k\neq i,k\neq j}^{p} \lambda_i/\{(\lambda_i - \lambda_j)(\lambda_i - \lambda_k)\}$. Then, I is written as $I = \sum_{i=1}^{p}\sum_{j\neq i} D_{ij}$, and we shall show that $I = 0$. From the identity (3.2), I is written as $I = \sum_{i=1}^{p}\sum_{j=1}^{i-1}(D_{ij} + D_{ji})$, and $D_{ij} + D_{ji}$ is rewritten as

$$\begin{aligned}D_{ij} + D_{ji} =& \sum_{k=1,k\neq i,k\neq j}^{p} \left\{\frac{\lambda_i}{(\lambda_i - \lambda_j)(\lambda_i - \lambda_k)} + \frac{\lambda_j}{(\lambda_j - \lambda_i)(\lambda_j - \lambda_k)}\right\}\\ =& - \sum_{k=1,k\neq i,k\neq j}^{p} \frac{\lambda_k}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)}.\end{aligned}$$

Hence,

$$\begin{aligned}I =& - \sum_{i=1}^{p}\sum_{j=1}^{i-1}\sum_{k=1,k\neq i,k\neq j}^{p} \frac{\lambda_k}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)}\\ =& - \frac{1}{2}\sum\nolimits_{i,j,k}^{*} \frac{\lambda_k}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)}\\ =& - I/2,\end{aligned}$$

which means that $I = 0$.

For the proof of (3), from the identity (3.2), it is observed that

$$\begin{aligned}\sum_{i=1}^{p} F_i =& \sum_{i=1}^{p}\sum_{j\neq i} \frac{1}{\lambda_i - \lambda_j}\\ =& \sum_{i=1}^{p}\sum_{j=1}^{i-1} \left\{\frac{1}{\lambda_i - \lambda_j} + \frac{1}{\lambda_j - \lambda_i}\right\},\end{aligned}$$

which is equal to zero. Therefore the proof of Lemma 3.1 is complete. ■

Lemma 3.2 (1) $\frac{\partial}{\partial \lambda_i}(\lambda_i \phi_i) = 2F_i - 2\lambda_i F_i^2 + 2 \sum_{j=1,j\neq i}^{p} \sum_{k=1,k\neq i,k\neq j}^{p} \frac{\lambda_i}{(\lambda_i - \lambda_j)(\lambda_i - \lambda_k)}$

(2) $\lambda_i \phi_i^2 = c_0 \phi_i + 2c_0 F_i + 4\lambda_i F_i^2$

(3) *Let* $\widehat{\Delta}_i = \lambda_i \phi_i^2 - 2c_0\phi_i - 4\frac{\partial}{\partial \lambda_i}(\lambda_i \phi_i) - 4\lambda_i \phi_i F_i$. *Then,* $\widehat{\Delta}_i$ *is expressed as*

$$\begin{aligned}\widehat{\Delta}_i =& - \lambda_i \phi_i^2 - 4\frac{\partial}{\partial \lambda_i}(\lambda_i \phi_i) \\ =& - c_0\phi_i - 2(c_0+4)F_i + 4\lambda_i F_i^2 - 8 \sum_{j=1,j\neq i}^{p} \sum_{k=1,k\neq i,k\neq j}^{p} \frac{\lambda_i}{(\lambda_i - \lambda_j)(\lambda_i - \lambda_k)}.\end{aligned} \tag{3.3}$$

Proof. Since $\lambda_i \phi_i = c_0 + 2\lambda_i F_i$, it is seen that

$$\frac{\partial}{\partial \lambda_i}(\lambda_i \phi_i) = 2F_i + 2\lambda_i \frac{\partial}{\partial \lambda_i} F_i.$$

and $(\partial/\partial\lambda_i)F_i$ is expressed as

$$\begin{aligned}\frac{\partial}{\partial \lambda_i} F_i =& - \sum_{j\neq i} \frac{1}{(\lambda_i - \lambda_j)^2} \\ =& - \left(\sum_{j\neq i} \frac{1}{\lambda_i - \lambda_j} \right)^2 + \sum_{j=1,j\neq i}^{p} \sum_{k=1,k\neq i,k\neq j}^{p} \frac{1}{(\lambda_i - \lambda_j)(\lambda_i - \lambda_k)},\end{aligned} \tag{3.4}$$

which proves (1). For (2), it is observed that $\lambda_i \phi_i^2 = \phi_i(c_0 + 2\lambda_i F_i) = c_0\phi_i + 2F_i(c_0 + 2\lambda_i F_i)$, which yields the r.h.s. of (2). For the proof of (3), completing square with respect to ϕ_i gives that

$$\begin{aligned}\widehat{\Delta}_i =& \lambda_i \left\{ \phi_i^2 - 2(c_0/\lambda_i + 2F_i)\phi_i \right\} - 4\frac{\partial}{\partial \lambda_i}(\lambda_i \phi_i) \\ =& \lambda_i \left\{ \phi_i - (c_0/\lambda_i + 2F_i) \right\}^2 - \lambda_i (c_0/\lambda_i + 2F_i)^2 - 4\frac{\partial}{\partial \lambda_i}(\lambda_i \phi_i),\end{aligned} \tag{3.5}$$

which yields the first equality in (3.3). The second equality in (3.3) can be obtained by using (1) and (2) of Lemma 3.2, and the proof is complete. ■

3.2 Proof of Proposition 2.1

From Lemma 3.2, $\widehat{\Delta}$ for $\boldsymbol{\delta}^{RM}$ is written as

$$\begin{aligned}\widehat{\Delta} =& \sum_{i=1}^{p} \widehat{\Delta}_i \\ =& \sum_{i=1}^{p} \left\{ -c_0\phi_i - 2(c_0+4)F_i + 4\lambda_i F_i^2 - 8 \sum_{j=1,j\neq i}^{p} \sum_{k=1,k\neq i,k\neq j}^{p} \frac{\lambda_i}{(\lambda_i - \lambda_j)(\lambda_i - \lambda_k)} \right\}.\end{aligned}$$

Using the equalities given in (2) and (3) of Lemma 3.1, we can rewrite it as

$$\widehat{\Delta} = -c_0^2 \sum_{i=1}^{p} \lambda_i^{-1} + 4 \sum_{i=1}^{p} \lambda_i F_i^2,$$

which is equal to (2.1). Let $F_i^* = \sum_{j=3}^{p} (\lambda_i - \lambda_j)^{-1}$ for $i = 1, 2$. Then, it is observed that

$$\begin{aligned}
\sum_{i=1}^{2} \lambda_i F_i^2 =& \lambda_1 \left(\frac{1}{\lambda_1 - \lambda_2} + F_1^* \right)^2 + \lambda_2 \left(\frac{1}{\lambda_1 - \lambda_2} - F_2^* \right)^2 \\
=& \frac{\lambda_1 + \lambda_2}{(\lambda_1 - \lambda_2)^2} + \frac{2}{\lambda_1 - \lambda_2} (\lambda_1 F_1^2 - \lambda_2 F_2^*) + \lambda_1 (F_1^*)^2 + \lambda_2 (F_2^*)^2 \\
=& \frac{\lambda_1 + \lambda_2}{(\lambda_1 - \lambda_2)^2} - 2 \sum_{j=3}^{p} \frac{\lambda_j}{(\lambda_1 - \lambda_j)(\lambda_2 - \lambda_j)} + \lambda_1 (F_1^*)^2 + \lambda_2 (F_2^*)^2,
\end{aligned}$$

so that $\widehat{\Delta}$ is expressed as

$$\begin{aligned}
\widehat{\Delta} =& - c_0^2 \sum_{i=1}^{p} \lambda_i^{-1} + 4 \frac{\lambda_1 + \lambda_2}{(\lambda_1 - \lambda_2)^2} - 8 \sum_{j=3}^{p} \frac{\lambda_j}{(\lambda_1 - \lambda_j)(\lambda_2 - \lambda_j)} \\
&+ 4\lambda_1 (F_1^*)^2 + 4\lambda_2 (F_2^*)^2 + \sum_{i=3}^{p} \lambda_i F_i^2.
\end{aligned}$$

This expression means that $\widehat{\Delta}$ tends to infinity as $\lambda_1 - \lambda_2 \to 0$. Hence, $\widehat{\Delta} > 0$ with a positive probability. ■

3.3 Proof of Proposition 2.2

Using (3) of Lemma 3.2, we observe that

$$\begin{aligned}
\widehat{\Delta} =& - c_0 \sum_{i=1}^{p} \phi_i - 2(c_0 + 4) \sum_{i=1}^{p} F_i + 4 \sum_{i=1}^{p} \lambda_i F_i^2 - 8 {\sum}^*_{i,j,k} \frac{\lambda_i}{(\lambda_i - \lambda_j)(\lambda_i - \lambda_k)} \\
=& - c_0 \sum_{i=1}^{p} \phi_i + 4 \sum_{i=1}^{p} \lambda_i F_i^2, \qquad (3.6)
\end{aligned}$$

where the second equality follows from (2) and (3) of Lemma 3.1. Using the identity similar to (3.2) gives that

$$\begin{aligned}
\sum_{i=1}^{p} \lambda_i F_i^2 =& \sum_{i=1}^{p} \lambda_i F_i \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} = \sum_{i=1}^{p} \sum_{j \neq i} \frac{\lambda_i F_i}{\lambda_i - \lambda_j} \\
=& \sum_{i=1}^{p} \sum_{j=i+1}^{p} \left\{ \frac{\lambda_i F_i}{\lambda_i - \lambda_j} + \frac{\lambda_j F_j}{\lambda_j - \lambda_i} \right\} \\
=& \sum_{i=1}^{p} \sum_{j=i+1}^{p} \frac{1}{\lambda_i - \lambda_j} (\lambda_i F_i - \lambda_j F_j).
\end{aligned}$$

Since $i < j$, it is noted that $\lambda_i \geq \lambda_j$ and $\phi_i \leq \phi_j$. We thus use the inequality (3.1) to obtain that

$$4\sum_{i=1}^{p} \lambda_i F_i^2 \leq \sum_{i=1}^{p} \sum_{j=i+1}^{p} (\phi_i + \phi_j) = \sum_{i=1}^{p} (p-1)\phi_i.$$

From (3.6), we can see that

$$\widehat{\Delta} \leq (p-1-c_0)\sum_{i=1}^{p} \phi_i = (p-1-c_0)c_0 \sum_{i=1}^{p} \frac{1}{\lambda_i},$$

which proves Proposition 2.2. ■

3.4 Proof of Proposition 2.3

We first treat the simple case $p = 2$ and prove the minimaxity of the Stein minimization estimator $\boldsymbol{\delta}^{RM*}$. For $p = 2$, ϕ_i^*'s are given as follows: $(\phi_1^*, \phi_2^*) = (\phi_1, \phi_2)$ on the set $\{\phi_1 \leq \phi_2\}$, and $\phi_1^* = \phi_2^* = (\lambda_1\phi_1 + \lambda_2\phi_2)/(\lambda_1 + \lambda_2)$ on the set $\{\phi_1 > \phi_2\}$. On the set of $\phi_1 \leq \phi_2$, Proposition 2.2 implies that $\widehat{\Delta} \leq 0$ for $c_0 \geq 1$. Hence, we shall show that $\widehat{\Delta} \leq 0$ on the set of $\phi_1 > \phi_2$. Let $\overline{\phi}_{12} = (\lambda_1\phi_1 + \lambda_2\phi_2)/(\lambda_1 + \lambda_2)$. Then, it can be written as $\overline{\phi}_{12} = 2(c_0 + \lambda_1 F_1 + \lambda_2 F_2)/(\lambda_1 + \lambda_2) = 2(c_0 + 1)/(\lambda_1 + \lambda_2)$. Since $\phi_1^* = \phi_2^* = \overline{\phi}_{12}$, (1.5) gives the expression

$$\begin{aligned}
\widehat{\Delta} =& (\lambda_1 + \lambda_2)\overline{\phi}_{12}^2 - 4c_0\overline{\phi}_{12} - 4\sum_{i=1}^{2} \frac{\partial}{\partial \lambda_i}(\lambda_i \overline{\phi}_{12}) - 4\overline{\phi}_{12} \\
=& (\lambda_1 + \lambda_2)\left\{\overline{\phi}_{12} - 2\frac{c_0 + 1}{\lambda_1 + \lambda_2}\right\}^2 - \frac{4(c_0+1)^2}{\lambda_1 + \lambda_2} - 4\sum_{i=1}^{2} \frac{\partial}{\partial \lambda_i}(\lambda_i \overline{\phi}_{12}) \\
=& -(\lambda_1 + \lambda_2)\overline{\phi}_{12}^2 - 4\sum_{i=1}^{2} \frac{\partial}{\partial \lambda_i}(\lambda_i \overline{\phi}_{12}).
\end{aligned} \tag{3.7}$$

Noting that $\sum_{i=1}^{2}(\partial/\partial\lambda_i)(\lambda_i\overline{\phi}_{12}) = 2(c_0 + 1)/(\lambda_1 + \lambda_2) = \overline{\phi}_{12}$, we can see that $\widehat{\Delta} = -(\lambda_1 + \lambda_2)\overline{\phi}_{12}^2 - 4\overline{\phi}_{12}$, which is not positive. Therefore, Proposition 2.3 is proved. ■

3.5 Proof of Proposition 2.4

We now handle the case of $p = 3$ and prove the minimaxity of the Stein minimization estimator $\boldsymbol{\delta}^{RM*}$. The isotonic functions ϕ_i^*'s take the following four cases:

(C1) In the case that $\phi_1 \leq \phi_2 \leq \phi_3$, $(\phi_1^*, \phi_2^*, \phi_3^*) = (\phi_1, \phi_2, \phi_3)$.

(C2) In the case that $\phi_2 > \phi_3$ and $\phi_1 \leq \overline{\phi}_{23}$ for $\overline{\phi}_{23} = (\lambda_2\phi_2 + \lambda_3\phi_3)/(\lambda_2 + \lambda_3)$, $(\phi_1^*, \phi_2^*, \phi_3^*) = (\phi_1, \overline{\phi}_{23}, \overline{\phi}_{23})$.

(C3) In the case that $\phi_1 > \phi_2$ and $\overline{\phi}_{12} \leq \phi_3$ for $\overline{\phi}_{12} = (\lambda_1\phi_1 + \lambda_2\phi_2)/(\lambda_1 + \lambda_2)$, $(\phi_1^*, \phi_2^*, \phi_3^*) = (\overline{\phi}_{12}, \overline{\phi}_{12}, \phi_3)$.

(C4) In the cases that $\{\phi_2 > \phi_3$ and $\phi_1 > \overline{\phi}_{23}\}$ or $\{\phi_1 > \phi_2$ and $\overline{\phi}_{12} > \phi_3\}$, $\phi_1^* = \phi_2^* = \phi_3^* = (\lambda_1\phi_1 + \lambda_2\phi_2 + \lambda_3\phi_3)/(\lambda_1 + \lambda_2 + \lambda_3) = \overline{\phi}_{13}$.

The result in the case (C1) follows from Proposition 2.2. For the case (C4), it is noted that $\overline{\phi}_{13}$ is expressed as $\overline{\phi}_{13} = (3c_0 + 2\sum_{i=1}^{3} \lambda_i F_i)/\sum_{i=1}^{3} \lambda_i$. Noting $F_i = \sum_{j\neq i} 1/(\lambda_i - \lambda_j)$ and using the identity (3.2), we can see that

$$\begin{aligned}\sum_{i=1}^{3} \lambda_i F_i =& \sum_{i=1}^{3}\sum_{j\neq i} \frac{\lambda_i}{\lambda_i - \lambda_j}\\ =& \sum_{i=1}^{3}\sum_{j=1}^{i-1} \left\{ \frac{\lambda_i}{\lambda_i - \lambda_j} + \frac{\lambda_j}{\lambda_j - \lambda_i} \right\} = 3,\end{aligned}$$

so that $\overline{\phi}_{13} = 3(c_0 + 2)/\sum_{i=1}^{3} \lambda_i$. The same arguments as in (3.7) is used to rewrite $\widehat{\Delta}$ as

$$\begin{aligned}\widehat{\Delta} =& \sum_{i=1}^{3} \left\{ \lambda_i \overline{\phi}_{13}^2 - 2c_0\overline{\phi}_{13} - 4\frac{\partial}{\partial \lambda_i}(\lambda_i \overline{\phi}_{13}) - 4\lambda_i F_i \overline{\phi}_{13} \right\}\\ =& -\sum_{i=1}^{3} \lambda_i \overline{\phi}_{13}^2 - 4\sum_{i=1}^{3} \frac{\partial}{\partial \lambda_i}(\lambda_i \overline{\phi}_{13}).\end{aligned}$$

The derivative $(\partial/\partial\lambda_1)(\lambda_1\overline{\phi}_{13})$ is written as

$$\frac{\partial}{\partial \lambda_1}(\lambda_1 \overline{\phi}_{13}) = 3(c_0 + 2)\frac{\partial}{\partial \lambda_1}\left\{ \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \right\} = 3(c_0 + 2)\frac{\lambda_2 + \lambda_3}{(\lambda_1 + \lambda_2 + \lambda_3)^2},$$

which implies that

$$\sum_{i=1}^{3} \frac{\partial}{\partial \lambda_i}(\lambda_i \overline{\phi}_{13}) = 3(c_0 + 2)\frac{2}{\lambda_1 + \lambda_2 + \lambda_3} = 2\overline{\phi}_{13}.$$

Hence, $\widehat{\Delta}$ is expressed by

$$\widehat{\Delta} = -(\lambda_1 + \lambda_2 + \lambda_3)\overline{\phi}_{13}^2 - 8\overline{\phi}_{13},$$

which is not positive.

Finally, we shall show that $\widehat{\Delta} \leq 0$ in the two cases (C2) and (C3). Since both cases can be proved by the same arguments, it is sufficient to show it in the case (C2). In the case (C2), the unbiased risk estimator (1.5) gives $\widehat{\Delta} = \widehat{\Delta}_1 + \widehat{\Delta}_{23}$, where

$$\begin{aligned}\widehat{\Delta}_1 =& \lambda_1 \phi_1^2 - 2c_0\phi_1 - 4\frac{\partial}{\partial \lambda_1}(\lambda_1 \phi_1) - 4\lambda_1 F_1 \phi_1,\\ \widehat{\Delta}_{23} =& \sum_{i=2}^{3} \left\{ \lambda_i \overline{\phi}^2 - 2c_0\overline{\phi} - 4\frac{\partial}{\partial \lambda_i}(\lambda_i \overline{\phi}) - 4\lambda_i F_i \overline{\phi} \right\},\end{aligned}$$

where the simple notation $\overline{\phi}$ is used here instead of $\overline{\phi}_{23}$. From (3) of Lemma 3.2, it follows that

$$\begin{aligned}\widehat{\Delta}_1 =& -\lambda_1 \phi_1^2 - 4\frac{\partial}{\partial \lambda_1}(\lambda_1 \phi_1)\\ =& -c_0\phi_1 - 2(c_0 + 4)F_1 + 4\lambda_1 F_1^2 - \frac{16\lambda_1}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)}.\end{aligned} \tag{3.8}$$

A similar argument gives the expression

$$\begin{aligned}\widehat{\Delta}_{23} =&(\lambda_2+\lambda_3)\overline{\phi}^2 - 4c_0\overline{\phi} - 4\sum_{i=2}^{3}\lambda_i F_i\overline{\phi} - 4\sum_{i=2}^{3}\frac{\partial}{\partial\lambda_i}(\lambda_i\overline{\phi})\\ =&-(\lambda_2+\lambda_3)\overline{\phi}^2 - 4\sum_{i=2}^{3}\frac{\partial}{\partial\lambda_i}(\lambda_i\overline{\phi}).\end{aligned}$$

It is noted that

$$\begin{aligned}(\lambda_2+\lambda_3)\overline{\phi}^2 =&\overline{\phi}(2c_0+2\lambda_2F_2+2\lambda_3F_3)\\ =&2c_0\overline{\phi}+4c_0\frac{\lambda_2F_2+\lambda_3F_3}{\lambda_2+\lambda_3}+4\frac{(\lambda_2F_2+\lambda_3F_3)^2}{\lambda_2+\lambda_3},\end{aligned} \tag{3.9}$$

and

$$\begin{aligned}\sum_{i=2}^{3}\frac{\partial}{\partial\lambda_i}(\lambda_i\overline{\phi}) =&\overline{\phi}+2\frac{\sum_{i=2}^{3}\lambda_i(\partial/\partial\lambda_i)(\sum_{j=2}^{3}\lambda_jF_j)}{\lambda_2+\lambda_3}\\ =&\frac{2c_0}{\lambda_2+\lambda_3}+4\frac{\lambda_2F_2+\lambda_3F_3}{\lambda_2+\lambda_3}\\ &+\frac{2}{\lambda_2+\lambda_3}\left\{\sum_{i=2}^{3}\lambda_i^2\frac{\partial}{\partial\lambda_i}F_i+\lambda_2\lambda_3\Big(\frac{\partial}{\partial\lambda_2}F_3+\frac{\partial}{\partial\lambda_3}F_2\Big)\right\}.\end{aligned} \tag{3.10}$$

Combining (3.9) and (3.10), we can rewrite $\widehat{\Delta}_{23}$ as

$$\begin{aligned}\widehat{\Delta}_{23} =&-2c_0\overline{\phi}-4(c_0+4)\frac{\lambda_2F_2+\lambda_3F_3}{\lambda_2+\lambda_3}-\frac{8c_0}{\lambda_2+\lambda_3}-4\frac{(\lambda_2F_2+\lambda_3F_3)^2}{\lambda_2+\lambda_3}\\ &-\frac{8}{\lambda_2+\lambda_3}\left\{\sum_{i=2}^{3}\lambda_i^2\frac{\partial}{\partial\lambda_i}F_i+\lambda_2\lambda_3\Big(\frac{\partial}{\partial\lambda_2}F_3+\frac{\partial}{\partial\lambda_3}F_2\Big)\right\}.\end{aligned} \tag{3.11}$$

From (3.4), it is noted that

$$\begin{aligned}&\lambda_2^2\frac{\partial}{\partial\lambda_2}F_2+\lambda_3^2\frac{\partial}{\partial\lambda_3}F_3\\ &=-\lambda_2^2F_2^2-\lambda_3^2F_3^2+\frac{2\lambda_2^2}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)}+\frac{2\lambda_3^2}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)},\end{aligned} \tag{3.12}$$

and $(\partial/\partial\lambda_2)F_3 = (\partial/\partial\lambda_3)F_2 = 1/(\lambda_2-\lambda_3)^2$. Also note that

$$\begin{aligned}&\frac{2\lambda_2(\lambda_2+\lambda_3-\lambda_3)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)}+\frac{2\lambda_3(\lambda_3+\lambda_2-\lambda_2)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}\\ &\quad=(\lambda_2+\lambda_3)\left\{\frac{2\lambda_2}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)}+\frac{2\lambda_3}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}\right\}\\ &\qquad-2\lambda_2\lambda_3\left\{\frac{1}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)}+\frac{1}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}\right\}.\end{aligned} \tag{3.13}$$

Combining (3.12) and (3.13) yields that

$$
\begin{aligned}
&-4\frac{(\lambda_2F_2+\lambda_3F_3)^2}{\lambda_2+\lambda_3}-\frac{8}{\lambda_2+\lambda_3}\left\{\sum_{i=2}^{3}\lambda_i^2\frac{\partial}{\partial\lambda_i}F_i+\lambda_2\lambda_3\Big(\frac{\partial}{\partial\lambda_2}F_3+\frac{\partial}{\partial\lambda_3}F_2\Big)\right\}\\
&\quad=-\frac{16\lambda_2}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)}-\frac{16\lambda_3}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}+\frac{4(\lambda_2^2F_2^2+\lambda_3^2F_3^2-2\lambda_2\lambda_3F_2F_3)}{\lambda_2+\lambda_3}\\
&\qquad+\frac{16\lambda_2\lambda_3}{\lambda_2+\lambda_3}\left\{\frac{1}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)}+\frac{1}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}-\frac{1}{(\lambda_2-\lambda_3)^2}\right\}.
\end{aligned}
$$

Hence, substituting this expression into (3.11) and combining it with $\widehat{\Delta}_1$, we obtain the expression

$$
\begin{aligned}
\widehat{\Delta}=&-c_0\phi_1-2c_0\overline{\phi}-2(c_0+4)F_1-4(c_0+4)\frac{\lambda_2F_2+\lambda_3F_3}{\lambda_2+\lambda_3}\\
&-\frac{16\lambda_1}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)}-\frac{16\lambda_2}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)}-\frac{16\lambda_3}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}\\
&+4\lambda_1F_1^2-\frac{8c_0}{\lambda_2+\lambda_3}\\
&+\frac{4}{\lambda_2+\lambda_3}\left\{\lambda_2^2F_2^2+\lambda_3^3F_3^2-2\lambda_2\lambda_3F_2F_3-\frac{4\lambda_2\lambda_3}{\lambda_2-\lambda_3}\left(\frac{1}{\lambda_1-\lambda_2}+\frac{1}{\lambda_2-\lambda_3}+\frac{1}{\lambda_3-\lambda_1}\right)\right\}.
\end{aligned}
\tag{3.14}
$$

We first evaluate the term $-4(c_0+4)(\lambda_2F_2+\lambda_3F_3)/(\lambda_2+\lambda_3)$, which is rewritten as $-2(c_0+4)\{(\lambda_2\phi_2+\lambda_3\phi_3)/(\lambda_2+\lambda_3)-2c_0/(\lambda_2+\lambda_3)\}$. It is observed that the inequality

$$
2\frac{\lambda_2\phi_2+\lambda_3\phi_3}{\lambda_2+\lambda_3}\geq\phi_2+\phi_3 \tag{3.15}
$$

is equivalent to $(\lambda_2-\lambda_3)(\phi_2-\phi_3)\geq 0$. Since $\phi_2>\phi_3$ in the case (C2), we can use the inequality (3.15) to show that

$$
\begin{aligned}
-4(c_0+4)\frac{\lambda_2F_2+\lambda_3F_3}{\lambda_2+\lambda_3}&\leq-(c_0+4)\left\{\phi_2+\phi_3-\frac{4c_0}{\lambda_2+\lambda_3}\right\}\\
&=-2(c_0+4)(F_2+F_3)-(c_0+4)c_0\left\{\frac{1}{\lambda_2}+\frac{1}{\lambda_3}-\frac{4}{\lambda_2+\lambda_3}\right\}\\
&\leq-2(c_0+4)(F_2+F_3),
\end{aligned}
$$

since $1/\lambda_2+1/\lambda_3\geq 4/(\lambda_2+\lambda_3)$. Then, we can use (3) of Lemma 3.1 to obtain that

$$
-2(c_0+4)F_1-4(c_0+4)\frac{\lambda_2F_2+\lambda_3F_3}{\lambda_2+\lambda_3}\leq-2(c_0+4)(F_1+F_2+F_3)=0.
$$

From (2) of Lemma 3.1, it follows that

$$
-\frac{16\lambda_1}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)}-\frac{16\lambda_2}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)}-\frac{16\lambda_3}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}=0.
$$

Hence, $\widehat{\Delta}$ in (3.14) is evaluated as

$$\begin{aligned}\widehat{\Delta} \leq & - c_0\phi_1 - 2c_0\overline{\phi} + 4\lambda_1 F_1^2 - \frac{8c_0}{\lambda_2+\lambda_3} \\ & + \frac{4}{\lambda_2+\lambda_3}\left\{\lambda_2^2F_2^2 + \lambda_3^2F_3^2 - 2\lambda_2\lambda_3F_2F_3 - \frac{4\lambda_2\lambda_3}{\lambda_2-\lambda_3}\left(\frac{1}{\lambda_1-\lambda_2}+\frac{1}{\lambda_2-\lambda_3}+\frac{1}{\lambda_3-\lambda_1}\right)\right\}.\end{aligned} \tag{3.16}$$

We shall evaluate the last term in the r.h.s. of (3.16). It is observed that

$$\begin{aligned}&\lambda_2^2F_2^2 + \lambda_3^2F_3^2 - 2\lambda_2\lambda_3F_2F_3 \\ &= \left(\frac{\lambda_2}{\lambda_2-\lambda_1}+\frac{\lambda_2}{\lambda_2-\lambda_3}\right)\lambda_2F_2 + \left(\frac{\lambda_3}{\lambda_3-\lambda_1}+\frac{\lambda_3}{\lambda_3-\lambda_2}\right)\lambda_3F_3 - 2\lambda_2\lambda_3F_2F_3 \\ &= \lambda_2F_2 + \lambda_3F_3 + \frac{\lambda_2}{\lambda_1-\lambda_2}(\lambda_1F_1-\lambda_2F_2) + \frac{\lambda_3}{\lambda_1-\lambda_3}(\lambda_1F_1-\lambda_3F_3) \\ &\quad - \left(\frac{\lambda_2}{\lambda_1-\lambda_2}\lambda_1F_1 + \frac{\lambda_3}{\lambda_1-\lambda_3}\lambda_1F_1\right) \\ &\quad + \lambda_2\lambda_3\left(\frac{1}{\lambda_2-\lambda_3}F_2 - \frac{1}{\lambda_2-\lambda_3}F_3 - 2F_2F_3\right).\end{aligned} \tag{3.17}$$

The second last term in the r.h.s. of (3.17) is further rewritten as

$$\begin{aligned}&-\left(\frac{\lambda_2}{\lambda_1-\lambda_2}\lambda_1F_1 + \frac{\lambda_3}{\lambda_1-\lambda_3}\lambda_1F_1\right) \\ &= -\frac{(\lambda_2+\lambda_3)-\lambda_3}{\lambda_1-\lambda_2}\lambda_1F_1 - \frac{(\lambda_3+\lambda_2)-\lambda_2}{\lambda_1-\lambda_3}\lambda_1F_1 \\ &= -(\lambda_2+\lambda_3)\left(\frac{1}{\lambda_1-\lambda_2}+\frac{1}{\lambda_1-\lambda_3}\right)\lambda_1F_1 + \frac{\lambda_3}{\lambda_1-\lambda_2}\lambda_1F_1 + \frac{\lambda_2}{\lambda_1-\lambda_3}\lambda_1F_1 \\ &= -(\lambda_2+\lambda_3)\lambda_1F_1^2 + \frac{\lambda_3}{\lambda_1-\lambda_2}(\lambda_1F_1-\lambda_2F_2) + \frac{\lambda_2}{\lambda_1-\lambda_3}(\lambda_1F_1-\lambda_3F_3) \\ &\quad + \lambda_2\lambda_3\left(\frac{1}{\lambda_1-\lambda_2}F_2 + \frac{1}{\lambda_1-\lambda_3}F_3\right).\end{aligned} \tag{3.18}$$

It is noted that

$$-2F_2F_3 = -\left(\frac{1}{\lambda_3-\lambda_1}+\frac{1}{\lambda_3-\lambda_2}\right)F_2 - \left(\frac{1}{\lambda_2-\lambda_1}+\frac{1}{\lambda_2-\lambda_3}\right)F_3.$$

Then, combining the last terms in (3.17) and (3.18), we see that

$$\begin{aligned}&\left(\frac{1}{\lambda_2-\lambda_3}F_2 - \frac{1}{\lambda_2-\lambda_3}F_3 - 2F_2F_3\right) + \left(\frac{1}{\lambda_1-\lambda_2}F_2 + \frac{1}{\lambda_1-\lambda_3}F_3\right) \\ &= \left(\frac{1}{\lambda_1-\lambda_2}+\frac{1}{\lambda_1-\lambda_3}\right)F_2 + \left(\frac{1}{\lambda_1-\lambda_2}+\frac{1}{\lambda_1-\lambda_3}\right)F_3 + \frac{2}{\lambda_2-\lambda_3}(F_2-F_3) \\ &= F_1(F_2+F_3) + \frac{2}{\lambda_2-\lambda_3}(F_2-F_3) \\ &= -F_1^2 + \frac{2}{\lambda_2-\lambda_3}(F_2-F_3),\end{aligned} \tag{3.19}$$

since $F_1 + F_2 + F_3 = 0$. From (3.17), (3.18) and (3.19), $\lambda_2^2 F_2^2 + \lambda_3^2 F_3^2 - 2\lambda_2\lambda_3 F_2 F_3$ is expressed as

$$\begin{aligned}
&\lambda_2^2 F_2^2 + \lambda_3^2 F_3^2 - 2\lambda_2\lambda_3 F_2 F_3 \\
&\quad = (\lambda_2 F_2 + \lambda_3 F_3) - (\lambda_2 + \lambda_3)\lambda_1 F_1^2 \\
&\qquad + (\lambda_2 + \lambda_3)\left(\frac{1}{\lambda_1 - \lambda_2}(\lambda_1 F_1 - \lambda_2 F_2) + \frac{1}{\lambda_1 - \lambda_3}(\lambda_1 F_1 - \lambda_3 F_3)\right) \\
&\qquad + \lambda_2\lambda_3\left(-F_1^2 + \frac{2}{\lambda_2 - \lambda_3}(F_2 - F_3)\right).
\end{aligned} \tag{3.20}$$

Substituting (3.20) into (3.16), we obtain that

$$\begin{aligned}
\widehat{\Delta} \leq & - c_0\phi_1 - 2c_0\overline{\phi} - \frac{8c_0}{\lambda_2 + \lambda_3} + 2\frac{2\lambda_2 F_2 + 2\lambda_3 F_3}{\lambda_2 + \lambda_3} \\
& + \left(\frac{4}{\lambda_1 - \lambda_2}(\lambda_1 F_1 - \lambda_2 F_2) + \frac{4}{\lambda_1 - \lambda_3}(\lambda_1 F_1 - \lambda_3 F_3)\right) \\
& + \frac{4\lambda_2\lambda_3}{\lambda_2 + \lambda_3}\left\{-F_1^2 - \frac{2}{\lambda_2 - \lambda_3}\left(F_3 - F_2 + \frac{2}{\lambda_1 - \lambda_2} + \frac{2}{\lambda_2 - \lambda_3} + \frac{2}{\lambda_3 - \lambda_1}\right)\right\}.
\end{aligned} \tag{3.21}$$

From Lemma 3.3 given below, it can be shown that

$$\frac{4}{\lambda_1 - \lambda_2}(\lambda_1 F_1 - \lambda_2 F_2) + \frac{4}{\lambda_1 - \lambda_3}(\lambda_1 F_1 - \lambda_3 F_3) \leq 2\phi_1 + 2\overline{\phi}.$$

It is also observed that

$$\begin{aligned}
& F_3 - F_2 + \frac{2}{\lambda_1 - \lambda_2} + \frac{2}{\lambda_2 - \lambda_3} + \frac{2}{\lambda_3 - \lambda_1} \\
& \quad = \frac{1}{\lambda_3 - \lambda_1} + \frac{1}{\lambda_3 - \lambda_2} - \frac{1}{\lambda_2 - \lambda_1} - \frac{1}{\lambda_2 - \lambda_3} \\
& \qquad + \frac{2}{\lambda_1 - \lambda_2} + \frac{2}{\lambda_2 - \lambda_3} + \frac{2}{\lambda_3 - \lambda_1} \\
& \quad = \frac{3}{\lambda_1 - \lambda_2} - \frac{3}{\lambda_1 - \lambda_3} \\
& \quad = \frac{3(\lambda_2 - \lambda_3)}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)}.
\end{aligned}$$

Hence from (3.21),

$$\begin{aligned}
\widehat{\Delta} \leq & - (c_0 - 2)(\phi_1 + 2\overline{\phi}) - \frac{12c_0}{\lambda_2 + \lambda_3} \\
& + \frac{4\lambda_2\lambda_3}{\lambda_2 + \lambda_3}\left\{-F_1^2 - \frac{6}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)}\right\},
\end{aligned} \tag{3.22}$$

which is not positive for $c_0 \geq 2$. Finally, we shall show the following lemma to complete the proof of Proposition 2.4.

Lemma 3.3 *Assume the case (C2), namely,* $\phi_2 > \phi_3$ *and* $\phi_1 \leq \overline{\phi}$ *for* $\overline{\phi} = \overline{\phi}_{23}$. *Then, the following inequality holds:*

$$\begin{aligned}
&\frac{2}{\lambda_1-\lambda_2}(\lambda_1 F_1 - \lambda_2 F_2) + \frac{2}{\lambda_1-\lambda_3}(\lambda_1 F_1 - \lambda_3 F_3)\\
&\quad = \frac{1}{\lambda_1-\lambda_2}(\lambda_1\phi_1 - \lambda_2\phi_2) + \frac{1}{\lambda_1-\lambda_3}(\lambda_1\phi_1 - \lambda_3\phi_3)\\
&\quad \leq \phi_1 + \overline{\phi}. \qquad (3.23)
\end{aligned}$$

Proof. It is easy to see that the inequality (3.23) is equivalent to

$$(\lambda_1-\lambda_3)\lambda_2\phi_2 + (\lambda_1-\lambda_2)\lambda_3\phi_3 + (\lambda_1-\lambda_2)(\lambda_1-\lambda_3)\overline{\phi} \geq (\lambda_1^2 - \lambda_2\lambda_3)\phi_1.$$

Since $\phi_1 \leq \overline{\phi}$, it is sufficient to show that

$$(\lambda_1-\lambda_3)\lambda_2\phi_2 + (\lambda_1-\lambda_2)\lambda_3\phi_3 + (\lambda_1-\lambda_2)(\lambda_1-\lambda_3)\overline{\phi} \geq (\lambda_1^2 - \lambda_2\lambda_3)\overline{\phi},$$

which is rewritten as

$$(\lambda_1-\lambda_3)\lambda_2\phi_2 + (\lambda_1-\lambda_2)\lambda_3\phi_3 + \{2\lambda_2\lambda_3 - (\lambda_2+\lambda_3)\lambda_1\}\overline{\phi} \geq 0. \qquad (3.24)$$

Substituting $\overline{\phi} = (\lambda_2\phi_2 + \lambda_3\phi_3)/(\lambda_2+\lambda_3)$ into (3.24) and simplifying the expression, we can see that the inequality (3.24) is equivalent to

$$\lambda_2\lambda_3(\lambda_2-\lambda_3)(\phi_2-\phi_3) \geq 0,$$

which is guaranteed by the condition $\phi_2 > \phi_3$. ∎

4 Monte Carlo studies

It is generally surmised that the Stein Risk-Minimization estimator $\boldsymbol{\delta}^{RM*} = \boldsymbol{U}\{\boldsymbol{L} - \boldsymbol{L}\boldsymbol{\Phi}^{RM*}(\boldsymbol{\Lambda})\}\boldsymbol{V}^t$ has nice risk performance notably when each element of the matrix mean $\boldsymbol{\Theta}$ is near zero. There is, however, a problem such that the diagonal elements of $\boldsymbol{L} - \boldsymbol{L}\boldsymbol{\Phi}^{RM*}(\boldsymbol{\Lambda})$ may be negative. Because all singular values l_i's are positive for the singular value decomposition $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{L}\boldsymbol{V}^t$, it would be necessary for the diagonal elements of $\boldsymbol{L} - \boldsymbol{L}\boldsymbol{\Phi}^{RM*}(\boldsymbol{\Lambda})$ not to be negative. This section presents some results of Monte Carlo experiments to compare the risk performances of various estimators and to inspect the so-called positive-part version of $\boldsymbol{\delta}^{RM*}$.

Our Monte Carlo experiments were based on 100,000 independent replications for $(p,m) = (2,5)$, $(2,15)$, $(5,8)$ and $(10,15)$. It is here noted that the minimaxity of the Stein RM estimator $\boldsymbol{\delta}^{RM*}$ is shown for $(p,m) = (2,5)$ and $(2,15)$, but not for $(p,m) = (5,8)$ and $(10,15)$. For true parameter values of the matrix mean $\boldsymbol{\Theta}$ to be estimated, we took typical values for squared singular values of $\boldsymbol{\Theta}$, that is, eigenvalues of $\boldsymbol{\Theta}\boldsymbol{\Theta}^t$, since the risk function of an equivariant estimator (1.3) depends only on singular values of $\boldsymbol{\Theta}$.

Here, for a diagonal matrix $\boldsymbol{\Psi} = \mathrm{diag}\,(\psi_i)$ with the i-th diagonal element ψ_i, $\{\boldsymbol{\Psi}\}_+$ denotes the positive-part of $\boldsymbol{\Psi}$, that is, $\{\boldsymbol{\Psi}\}_+ = \mathrm{diag}\,(\max(0,\psi_i))$, and also $\{\boldsymbol{\Psi}\}_+^O =$

diag $(\max(0, \psi_{(i)}))$ where $\psi_{(i)}$ is the i-th largest value of ψ_i's. The objects of this numerical studies are the following six estimators: $\boldsymbol{\delta}^{ML} = \boldsymbol{X}$, $\boldsymbol{\delta}^{EM}$, $\boldsymbol{\delta}^{ST}$, $\boldsymbol{\delta}_+^{ST*} = \boldsymbol{U}\{\boldsymbol{L}(\boldsymbol{I}_p - \boldsymbol{\Phi}^{ST})\}_+^O \boldsymbol{V}^t$, $\boldsymbol{\delta}^{RM*}$ and $\boldsymbol{\delta}_+^{RM*} = \boldsymbol{U}\{\boldsymbol{L}(\boldsymbol{I}_p - \boldsymbol{\Phi}^{RM*})\}_+ \boldsymbol{V}^t$. Note from Tsukuma (2008) that $\boldsymbol{\delta}_+^{ST*}$ and $\boldsymbol{\delta}_+^{RM*}$ dominate $\boldsymbol{\delta}^{ST}$ and $\boldsymbol{\delta}^{RM*}$, respectively, relative to the loss (1.2).

The estimated risk values by means of the Monte Carlo experiments are given in Tables 1, 2 and 3. The experimental results can now be summarized as follows.

1. As the overall impression of the experimental results, each shrinkage estimator enormously reduces the risk over $\boldsymbol{\delta}^{ML}$ in the cases that all eigenvalues of $\boldsymbol{\Theta\Theta}^t$ are zeros, namely, $\boldsymbol{\Theta}$ is the zero matrix. On the other hand, the larger the eigenvalues of $\boldsymbol{\Theta\Theta}^t$ are, the poorer the risk reduction of shrinkage estimator becomes. Tables 1 and 2 suggest that the improvement in risk increases with m for fixed p.

2. The Stein RM estimator $\boldsymbol{\delta}^{RM*}$ is superior to $\boldsymbol{\delta}^{ML}$ and $\boldsymbol{\delta}^{EM}$. Moreover $\boldsymbol{\delta}^{RM*}$ is better than $\boldsymbol{\delta}^{ST}$ except the cases that all eigenvalues of $\boldsymbol{\Theta\Theta}^t$ are much large or widely scatter. As far as we look through the experimental results for $(p, m) = (5, 8)$ and $(10, 15)$, we are considerably expecting $\boldsymbol{\delta}^{RM*}$ to be minimax for $p \geq 4$.

3. Comparing the estimated risk values of $\boldsymbol{\delta}_+^{RM*}$ and $\boldsymbol{\delta}^{RM*}$, we can see that taking the positive-part of $\boldsymbol{L}(\boldsymbol{I}_p - \boldsymbol{\Phi}^{RM*})$ has a sufficient effect on the risk savings. In the cases that all eigenvalues of $\boldsymbol{\Theta\Theta}^t$ are zeros, the risk reduction by $\boldsymbol{\delta}_+^{RM*}$ over $\boldsymbol{\delta}^{ML}$ is the most substantial in the five shrinkage estimators and their improvements are about 75%, 93%, 87% and 94% for $(p, m) = (2, 5)$, $(2, 15)$, $(5, 8)$ and $(10, 15)$, respectively.

4. The risk performance of $\boldsymbol{\delta}_+^{ST*}$ is not as rich as that of $\boldsymbol{\delta}_+^{RM*}$ in almost all cases. $\boldsymbol{\delta}_+^{ST*}$ is, however, slightly better than $\boldsymbol{\delta}_+^{RM*}$ when all eigenvalues of $\boldsymbol{\Theta\Theta}^t$ are extremely large (e.g., all eigenvalues of $\boldsymbol{\Theta\Theta}^t$ are 10^3 for $p = 10$).

Table 1: Risk values in estimation of the normal mean matrix ($p = 2$ and $m = 5$).

Eigenvalues of $\boldsymbol{\Theta\Theta}^t$	$\boldsymbol{\delta}^{ML}$	$\boldsymbol{\delta}^{EM}$	$\boldsymbol{\delta}^{ST}$	$\boldsymbol{\delta}_+^{ST*}$	$\boldsymbol{\delta}^{RM*}$	$\boldsymbol{\delta}_+^{RM*}$
$(0, 0)$	10.01	6.00	4.00	2.65	3.60	2.50
$(2, 0)$	10.01	6.61	4.99	3.87	4.69	3.78
$(2, 2)$	10.01	7.33	5.87	4.97	5.55	4.82
$(5, 0)$	10.01	7.14	5.97	5.06	5.82	5.06
$(5, 2)$	10.01	7.84	6.66	6.02	6.40	5.88
$(5, 5)$	10.01	8.38	7.28	6.89	6.98	6.68
$(10, 0)$	10.01	7.51	6.83	6.09	6.87	6.22
$(10, 5)$	10.01	8.76	7.90	7.67	7.65	7.47
$(10, 10)$	10.01	9.13	8.29	8.21	8.02	7.96
$(20, 0)$	10.01	7.74	7.45	6.83	7.66	7.09
$(20, 10)$	10.01	9.37	8.78	8.75	8.58	8.55
$(20, 20)$	10.01	9.58	8.94	8.94	8.79	8.79

Table 2: Risk values in estimation of the normal mean matrix ($p = 2$ and $m = 15$).

Eigenvalues of $\mathbf{\Theta\Theta}^t$	$\boldsymbol{\delta}^{ML}$	$\boldsymbol{\delta}^{EM}$	$\boldsymbol{\delta}^{ST}$	$\boldsymbol{\delta}_+^{ST*}$	$\boldsymbol{\delta}^{RM*}$	$\boldsymbol{\delta}_+^{RM*}$
$(0, 0)$	29.98	5.98	4.84	2.63	3.72	2.17
$(2, 0)$	29.98	7.41	6.36	4.32	5.34	3.92
$(2, 2)$	29.98	8.86	7.85	5.99	6.85	5.56
$(5, 0)$	29.98	9.06	8.14	6.40	7.35	6.12
$(5, 2)$	29.98	10.52	9.61	8.04	8.73	7.65
$(5, 5)$	29.98	12.21	11.33	10.05	10.43	9.58
$(10, 0)$	29.98	10.93	10.26	8.88	9.86	8.86
$(10, 5)$	29.98	14.13	13.37	12.47	12.62	12.02
$(10, 10)$	29.98	16.07	15.33	14.78	14.50	14.15
$(20, 0)$	29.98	13.06	12.70	11.65	12.82	11.99
$(20, 10)$	29.98	18.25	17.68	17.40	17.12	16.93
$(20, 20)$	29.98	20.43	19.84	19.78	19.08	19.04

Table 3: Risk values in estimation of the normal mean matrix ($p = 5$ and $m = 8$).

Eigenvalues of $\mathbf{\Theta\Theta}^t$	$\boldsymbol{\delta}^{ML}$	$\boldsymbol{\delta}^{EM}$	$\boldsymbol{\delta}^{ST}$	$\boldsymbol{\delta}_+^{ST*}$	$\boldsymbol{\delta}^{RM*}$	$\boldsymbol{\delta}_+^{RM*}$
$(0, 0, 0, 0, 0)$	40.02	30.01	12.12	6.06	8.28	5.33
$(2, 0, 0, 0, 0)$	40.02	30.40	13.35	7.61	9.74	6.95
$(2, 2, 1, 0, 0)$	40.02	31.04	15.14	9.89	11.75	9.23
$(2, 2, 2, 2, 2)$	40.02	32.29	17.97	13.40	14.76	12.57
$(5, 0, 0, 0, 0)$	40.02	30.81	14.85	9.46	11.67	9.00
$(5, 4, 3, 2, 1)$	40.02	33.07	20.06	16.12	17.11	15.21
$(5, 5, 5, 5, 5)$	40.02	34.81	23.56	20.93	20.72	19.51
$(10, 0, 0, 0, 0)$	40.02	31.18	16.61	11.62	14.11	11.58
$(10, 8, 6, 4, 2)$	40.02	35.01	24.69	22.31	22.13	20.98
$(10, 10, 10, 10, 10)$	40.02	37.20	28.41	27.54	26.15	25.74
$(20, 0, 0, 0, 0)$	40.02	31.49	18.50	13.91	16.91	14.51
$(20, 15, 10, 5, 0)$	40.02	35.54	27.96	26.19	26.22	25.21
$(20, 20, 20, 20, 20)$	40.02	38.78	32.00	31.95	31.23	31.20
$(50, 0, 0, 0, 0)$	40.02	31.71	19.99	15.75	18.52	16.22
$(50, 40, 30, 20, 10)$	40.02	38.94	34.58	34.51	33.53	33.49
$(50, 50, 50, 50, 50)$	40.02	39.58	34.80	34.80	36.11	36.11
$(100, 0, 0, 0, 0)$	40.02	31.82	20.44	16.33	18.59	16.32
$(100, 80, 60, 40, 20)$	40.02	39.52	36.89	36.89	36.45	36.45
$(100, 100, 100, 100, 100)$	40.02	39.81	36.18	36.18	37.95	37.95

Table 4: Risk values in estimation of the normal mean matrix ($p = 10$ and $m = 15$).

Eigenvalues of $\mathbf{\Theta\Theta}^t$	$\boldsymbol{\delta}^{ML}$	$\boldsymbol{\delta}^{EM}$	$\boldsymbol{\delta}^{ST}$	$\boldsymbol{\delta}_+^{ST*}$	$\boldsymbol{\delta}^{RM*}$	$\boldsymbol{\delta}_+^{RM*}$
$(0,0,0,0,0,0,0,0,0,0)$	150.02	110.02	44.19	15.08	16.15	9.37
$(2,0,0,0,0,0,0,0,0,0)$	150.02	110.55	45.57	16.81	17.95	11.21
$(2,2,2,1,1,1,1,0,0,0)$	150.02	112.59	50.86	23.67	24.54	18.21
$(2,2,2,2,2,2,2,2,2,2)$	150.02	115.20	57.11	31.87	32.07	26.24
$(5,0,0,0,0,0,0,0,0,0)$	150.02	111.10	47.34	19.13	20.44	13.79
$(5,5,4,4,3,3,2,2,1,1)$	150.02	117.32	62.46	39.10	38.73	33.33
$(5,5,5,5,5,5,5,5,5,5)$	150.02	121.90	72.32	52.92	50.08	45.79
$(10,0,0,0,0,0,0,0,0,0)$	150.02	111.73	49.77	22.34	24.15	17.62
$(10,10,8,8,6,6,4,4,2,2)$	150.02	122.97	75.83	57.95	54.81	50.80
$(10,10,10,10,10,10,10,10,10,10)$	150.02	130.45	89.72	78.62	70.52	68.20
$(20,0,0,0,0,0,0,0,0,0)$	150.02	112.46	53.23	26.86	30.04	23.63
$(20,20,15,15,10,10,5,5,0,0)$	150.02	126.42	87.25	73.99	70.06	66.75
$(20,20,20,20,20,20,20,20,20,20)$	150.02	139.63	107.44	105.08	94.82	94.36
$(50,0,0,0,0,0,0,0,0,0)$	150.02	113.27	57.60	32.52	36.88	30.60
$(50,50,40,40,30,30,20,20,10,10)$	150.02	141.40	117.15	115.71	106.76	106.38
$(50,50,50,50,50,50,50,50,50,50)$	150.02	146.36	122.84	122.83	126.07	126.07
$(100,0,0,0,0,0,0,0,0,0)$	150.02	113.65	59.32	34.86	37.92	31.69
$(100,100,80,80,60,60,40,40,20,20)$	150.02	145.92	129.73	129.67	125.60	125.58
$(100,100,100,100,100,$ $100,100,100,100,100)$	150.02	148.30	129.86	129.86	138.42	138.42
$(10^3,0,0,0,0,0,0,0,0,0)$	150.02	114.06	60.72	36.93	38.61	32.44
$(10^3,10^3,800,800,600,$ $600,400,400,200,200)$	150.02	149.65	147.04	147.04	147.74	147.74
$(10^3,10^3,10^3,10^3,10^3,$ $10^3,10^3,10^3,10^3,10^3)$	150.02	149.86	142.55	142.55	148.53	148.53

References

[1] Efron, B. and Morris, C. (1972). Empirical Bayes on vector observations: An extension of Stein's method, *Biometrika*, **59**, 335-347.

[2] Lin, S.P. and Perlman, M.D. (1985). A Monte Carlo comparison of four estimators for a covariance matrix, in *Multivariate Analysis VI*, 411-429, North-Holland, Amsterdam.

[3] Robertson, T., Wright, F.T. and Dykstra, R. (1988). *Order Restricted Statistical Inference*, Wiley, Chichester.

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

[5] Stein, C. (1977). Lectures on the theory of estimation of many parameters, In *Studies in the Statistical Theory of Estimation I*, I.A. Ibragimov and M.S. Nikulin (eds.), **74**, 4–65, Nauka, Leningrad.

[6] Tsukuma, H. (2008). Admissibility and minimaxity of Bayes estimators for a normal mean matrix. *J. Multivariate Analysis*, to appear.