Alternative estimators of the common regression matrix in two GMANOVA models under weighted quadratic losses

Hisayuki Tsukuma [a] Yoshihiko Konno [b]

[a] *The Institute of Statistical Mathematics, 4-6-7, Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan*

[b] *Faculty of Science, Japan Women's University, 2-8-1 Mejirodai, Bunkyo-ku, Tokyo, 112-869 Japan*

Abstract

We consider the problem of estimating the common regression matrix of two GMANOVA models with different unknown covariance matrices under a certain type of loss functions which include a weighted quadratic loss function as a special case. We consider a class of estimators, which contains the Graybill-Deal type estimator proposed by Sugiura and Kubokawa (1988), and we give its risk representation via Kubokawa and Srivastava's (1999, 2001) identities when the error matrices follow the elliptically contoured distributions. Using the method similar to an approximate minimization of the unbiased risk estimate due to Stein (1977), we obtain an alternative estimator to the Graybill-Deal type estimator which was given under the normality assumption. However, it seems difficult to evaluate the risk of our proposed estimator analytically because of complex nature of its risk function. Instead, we conduct a Monte-Carlo simulation to evaluate the performance of our proposed estimator. The results indicate that our proposed estimator compares favorably with the Graybill-Deal type estimator.

Key words: common mean, Stein's loss, Stein-Haff identity, two-sample problem, elliptically contoured distribution
MSC: primary: 62H12 ; secondary: 62F10, 62J07

Email addresses: tsukuma@ism.ac.jp (Hisayuki Tsukuma), konno@fc.jwu.ac.jp (Yoshihiko Konno).
URL: http://mp-w3math.jwu.ac.jp/~konno/index.html (Yoshihiko Konno).

1 Introduction

This paper is mainly concerned with estimating the common regression matrix of two GMANOVA models with different covariance matrices n a decision-theoretic point of view. Sugiura and Kubokawa (1988) first considered this problem and proposed the Graybill-Deal type estimator of the common regression matrix of two GMANOVA models. In this paper, using the techniques of Stein (1977), Loh (1991), and Kubokawa and Srivastava (1999, 2001), we shall derive an alternative estimator for the common regression matrix and show, using Monte Carlo simulation, that our proposed estimator compares favorably with that of Sugiura and Kubokawa (1988). The precise formulation of this problem is as follows.

Let $\boldsymbol{Y}_i,\ i = 1,\ 2,$ be $N_i \times p_i$ matrices of response variables and consider the two GMANOVA models

$$\boldsymbol{Y}_1 = \boldsymbol{A}_{11}\Xi\boldsymbol{A}_{12} + \boldsymbol{\epsilon}_1 \qquad \text{and} \qquad \boldsymbol{Y}_2 = \boldsymbol{A}_{21}\Xi\boldsymbol{A}_{22} + \boldsymbol{\epsilon}_2, \tag{1}$$

where $\boldsymbol{A}_{i1}$ and $\boldsymbol{A}_{i2}$ are, respectively, $N_i \times m$ and $q \times p_i$ known full-rank matrices with $N_i > m$ and $p_i \geq q$, Ξ is an $m \times q$ matrix of unknown parameters, and $\boldsymbol{\epsilon}_i$ are $N_i \times p_i$ error matrices with mean zero matrices. We assume that $\boldsymbol{A}_{i1}'\boldsymbol{A}_{i1},\ i = 1, 2,$ are $m \times m$ diagonal matrices and that the error matrices $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$ are jointly distributed as the elliptically contoured distribution with the density function

$$|\boldsymbol{\Omega}_1|^{-N_1/2}|\boldsymbol{\Omega}_2|^{-N_2/2}g(\operatorname{tr}(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\epsilon}_1'\boldsymbol{\epsilon}_1) + \operatorname{tr}(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\epsilon}_2'\boldsymbol{\epsilon}_2)), \tag{2}$$

where g is a nonnegative unknown function and $\boldsymbol{\Omega}_i,\ i = 1,\ 2,$ are $p_i \times p_i$ scale matrices. We also assume that $\boldsymbol{\Omega}_i$ are unknown positive definite $p_i \times p_i$ matrices. Here we denote by $\boldsymbol{B}'$, $|\boldsymbol{B}|$, and $\operatorname{tr}(\boldsymbol{B})$ the transpose, determinant, and trace of a squared matrix $\boldsymbol{B}$, respectively. Here we note that the model (1) occurs in missing data model of one-sample growth curve model.

We consider the problem of estimating Ξ under the loss function

$$\begin{aligned}\widetilde{\boldsymbol{L}}((\Xi,\ \boldsymbol{\Omega}_1,\ \boldsymbol{\Omega}_2),\ \widehat{\Xi}) = &\operatorname{tr}\{\boldsymbol{A}_{11}(\widehat{\Xi} - \Xi)\boldsymbol{A}_{12}\boldsymbol{\Omega}_1^{-1}\boldsymbol{A}_{12}'(\widehat{\Xi} - \Xi)'\boldsymbol{A}_{11}'\} \\ &+ \operatorname{tr}\{\widetilde{\boldsymbol{C}}(\widehat{\Xi} - \Xi)\boldsymbol{A}_{22}\boldsymbol{\Omega}_2^{-1}\boldsymbol{A}_{22}'(\widehat{\Xi} - \Xi)'\widetilde{\boldsymbol{C}}'\},\end{aligned} \tag{3}$$

where $\widehat{\Xi}$ is an estimator of Ξ and $\widetilde{\boldsymbol{C}}$ is an $N_2 \times m$ known matrix of full rank. When $\widetilde{\boldsymbol{C}} = \boldsymbol{A}_{21}$, the above loss function is a natural extension of an invariant loss function of the regression matrix of the GMANOVA model, which was used by Kariya, et al. (1996, 1999). This loss function includes a quadratic loss which was used by Loh (1991) in estimating the common mean of the multivariate normal distributions. Then the inaccuracy of an estimator

$\widehat{\Xi}$ is measured by the risk function $\mathbb{E}[\widetilde{\boldsymbol{L}}((\Xi, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2), \widehat{\Xi})]$. On the other hand, Kubokawa (1989) considered the problem of estimating the common regression matrix of several GMANOVA models and employed the quadratic loss function $\mathrm{tr}\,\{(\widehat{\Xi}-\Xi)\boldsymbol{Q}(\widehat{\Xi}-\Xi)'\}$ for a $q \times q$ known positive definite matrix $\boldsymbol{Q}$.

In Section 2, we first give a canonical form of two sample problem of estimating the common regression matrix of the model (1). Next we derive a class of a fully equivariant estimators under the group of affine transformations. Then we modify this class to obtain our class of estimators, which contains the Graybill-Deal type estimator by Sugiura and Kubokawa (1988). Using the method of Kubokawa and Srivastava (1999, 2001), we obtain a risk representation for the class of estimators. In Section 3, using an approximate minimization of the risk representation, we obtain an alternative estimator to the Graybill-Deal type estimator. However our proposed estimator is difficult to treat analytically. Instead, we carry out Monte-Carlo simulation to compare the risks of these estimators in Section 4. We first give a numerical result that our estimators proposed in Section 2 reduce the risk substantially over the Graybill-Deal type estimator under normality assumption. Next we conduct a simulation for a special case of two GMANOVA models with elliptical error matrices. Since the model (2) is not i.i.d. sampling set-up of two sample problems, we carry out Monte-Carlo simulation to show that our proposed estimators reduce the risk under the i.i.d. sampling from two independent multivariate elliptically contoured distributions instead of sampling from the model (2) in order to justify our derivation of alternative estimators under the model (2). In Section 5, we give technical lemmas and the proofs of the main results.

2 A class of estimators and its risk representation

2.1 A canonical form

Kubokawa and Srivastava (2001, page 149) give a canonical form of one-sample GMANOVA model with an elliptical error matrix. In the same way as the canonical reduction of Kubokawa and Srivastava (2001), we can rewrite the elliptical density (2) as the following lemma. For a detail proof, see Tsukuma and Konno (2002).

Lemma 1. *The density function of the model* (1) *with* (2) *is written as*

$$
\begin{aligned}
&(|\boldsymbol{\Sigma}_1| \times |\boldsymbol{\Lambda}_1|)^{-N_1/2}(|\boldsymbol{\Sigma}_2| \times |\boldsymbol{\Lambda}_2|)^{-N_2/2} \\
&\times g\Big\{ \operatorname{tr}\Big[\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{X}_1 - \boldsymbol{Z}_1\boldsymbol{\gamma}_1 - \boldsymbol{\Theta})'(\boldsymbol{X}_1 - \boldsymbol{Z}_1\boldsymbol{\gamma}_1 - \boldsymbol{\Theta}) \\
&\qquad +\boldsymbol{\Sigma}_1^{-1}\boldsymbol{s}_1'\boldsymbol{s}_1 + \boldsymbol{\Sigma}_1^{-1}(\boldsymbol{u}_1 - (\boldsymbol{w}_1'\boldsymbol{w}_1)^{1/2}\boldsymbol{\gamma}_1)'(\boldsymbol{u}_1 - (\boldsymbol{w}_1'\boldsymbol{w}_1)^{1/2}\boldsymbol{\gamma}_1)\Big] \\
&\qquad + \operatorname{tr}\Big[\boldsymbol{\Lambda}_1^{-1}\{\boldsymbol{Z}_1'\boldsymbol{Z}_1 + \boldsymbol{w}_1'\boldsymbol{w}_1\}\Big] + \operatorname{tr}\Big[\boldsymbol{\Sigma}_2^{-1}(\boldsymbol{X}_2 - \boldsymbol{Z}_2\boldsymbol{\gamma}_2 - \boldsymbol{A}\boldsymbol{\Theta})' \\
&\qquad \times(\boldsymbol{X}_2 - \boldsymbol{Z}_2\boldsymbol{\gamma}_2 - \boldsymbol{A}\boldsymbol{\Theta}) + \boldsymbol{\Sigma}_2^{-1}\boldsymbol{s}_2'\boldsymbol{s}_2 + \boldsymbol{\Sigma}_2^{-1}(\boldsymbol{u}_2 - (\boldsymbol{w}_2'\boldsymbol{w}_2)^{1/2}\boldsymbol{\gamma}_2)' \\
&\qquad \times(\boldsymbol{u}_2 - (\boldsymbol{w}_2'\boldsymbol{w}_2)^{1/2}\boldsymbol{\gamma}_2)\Big] + \operatorname{tr}\Big[\boldsymbol{\Lambda}_2^{-1}\{\boldsymbol{Z}_2'\boldsymbol{Z}_2 + \boldsymbol{w}_2'\boldsymbol{w}_2\}\Big]\Big\}. \qquad (4)
\end{aligned}
$$

Here $\boldsymbol{X}_i$, $\boldsymbol{Z}_i$, $\boldsymbol{s}_i$, $\boldsymbol{u}_i$ and $\boldsymbol{w}_i\,(i = 1, 2)$ are, respectively, $m \times q$, $m \times (p_i - q)$, $(N_i - m - p_i + q) \times q$, $(p_i - q) \times q$ and $(N_i - m) \times (p_i - q)$ random matrices and $\boldsymbol{\Theta}$, $\boldsymbol{\Sigma}_i$, $\boldsymbol{\Lambda}_i$ and $\boldsymbol{\gamma}_i$ are, respectively, $m \times q$, $q \times q$, $(p_i - q) \times (p_i - q)$ and $(p_i - q) \times q$ matrices of unknown parameters and $\boldsymbol{A}$ is an $m \times m$ diagonal matrix with positive diagonal elements.

For $i = 1, 2$, put

$$
\boldsymbol{S}_i = \boldsymbol{s}_i'\boldsymbol{s}_i, \quad \boldsymbol{W}_i = \boldsymbol{w}_i'\boldsymbol{w}_i, \quad \widehat{\gamma}_i = \boldsymbol{W}_i^{-1/2}\boldsymbol{u}_i, \quad \widehat{\boldsymbol{\Theta}}_i = \boldsymbol{X}_i - \boldsymbol{Z}_i\widehat{\gamma}_i \qquad (5)
$$

and $n_i = N_i - m - p_i + q$. Furthermore, the loss function (3) turns into

$$
\begin{aligned}
\boldsymbol{L}((\boldsymbol{\Theta}, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2), \widehat{\boldsymbol{\Theta}}) = &\operatorname{tr}[(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})\boldsymbol{\Sigma}_1^{-1}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})'] \\
&+ \operatorname{tr}[\boldsymbol{C}'\boldsymbol{C}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})\boldsymbol{\Sigma}_2^{-1}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})'], \qquad (6)
\end{aligned}
$$

where $\widehat{\boldsymbol{\Theta}}$ is an estimator of $\boldsymbol{\Theta}$ and $\boldsymbol{C}$ is an $N_2 \times m$ known matrix of full-rank. Under this canonical form, the problem of estimating $\boldsymbol{\Xi}$ in (1) changes into that of estimating $\boldsymbol{\Theta}$ based on $(\boldsymbol{X}_i, \boldsymbol{Z}_i, \boldsymbol{S}_i, \widehat{\gamma}_i, \boldsymbol{W}_i |\, i = 1, 2)$ under the loss function (6). Then the risk function is defined by

$$
\boldsymbol{R}((\boldsymbol{\Theta}, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2), \widehat{\boldsymbol{\Theta}}) = \mathbb{E}[\boldsymbol{L}((\boldsymbol{\Theta}, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2), \widehat{\boldsymbol{\Theta}})],
$$

where the expectation is taken with respect to the density function given by (4).

Remark 1. If the errors have normal distributions, then a set of random matrices $(\boldsymbol{X}_i, \boldsymbol{Z}_i, \boldsymbol{S}_i, \widehat{\gamma}_i, \boldsymbol{W}_i)$ is distributed as

$$
\boldsymbol{X}_1 |\, \boldsymbol{Z}_1 \sim N_{m\times q}(\boldsymbol{\Theta} + \boldsymbol{Z}_1\boldsymbol{\gamma}_1, \boldsymbol{I}_m \otimes \boldsymbol{\Sigma}_1), \qquad (7a)
$$

$$
\boldsymbol{X}_2 |\, \boldsymbol{Z}_2 \sim N_{m\times q}(\boldsymbol{A}\boldsymbol{\Theta} + \boldsymbol{Z}_2\boldsymbol{\gamma}_2, \boldsymbol{I}_m \otimes \boldsymbol{\Sigma}_2) \qquad (7b)
$$

and, for $i = 1, 2$,

$$\boldsymbol{Z}_i \sim N_{m\times(p_i-q)}(\boldsymbol{0}, \boldsymbol{I}_m \otimes \boldsymbol{\Lambda}_i), \tag{8a}$$
$$\boldsymbol{S}_i \sim W_q(\boldsymbol{\Sigma}_i,\, n_i),\;\; n_i = N_i - m - p_i + q, \tag{8b}$$
$$\widehat{\boldsymbol{\gamma}}_i|\, \boldsymbol{W}_i \sim N_{(p_i-q)\times q}(\boldsymbol{\gamma}_i,\, \boldsymbol{W}_i^{-1} \otimes \boldsymbol{\Sigma}_i), \tag{8c}$$
$$\boldsymbol{W}_i \sim W_{p_i-q}(\boldsymbol{\Lambda}_i, n_i + p_i - q). \tag{8d}$$

where and $\boldsymbol{G} \otimes \boldsymbol{H}$ stands for the Kronecker product of matrices $\boldsymbol{G}$ and $\boldsymbol{H}$ defined by $(g_{ij}\boldsymbol{H})$ for $\boldsymbol{G} = (g_{ij})$. Here, note that $(\boldsymbol{X}_i,\, \boldsymbol{Z}_i)$, $(\boldsymbol{W}_i,\, \widehat{\boldsymbol{\gamma}}_i)$ and $\boldsymbol{S}_i$ are independent and that two statistics $\widehat{\boldsymbol{\Theta}}_1$ and $\widehat{\boldsymbol{\Theta}}_2$ are the maximum likelihood estimators of $\boldsymbol{\Theta}$ and $\boldsymbol{A\Theta}$ for one-sample problem, respectively.

2.2 A class of estimators and Graybill-Deal type estimator

Under normality assumption, Sugiura and Kubokawa (1988) proposed the Graybill-Deal type estimator of the form

$$\begin{aligned}\mathrm{vec}(\widehat{\boldsymbol{\Theta}}^{SK}) = &\{\boldsymbol{I}_m \otimes (\boldsymbol{S}_1/n_1)^{-1} + \boldsymbol{A}^2 \otimes (\boldsymbol{S}_2/n_2)^{-1}\}^{-1} \\ &\times\{\boldsymbol{I}_m \otimes (\boldsymbol{S}_1/n_1)^{-1}\mathrm{vec}(\widehat{\boldsymbol{\Theta}}_1) + \boldsymbol{A}^2 \otimes (\boldsymbol{S}_2/n_2)^{-1}\mathrm{vec}(\boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2)\},\end{aligned} \tag{9}$$

where we denote by $\mathrm{vec}(\boldsymbol{U})$ an $mq \times 1$ vector consisting of $(u_1,\, u_2,\, \ldots\, ,\, u_m)'$ for $\boldsymbol{U} = (u_1',\, u_2',\, \ldots\, ,\, u_m')'$. Note that the above estimator (9) is unbiased estimator of $\boldsymbol{\Theta}$. In this subsection, we obtain a class of estimators which includes the Graybill-Deal type estimator.

To clarify our class of estimators of $\boldsymbol{\Theta}$ considered in the sequel of this paper, we first derive a class of equivariant estimators. To this end, let G be a group of transformations on the sample space. Each element of G consists of triples $(\boldsymbol{D},\, \boldsymbol{P}_1,\, \boldsymbol{P}_2)$, where $\boldsymbol{D}$ is an $m \times q$ matrix and

$$\boldsymbol{P}_i = \begin{pmatrix} \boldsymbol{P}_{11} & \boldsymbol{P}_{i\cdot 12} \\ \boldsymbol{0}_{(p_i-q)\times q} & \boldsymbol{P}_{i\cdot 22} \end{pmatrix}, \qquad i = 1,\, 2.$$

Here $\boldsymbol{P}_{11}$ and $\boldsymbol{P}_{i\cdot 22}$ are $q \times q$ and $(p_i - q) \times (p_i - q)$ nonsingular matrices, respectively, and $\boldsymbol{P}_{i\cdot 12}$ are $q \times (p_i - q)$ matrices. Here note that the left-upper blocks of $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are identical so as to capture the structure of estimating the common regression matrix in two GMANOVA models. The group composition is given by $(\boldsymbol{D},\, \boldsymbol{P}_1,\, \boldsymbol{P}_2)(\tilde{\boldsymbol{D}},\, \tilde{\boldsymbol{P}}_1,\, \tilde{\boldsymbol{P}}_2) = (\boldsymbol{D} + \tilde{\boldsymbol{D}},\, \boldsymbol{P}_1\tilde{\boldsymbol{P}}_2,\, \boldsymbol{P}_2\tilde{\boldsymbol{P}}_2)$ where $(\boldsymbol{D},\, \boldsymbol{P}_1,\, \boldsymbol{P}_2)$ and $(\tilde{\boldsymbol{D}},\, \tilde{\boldsymbol{P}}_1,\, \tilde{\boldsymbol{P}}_2)$ are elements of G. Then the actions of $(\boldsymbol{D},\, \boldsymbol{P}_1,\, \boldsymbol{P}_2)$ on $(\widehat{\boldsymbol{\Theta}}_i,\, \boldsymbol{Z}_i,\, \boldsymbol{S}_i,\, \widehat{\boldsymbol{\gamma}}_i,\, \boldsymbol{W}_i|\, i = 1,\, 2)$ and $(\boldsymbol{\Theta},\, \boldsymbol{\Sigma}_i,\, \boldsymbol{\gamma}_i,\, \boldsymbol{\Lambda}_i|\, i = 1,\, 2)$ are defined as

$\boldsymbol{\Theta} \to \boldsymbol{\Theta}\boldsymbol{P}_{11}' + \boldsymbol{D}$
$(\boldsymbol{\Sigma}_i,\, \boldsymbol{\Lambda}_i,\, \boldsymbol{\gamma}_i) \to (\boldsymbol{P}_{11}\boldsymbol{\Sigma}_i\boldsymbol{P}_{11}',\, \boldsymbol{P}_{i\cdot 22}\boldsymbol{\Lambda}_i\boldsymbol{P}_{i\cdot 22}',\, (\boldsymbol{P}_{i\cdot 22}')^{-1}\boldsymbol{\gamma}_i\boldsymbol{P}_{11}' + (\boldsymbol{P}_{i\cdot 22}')^{-1}\boldsymbol{P}_{i\cdot 12}'),$
$(\widehat{\boldsymbol{\Theta}}_1,\, \boldsymbol{Z}_1,\, \widehat{\boldsymbol{\Theta}}_2,\, \boldsymbol{Z}_2) \to (\widehat{\boldsymbol{\Theta}}_1\boldsymbol{P}_{11}' + \boldsymbol{D},\, \boldsymbol{Z}_1\boldsymbol{P}_{1\cdot 22}',\, \widehat{\boldsymbol{\Theta}}_2\boldsymbol{P}_{11}' + \boldsymbol{A}\boldsymbol{D},\, \boldsymbol{Z}_2\boldsymbol{P}_{2\cdot 22}'),$
$(\boldsymbol{S}_i,\, \boldsymbol{W}_i,\, \widehat{\gamma}_i) \to (\boldsymbol{P}_{11}\boldsymbol{S}_i\boldsymbol{P}_{11}',\, \boldsymbol{P}_{i\cdot 22}\boldsymbol{W}_i\boldsymbol{P}_{i\cdot 22}',\, (\boldsymbol{P}_{i\cdot 22}')^{-1}\widehat{\gamma}_i\boldsymbol{P}_{11}' + (\boldsymbol{P}_{i\cdot 22}')^{-1}\boldsymbol{P}_{i\cdot 12}')$

for $i = 1, 2$. Denote by $g \circ (\widehat{\boldsymbol{\Theta}}_i,\, \boldsymbol{Z}_i,\, \boldsymbol{S}_i,\, \widehat{\gamma}_i,\, \boldsymbol{W}_i |\, i = 1, 2)$ the action of g on this sample where g is an element of G, i.e., $g = (\boldsymbol{D},\, \boldsymbol{P}_1,\, \boldsymbol{P}_2)$. Then the model is easily shown to be invariant under the group of transformations. It is reasonable to require that an equivariant estimator $\widehat{\boldsymbol{\Theta}}^{EQI}$ should satisfy

$$\widehat{\boldsymbol{\Theta}}^{EQI}(g \circ (\widehat{\boldsymbol{\Theta}}_i,\, \boldsymbol{Z}_i,\, \boldsymbol{S}_i,\, \widehat{\gamma}_i,\, \boldsymbol{W}_i)) = \widehat{\boldsymbol{\Theta}}^{EQI}(\widehat{\boldsymbol{\Theta}}_i,\, \boldsymbol{Z}_i,\, \boldsymbol{S}_i,\, \boldsymbol{W}_i,\, \widehat{\gamma}_i)\boldsymbol{P}_{11}' + \boldsymbol{D},$$

so that $\widehat{\boldsymbol{\Theta}}^{EQI}(g \circ (\widehat{\boldsymbol{\Theta}}_i,\, \boldsymbol{Z}_i,\, \boldsymbol{S}_i,\, \widehat{\gamma}_i))$ estimates the parameter $\boldsymbol{\Theta}\boldsymbol{P}_{11}' + \boldsymbol{D}$ as does $\widehat{\boldsymbol{\Theta}}^{EQI}(\widehat{\boldsymbol{\Theta}}_i,\, \boldsymbol{Z}_i,\, \boldsymbol{S}_i,\, \boldsymbol{W}_i,\, \widehat{\gamma}_i)\boldsymbol{P}_{11}' + \boldsymbol{D}$. The next theorem characterizes the form of equivariant estimators.

Theorem 1. *Let $\boldsymbol{B}$ be a $q \times q$ nonsingular matrix such that $\boldsymbol{B}(\boldsymbol{S}_1 + \boldsymbol{S}_2)\boldsymbol{B}' = \boldsymbol{I}_q$, and let $\boldsymbol{F} = \mathbf{diag}(f_1, \ldots, f_q)$ be a $q \times q$ diagonal matrix such that $\boldsymbol{B}\boldsymbol{S}_2\boldsymbol{B}' = \boldsymbol{F}$ and $f_1 \geq \cdots \geq f_q > 0$. Then under the group of transformations, an equivariant estimator of $\boldsymbol{\Theta}$ is given by*

$$\widehat{\boldsymbol{\Theta}}^{EQI} = \widehat{\boldsymbol{\Theta}}_1\boldsymbol{B}'\tilde{\boldsymbol{\Phi}}(\boldsymbol{B}')^{-1} + \boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2\boldsymbol{B}'(\boldsymbol{I}_q - \tilde{\boldsymbol{\Phi}})(\boldsymbol{B}')^{-1}, \tag{10}$$

where $\tilde{\boldsymbol{\Phi}} \equiv \tilde{\boldsymbol{\Phi}}((\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2)\boldsymbol{B}',\, \boldsymbol{F},\, \boldsymbol{Z}_1\boldsymbol{W}_1^{-1/2},\, \boldsymbol{Z}_2\boldsymbol{W}_2^{-1/2})$ is a $q \times q$ matrix.

Proof. The proof of the theorem can be obtained in much similar way as that given in Loh (1988) and is omitted. □

Unfortunately, the Graybill-Deal type estimator (9) does not belong to the class given by (10). In addition, it is too large to obtain a risk representation of estimate for this class. In consideration of aspects of the class of estimators, we restrict ourselves to following class of unbiased estimators

$$\begin{aligned}\mathrm{vec}(\widehat{\boldsymbol{\Theta}}^{RE}) = &\{\boldsymbol{I}_m \otimes (\boldsymbol{B}'\,\mathbf{diag}\,(\beta_j)\boldsymbol{B}) + \boldsymbol{T} \otimes (\boldsymbol{B}'\,\mathbf{diag}\,(\alpha_j)\boldsymbol{B})\}^{-1}\\ &\times\{\boldsymbol{I}_m \otimes (\boldsymbol{B}'\,\mathbf{diag}\,(\beta_j)\boldsymbol{B})\mathrm{vec}(\widehat{\boldsymbol{\Theta}}_1)\\ &\quad + \boldsymbol{T} \otimes (\boldsymbol{B}'\,\mathbf{diag}\,(\alpha_j)\boldsymbol{B})\mathrm{vec}(\boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2)\},\end{aligned} \tag{11}$$

where $\widehat{\boldsymbol{\Theta}}_i$, $i = 1, 2$, is given by (5), $\alpha_j \equiv \alpha_j(\boldsymbol{F})$ and $\beta_j \equiv \beta_j(\boldsymbol{F})$ depend only on $\boldsymbol{F} = \mathbf{diag}\,(f_1,\, f_2,\, \ldots,\, f_q)$ with $f_1 \geq f_2 \geq \cdots \geq f_q$, the eigenvalues of $\boldsymbol{S}_2(\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1}$, and $\boldsymbol{T}$ is an $m \times m$ diagonal matrix with positive diagonal elements such that $\boldsymbol{T} = \mathbf{diag}\,(t_1, \ldots, t_m)$ with $t_l > 0$, $l = 1, 2, \ldots, m$. Here we denote by $\mathbf{diag}(\tilde{\beta}_j)$ a $q \times q$ diagonal matrix whose j-th diagonal elements are given by $\tilde{\beta}_j$. Note that the estimator (11) reduces to the estimator (9) if we put $\alpha_j(\boldsymbol{F}) = n_2/f_j$, $\beta_j(\boldsymbol{F}) = n_1/(1-f_j)$, and $\boldsymbol{T} = \boldsymbol{A}^2$ and that the estimator (11) belongs to the class of estimators given by (10) if $\boldsymbol{T}$ is a constant multiple of identity matrix. Furthermore, it is well known that the eigenvalues of $\boldsymbol{S}_2(\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1}$, i.e., $\boldsymbol{F}$, are more spread than the eigenvalues of the expected value of $\boldsymbol{S}_2(\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1}$. Hence we look for alternative estimators for $\boldsymbol{\Theta}$ by correcting the eigenvalues of $\boldsymbol{S}_2(\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1}$, i.e., looking for a promising estimator with the form (9) having natural ordering properties

$$0 < \alpha_1(\boldsymbol{F}) \le \cdots \le \alpha_q(\boldsymbol{F}), \qquad \beta_1(\boldsymbol{F}) \ge \cdots \ge \beta_q(\boldsymbol{F}) > 0. \tag{12}$$

2.3 Risk representation

We next give a risk representation of the estimators (11) when the matrix $\boldsymbol{C}'\boldsymbol{C}$ in the loss function (6) is a diagonal matrix such that $\boldsymbol{C}'\boldsymbol{C} = \mathbf{diag}\,(c_1^2, c_2^2, \ldots, c_m^2)$. To evaluate the risk of the estimators (11), we need the following notation which is used for the extended Wishart identity for the elliptically contoured distribution due to Kubokawa and Srivastava (1999). Let U be an integrable function of $(\boldsymbol{X}_i, \boldsymbol{Z}_i, \boldsymbol{s}_i, \boldsymbol{u}_i, \boldsymbol{w}_i |\, i = 1, 2)$ and define

$$\mathbb{E}_G[U] = \int U \times |\boldsymbol{\Sigma}_1|^{-N_1/2} |\boldsymbol{\Sigma}_2|^{-N_2/2} G(d) \prod_{i=1}^{2} d\boldsymbol{X}_i d\boldsymbol{Z}_i d\boldsymbol{s}_i d\boldsymbol{u}_i d\boldsymbol{w}_i,$$

where $G(x) = \frac{1}{2}\int_x^{+\infty} g(t)dt$ and d is given by the terms inside large curly bracket of (4).

Theorem 2. *Let $\widehat{\boldsymbol{\Theta}}_i = (\hat{\boldsymbol{\theta}}_{i1}, \ldots, \hat{\boldsymbol{\theta}}_{im})'$ where $\hat{\boldsymbol{\theta}}_{il}$ are $q \times 1$ vectors and $\boldsymbol{Z}_i = (\boldsymbol{z}_{i1}, \ldots, \boldsymbol{z}_{im})'$ where $\boldsymbol{z}_{il}$ are $(p_i - q) \times 1$ vectors for $i = 1, 2$ and $l = 1, 2, \ldots, m$. Put $\boldsymbol{A} = \mathbf{diag}\,(a_1, \ldots, a_m)$. Then the risk of the estimator (11) is written as*

$$\begin{aligned}
&\boldsymbol{R}((\boldsymbol{\Theta}, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2), \widehat{\boldsymbol{\Theta}}^{RE}) \\
&= \sum_{l=1}^{m} \mathbb{E}_G \Big[-q\hat{r}_{1l} + q\hat{r}_{2l}\frac{c_l^2}{a_l^2} + \sum_{j=1}^{q} \Big\{ 2\Big(\hat{r}_{1l} - \hat{r}_{2l}\frac{c_l^2}{a_l^2}\Big)\phi_{lj} \\
&\qquad + h_{lj}^2 \Big\{ (n_1 - q - 1)\frac{(1-\phi_{lj})^2}{1-f_j} + 4(1-\phi_{lj}) f_j \frac{\partial \phi_{lj}}{\partial f_j} \\
&\qquad + 2\sum_{k \ne j} (1-\phi_{lj})(\phi_{lj} - \phi_{lk}) \frac{f_k}{f_j - f_k} + (n_2 - q - 1) c_l^2 \frac{\phi_{lj}^2}{f_j} \\
&\qquad + 4c_l^2 \phi_{lj}(1-f_j)\frac{\partial \phi_{lj}}{\partial f_j} + 2c_l^2 \sum_{k \ne j} \phi_{lj}(\phi_{lj} - \phi_{lk}) \frac{1-f_k}{f_j - f_k} \Big\}\Big\}\Big],
\end{aligned} \tag{13}$$

where h_{lj} is the j-th element of $\boldsymbol{B}(\hat{\boldsymbol{\theta}}_{1l} - a_l^{-1}\hat{\boldsymbol{\theta}}_{2l})$, $\phi_{lj} = \beta_j/(\beta_j + t_l\alpha_j)$, and

$\hat{r}_{il} = 1 + \boldsymbol{z}'_{il}\boldsymbol{W}_i^{-1}\boldsymbol{z}_{il}$ *for* $l = 1, 2, \ldots, m$ *and* $j = 1, 2, \ldots, q$.

Proof. The proof of the theorem is put into Section 5. □

3 Alternative estimator

Although we obtain the risk representation (13) for the class of estimators (11), it is still difficult to deal with it to derive an alternative estimator. We adapt the argument given by Loh (1991) for obtaining more feasible risk representation from (13).

Denote by $\widehat{\boldsymbol{R}}_l$ the terms inside large bracket in the right-hand side of (13), i.e.,

$$\begin{aligned}\widehat{\boldsymbol{R}}_l = -q\hat{r}_{1l} + q\hat{r}_{2l}\frac{c_l^2}{a_l^2} + \sum_{j=1}^{q}\Big\{ & 2\Big(\hat{r}_{1l} - \hat{r}_{2l}\frac{c_l^2}{a_l^2}\Big)\phi_{lj} \\ & + h_{lj}^2\Big\{(n_1 - q - 1)\frac{(1-\phi_{lj})^2}{1-f_j} + 4(1-\phi_{lj})f_j\frac{\partial \phi_{lj}}{\partial f_j} \\ & + 2\sum_{k\neq j}(1-\phi_{lj})(\phi_{lj}-\phi_{lk})\frac{f_k}{f_j - f_k} + (n_2 - q - 1)c_l^2\frac{\phi_{lj}^2}{f_j} \\ & + 4c_l^2\phi_{lj}(1-f_j)\frac{\partial \phi_{lj}}{\partial f_j} + 2c_l^2\sum_{k\neq j}\phi_{lj}(\phi_{lj}-\phi_{lk})\frac{1-f_k}{f_j - f_k}\Big\}\end{aligned}$$

for $l = 1, \ldots, m$ and $j = 1, \ldots, q$. Here, note that

$$\frac{\partial \phi_{lj}}{\partial f_j} = f_j\frac{\partial}{\partial f_j}\left(\frac{\phi_{lj}}{f_j}\right) + \frac{\phi_{lj}}{f_j} = (1-f_j)\frac{\partial}{\partial(1-f_j)}\left(\frac{1-\phi_{lj}}{1-f_j}\right) + \frac{1-\phi_{lj}}{1-f_j}.$$

From the above relation without derivative terms, $\widehat{\boldsymbol{R}}_l$ can be approximated by

$$
\begin{aligned}
\widehat{\boldsymbol{R}}_l \approx & -q\hat{r}_{1l} + q\hat{r}_{2l}\frac{c_l^2}{a_l^2} + \sum_{j=1}^{q}\Big\{2\Big(\hat{r}_{1l} - \hat{r}_{2l}\frac{c_l^2}{a_l^2}\Big)\phi_{lj} + h_{lj}^2\Big\{(n_1 - q - 1)\frac{(1-\phi_{lj})^2}{1-f_j} \\
& +4\frac{(1-\phi_{lj})^2 f_j}{1-f_j} + 2\sum_{k\neq j}(1-\phi_{lj})(\phi_{lj}-\phi_{lk})\frac{f_k}{f_j - f_k} \\
& +(n_2 - q - 1)c_l^2\frac{\phi_{lj}^2}{f_j} + 4c_l^2\phi_{lj}^2\frac{1-f_j}{f_j} + 2c_l^2\sum_{k\neq j}\phi_{lj}(\phi_{lj}-\phi_{lk})\frac{1-f_k}{f_j-f_k}\Big\}\Big\} \\
= & -q\hat{r}_{1l} + q\hat{r}_{2l}\frac{c_l^2}{a_l^2} + \sum_{j=1}^{q}\Big\{2\Big(\hat{r}_{1l} - \hat{r}_{2l}\frac{c_l^2}{a_l^2}\Big)\phi_{lj} + h_{lj}^2\Big\{(n_1 - q - 1)\frac{(1-\phi_{lj})^2}{1-f_j} \\
& +4\frac{(1-\phi_{lj})^2 f_j}{1-f_j} - 2\sum_{k\neq j}\frac{(1-\phi_{lj})^2 f_k}{f_j - f_k} + 2\sum_{k\neq j}(1-\phi_{lj})(1-\phi_{lk})\frac{f_k}{f_j - f_k} \\
& +(n_2 - q - 1)c_l^2\frac{\phi_{lj}^2}{f_j} + 4c_l^2\phi_{lj}^2\frac{1-f_j}{f_j} + 2c_l^2\sum_{k\neq j}\phi_{lj}(\phi_{lj}-\phi_{lk})\frac{1-f_k}{f_j-f_k}\Big\}\Big\} \\
= & \widetilde{\boldsymbol{R}}_l, \quad \text{say.}
\end{aligned}
$$

Here we minimize $\widetilde{\boldsymbol{R}}_l$ with respect to ϕ_{lj} $(j = 1, \ldots, q)$, which yields

$$
\begin{aligned}
0 = \frac{1}{2} \times \frac{\partial \widetilde{\boldsymbol{R}}_l}{\partial \phi_{lj}} = & \hat{r}_{1l} - \hat{r}_{2l}\frac{c_l^2}{a_l^2} - (n_1 - q - 1)h_{lj}^2\frac{1-\phi_{lj}}{1-f_j} - 4h_{lj}^2(1-\phi_{lj})\frac{f_j}{1-f_j} \\
& +2h_{lj}^2(1-\phi_{lj})\sum_{k\neq j}\frac{f_k}{f_j - f_k} - h_{lj}^2\sum_{k\neq j}(1-\phi_{lk})\frac{f_k}{f_j-f_k} \\
& +(n_2 - q - 1)c_l^2 h_{lj}^2\frac{\phi_{lj}}{f_j} + 4c_l^2 h_{lj}^2\frac{1-f_j}{f_j}\phi_{lj} \\
& +2c_l^2 h_{lj}^2\phi_{lj}\sum_{k\neq j}\frac{1-f_k}{f_j-f_k} - c_l^2 h_{lj}^2\sum_{k\neq j}\phi_{lk}\frac{1-f_k}{f_j-f_k}.
\end{aligned}
$$

Hence, solving for ϕ_{lj} $(j = 1, \ldots, q)$ with ignoring the first, the second, the sixth, and the tenth terms in the last right-hand side above, we finally get

$$
\hat{\phi}_{lj}^{ST} = \frac{\hat{\beta}_j^{ST}}{\hat{\beta}_j^{ST} + c_l^2\hat{\alpha}_j^{ST}}, \tag{14}
$$

where

$$
\begin{aligned}
\hat{\alpha}_j^{ST} = & \Big((n_2 - q - 1) + 4(1-f_j) + 2\sum_{k\neq j}\frac{f_j(1-f_k)}{f_j - f_k}\Big)\Big/ f_j, \\
\hat{\beta}_j^{ST} = & \Big((n_1 - q - 1) + 4f_j - 2\sum_{k\neq j}\frac{(1-f_j)f_k}{f_j - f_k}\Big)\Big/(1-f_j)
\end{aligned}
$$

for $j = 1, \ldots, q$. Consequently, we take $t_l = c_l^2$, i.e., $\boldsymbol{T} = \boldsymbol{C}'\boldsymbol{C}$ for estimator of $\boldsymbol{\Theta}$.

Here, $\{\hat{\alpha}_j^{ST}\}_{j=1}^q$ and $\{\hat{\beta}_j^{ST}\}_{j=1}^q$ do not usually have the ordering properties (12). To correct this, we apply Stein's isotonic regression on $\{\hat{\alpha}_j^{ST}\}_{j=1}^q$ and $\{\hat{\beta}_j^{ST}\}_{j=1}^q$ to arrive at new sets $\{\bar{\alpha}_j^{ST}\}_{j=1}^q$ and $\{\bar{\beta}_j^{ST}\}_{j=1}^q$ which satisfy the natural ordering properties (12).

For a detailed description of Stein's isotonic regression, see Lin and Perlman (1985).

Therefore, we obtain

$$\begin{aligned}\mathrm{vec}(\widehat{\boldsymbol{\Theta}}^{ST}) = {} & [\boldsymbol{I}_m \otimes (\boldsymbol{B}' \,\mathbf{diag}\,(\bar{\beta}_j^{ST})\boldsymbol{B}) + (\boldsymbol{C}'\boldsymbol{C}) \otimes (\boldsymbol{B}' \,\mathbf{diag}\,(\bar{\alpha}_j^{ST})\boldsymbol{B})]^{-1} \\ & \times [\{\boldsymbol{I}_m \otimes (\boldsymbol{B}' \,\mathbf{diag}\,(\bar{\beta}_j^{ST})\boldsymbol{B})\}\mathrm{vec}(\widehat{\boldsymbol{\Theta}}_1) \\ & + \{(\boldsymbol{C}'\boldsymbol{C}) \otimes (\boldsymbol{B}' \,\mathbf{diag}\,(\bar{\alpha}_j^{ST})\boldsymbol{B})\}\mathrm{vec}(\boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2)], \qquad (15)\end{aligned}$$

where $\{\bar{\alpha}_j^{ST}\}_{j=1}^q$ and $\{\bar{\beta}_j^{ST}\}_{j=1}^q$ are made from Stein's isotonic regressions on $\{\hat{\alpha}_j^{ST}\}_{j=1}^q$ and on $\{\hat{\beta}_j^{ST}\}_{j=1}^q$, respectively.

4 Numerical studies

4.1 Numerical study for two GMANOVA models under normal errors

Since the risk of the Stein type estimator (15) is complicated, we have not been able to compare risks of the Stein type and the Graybill-Deal type estimators analytically. Therefore we investigate the risk performance of these estimators via a Monte-Carlo simulation.

Our simulation is based on 10,000 independent replications and these replications are generated from (7a)–(8d) with special cases for parameter matrices and (N_1, N_2, p_1, p_2, m, q). These results are given in Table 1.

For example, in case of $N_1 = N_2 = 12$, we assume that $\boldsymbol{A}^2 = \mathbf{diag}\,(1,\, 1)$ and $\boldsymbol{A}^2 = \mathbf{diag}\,(3,\, 1/3)$ are chosen in consideration of, respectively,

$$\boldsymbol{A}_{11} = \boldsymbol{A}_{21} = \begin{pmatrix} \mathbf{1}_6 & \mathbf{0}_6 \\ \mathbf{0}_6 & \mathbf{1}_6 \end{pmatrix}$$

and

$$\boldsymbol{A}_{11} = \begin{pmatrix} \mathbf{1}_3 & \mathbf{0}_3 \\ \mathbf{0}_9 & \mathbf{1}_9 \end{pmatrix} \quad \text{and} \quad \boldsymbol{A}_{21} = \begin{pmatrix} \mathbf{1}_9 & \mathbf{0}_9 \\ \mathbf{0}_3 & \mathbf{1}_3 \end{pmatrix}.$$

For $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$, we assume that the eigenvalues of $\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{-1}$ are close together and that these eigenvalues are widely spread out. Furthermore, we put $\boldsymbol{\Theta} = \mathbf{0}$, $\boldsymbol{\Lambda}_1 = \boldsymbol{\Lambda}_2 = \boldsymbol{I}_2$, and $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_2 = \mathbf{0}$.

In Table 1, "SK" and "ST" denote the Graybill-Deal type estimator (9) by Sugiura and Kubokawa (1988) and the Stein type estimator (15), respectively, and estimated standard errors are in parentheses. Furthermore, "AV" in Table 1 indicates the average of improvement in risk of ST against SK, i.e., $\text{AV} = 100(1 - \widehat{R}^{*ST}/\widehat{R}^{*SK})\,\%$, where $\widehat{R}^{*SK}$ and $\widehat{R}^{*ST}$ are, respectively, values of estimated risks for the Graybill-Deal type and the Stein type estimators by our simulations.

These simulation results are summarized as follows:

1. In Table 1, when the eigenvalues of $\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{-1}$ are close together, the AVs are large. Specially, in case when $\boldsymbol{A}^2 = \boldsymbol{C}'\boldsymbol{C} = \textbf{diag}\,(3, 1/3)$, $N_1 = N_2 = 20$, $p_1 = p_2 = 12$, $m = 2$, $q = 10$, and these eigenvalues are equal to 1, the AV is 26.3%.
2. On the contrary, when the eigenvalues of $\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{-1}$ are widely spread out, the AVs are negative. Furthermore, if one of these eigenvalues is extremely different from the others, it seems that the AV is equal to zero.
3. The AVs increase with increasing dimension q and fixed sample-size N.

Remark 2. Under another assumptions for $\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{-1}$ as examined by Loh (1991), we simulated the risk values of SK and ST and obtained the results that ST performs better than SK.

Remark 3. We also simulated risk values of the Kubokawa (1989) estimator which is an alternative estimator to the maximum likelihood estimator in the one-sample setup. Furthermore we simulated risk values of the estimator ST without Stein's isotonic regression. However, since these estimator performed worse than ST in most cases, we omitted risk values of these estimator for lack of space.

4.2 Numerical study for estimating the common mean under elliptical errors

First we illustrate the model (1) with the density (2) and the Graybill-Deal type and the Stein type estimators when $N_1 = N_2 = N$, $m = 1$, $p_1 = p_2 =$

Table 1: Estimated risks in GMANOVA models with normal errors
(Estimated standard errors are in parentheses)

Eigenvalues of $\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{-1}$	SK	ST	AV
$\boldsymbol{A}^2 = \boldsymbol{C}'\boldsymbol{C} =$ **diag (1, 1)**			
$\boldsymbol{N_1 = N_2 = 12,\ p_1 = p_2 = 7,\ m = 2,\ q = 5}$			
(1, 1, 1, 1, 1)	19.39 (0.106)	16.61 (0.090)	14.4 %
(10, 0.1, 0.1, 0.1, 0.1)	20.62 (0.132)	18.60 (0.116)	9.8 %
$(10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10})$	18.00 (0.156)	18.00 (0.156)	0.0 %
$(10^8, 10^4, 1, 10^{-4}, 10^{-8})$	20.27 (0.128)	20.44 (0.129)	−0.8 %
$\boldsymbol{N_1 = N_2 = 20,\ p_1 = p_2 = 12,\ m = 2,\ q = 10}$			
(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	33.67 (0.120)	26.06 (0.090)	22.6 %
(10, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)	36.07 (0.157)	28.75 (0.108)	20.3 %
$(10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10})$	27.52 (0.106)	27.52 (0.106)	0.0 %
$(10^5, 10^4, 10^3, 10^2, 10, 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4})$	33.97 (0.128)	34.01 (0.128)	−0.1 %
$\boldsymbol{A}^2 = \boldsymbol{C}'\boldsymbol{C} =$ **diag (3, 1/3)**			
$\boldsymbol{N_1 = N_2 = 12,\ p_1 = p_2 = 7,\ m = 2,\ q = 5}$			
(1, 1, 1, 1, 1)	20.47 (0.120)	16.94 (0.097)	17.3 %
(10, 0.1, 0.1, 0.1, 0.1)	20.33 (0.136)	18.41 (0.118)	9.4 %
$(10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10})$	18.00 (0.156)	18.00 (0.156)	0.0 %
$(10^8, 10^4, 1, 10^{-4}, 10^{-8})$	20.32 (0.128)	20.51 (0.130)	−0.9 %
$\boldsymbol{N_1 = N_2 = 20,\ p_1 = p_2 = 12,\ m = 2,\ q = 10}$			
(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	35.33 (0.135)	26.04 (0.092)	26.3 %
(10, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)	34.85 (0.149)	28.40 (0.105)	18.5 %
$(10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10})$	27.52 (0.106)	27.52 (0.106)	0.0 %
$(10^5, 10^4, 10^3, 10^2, 10, 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4})$	33.95 (0.128)	33.94 (0.128)	0.0 %

$q_1 = q_2 = p$, $\boldsymbol{A}_{11} = \boldsymbol{A}_{21} = \mathbf{1}_N$ and $\boldsymbol{A}_{12} = \boldsymbol{A}_{22} = \boldsymbol{I}_p$. Then, from Theorem 1, we can write the density function (2) as

$$|\boldsymbol{\Sigma}_1|^{-N/2}|\boldsymbol{\Sigma}_2|^{-N/2} g\Big(\sum_{i=1}^{2}[\,\mathrm{tr}\,\{\boldsymbol{\Sigma}_i^{-1}(\boldsymbol{X}_i-\boldsymbol{\theta})(\boldsymbol{X}_i-\boldsymbol{\theta})' + \boldsymbol{\Sigma}_i^{-1}\boldsymbol{S}_i\}]\Big), \tag{16}$$

where $\boldsymbol{\theta}$ and $\boldsymbol{X}_i$ are $p\times 1$ vectors and $\boldsymbol{\Sigma}_i$ and $\boldsymbol{S}_i$ are $p\times p$ matrices. Therefore, the problem of estimating $\boldsymbol{\Xi}$ in (2) turns into that of estimating the common mean vector $\boldsymbol{\theta}$ in (16). Then, the Graybill-Deal type estimator can be written as

$$\hat{\boldsymbol{\theta}}^{GD} = (\boldsymbol{S}_1^{-1} + \boldsymbol{S}_2^{-1})^{-1}(\boldsymbol{S}_1^{-1}\boldsymbol{X}_1 + \boldsymbol{S}_2^{-1}\boldsymbol{X}_2) \tag{17}$$

and also the Stein type estimator as

$$\hat{\boldsymbol{\theta}}^{ST} = \boldsymbol{B}^{-1}\boldsymbol{\Phi}^{ST}\boldsymbol{B}\boldsymbol{X}_1 + \boldsymbol{B}^{-1}(\boldsymbol{I}_p - \boldsymbol{\Phi}^{ST})\boldsymbol{B}\boldsymbol{X}_2, \tag{18}$$

where $\boldsymbol{\Phi}^{ST} = \mathbf{diag}\,(\bar{\phi}_1^{ST}, \cdots, \bar{\phi}_p^{ST})$, $\bar{\phi}_j^{ST} = \bar{\beta}_j^{ST}/(\bar{\beta}_j^{ST} + \bar{\alpha}_j^{ST})$. Here $\{\bar{\alpha}_j^{ST}\}_{j=1}^q$ and $\{\bar{\beta}_j^{ST}\}_{j=1}^q$ are given by (15).

Since the model (16) is not i.i.d. sampling set-up of two sample problems, we carry out Monte-Carlo simulation to show that our proposed estimator (18) reduces the risk over the Graybill-Deal estimator (17) under the i.i.d. sampling from two independent multivariate elliptically contoured distributions instead of sampling from the model (16). Hence, we carry out Monte-Carlo simulation when we sample $(\boldsymbol{Y}_1, \boldsymbol{Y}_2)$ which can be represented as

$$\boldsymbol{Y}_1 = \mathbf{1}_N\boldsymbol{\xi}' + \boldsymbol{\epsilon}_1 \qquad \text{and} \qquad \boldsymbol{Y}_2 = \mathbf{1}_N\boldsymbol{\xi}' + \boldsymbol{\epsilon}_2,$$

where $\boldsymbol{Y}_1$, $\boldsymbol{Y}_2$, $\boldsymbol{\epsilon}_1$, and $\boldsymbol{\epsilon}_2$ are $N\times p$ random matrices and $\boldsymbol{\xi}$ is a $p\times 1$ unknown vector. Here, the rows of $\boldsymbol{\epsilon}_i$ have densities

$$|\boldsymbol{\Sigma}_i|^{-N/2}h(\boldsymbol{e}_{ij}'\boldsymbol{\Sigma}_i^{-1}\boldsymbol{e}_{ij}), \qquad i=1,\,2,\;\; j=1,\ldots,N, \tag{19}$$

where $\boldsymbol{\epsilon}_i = (\boldsymbol{e}_{i1}, \boldsymbol{e}_{i2}, \ldots, \boldsymbol{e}_{iN})'$ and h is an unknown, positive-valued function on $[0, \infty)$. That is, it means that the rows of each error matrix $\boldsymbol{\epsilon}_i$ are independently and identically distributed (i.i.d.) as an elliptically contoured distribution. As it is difficult to derive an improved estimator under the density function (19), we consider an improvement under density (16). However, our simulation results justify our derivation of alternative estimator under the model (16).

For Monte Carlo simulations, we suppose that $\boldsymbol{e}_{ij}$, $i = 1, 2$, $j = 1, 2, \ldots, N$, follow the multivariate t-distribution whose density function is given by

$$\kappa_1\,|\boldsymbol{\Sigma}_i|^{-1/2}(1 + \boldsymbol{e}_{ij}'\boldsymbol{\Sigma}_i^{-1}\boldsymbol{e}_{ij}/v)^{-(v+p)/2},$$

where $v > 0$ and $\kappa_1 = \Gamma[(v+p)/2]/\{(\pi v)^{p/2}\Gamma[v/2]\}$, and we also suppose that $\boldsymbol{e}_{ij}$, $i = 1,\ 2,\ j = 1,\ 2,\ \ldots,N$, follow the vector-valued Kotz-type distribution whose density function is given by

$$\kappa_2|\boldsymbol{\Sigma}_i|^{-1/2}\{\boldsymbol{e}_{ij}'\boldsymbol{\Sigma}_i^{-1}\boldsymbol{e}_{ij}\}^{u-1}\exp[-r\{\boldsymbol{e}_{ij}'\boldsymbol{\Sigma}_i^{-1}\boldsymbol{e}_{ij}\}^s],$$

where $r > 0,\ s > 0,\ 2u+p > 2$, and

$$\kappa_2 = \frac{s\Gamma[p/2]r^{\{u+p/2-1\}/s}}{\pi^{p/2}\Gamma[\{u+p/2-1\}/s]}.$$

For generating a random number of the Kotz-type distribution above, see Fang, Kotz, and Ng (1990) for example.

In our simulations, we assume that $\boldsymbol{\xi} = \boldsymbol{0}$ and that $\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{-1}$ is a diagonal matrix with typical elements. We also take $(N,\ p) = (8,\ 5)$ and $(13,\ 10)$ and put $v = 3$ for t-distribution and $(u, r, s) = (5, 0.5, 2)$ for Kotz-type distributions. These simulation results are given in Tables 2 and 3, respectively. In tables, "GD" and "ST" denote $\hat{\boldsymbol{\theta}}^{GD}$ and $\hat{\boldsymbol{\theta}}^{ST}$, respectively, and "AV" is the average of improvement in risk of ST against GD. We summarize these results as follows:

Table 2: Estimated risks under t-distributions with $v = 3$
(Estimated standard errors are in parentheses)

Eigenvalues of $\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{-1}$	GD	ST	AV
$\boldsymbol{N = 8,\quad p = 5}$			
$(1,\ 1,\ 1,\ 1,\ 1)$	26.927 (1.272)	24.271 (1.423)	9.86 %
$(10,\ 0.1,\ 0.1,\ 0.1,\ 0.1)$	32.441 (2.726)	28.716 (2.292)	11.48 %
$(10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10})$	29.349 (2.148)	29.349 (2.148)	0.00 %
$(10^8, 10^4, 1, 10^{-4}, 10^{-8})$	29.434 (0.992)	29.784 (1.017)	−1.19 %
$\boldsymbol{N = 13,\quad p = 10}$			
$(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$	62.767 (2.137)	50.855 (1.713)	18.98 %
$(10, 0.1, 0.1, 0.1, 0.1,$ $0.1, 0.1, 0.1, 0.1, 0.1)$	84.003 (2.670)	60.471 (1.803)	28.01 %
$(10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10},$ $10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10})$	47.988 (1.340)	47.988 (1.340)	0.00 %
$(10^5, 10^4, 10^3, 10^2, 10,$ $1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4})$	63.451 (2.292)	63.348 (2.267)	0.16 %

1. In almost cases, the AVs are positive. These are large when the eigenvalues of $\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{-1}$ are close together, and particularly, when only one of these

Table 3: Estimated risks under Kotz-type distributions with $(u, r, s) = (5, 0.5, 2)$
(Estimated standard errors are in parentheses)

Eigenvalues of $\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{-1}$	GD	ST	AV
$N = 8, \quad p = 5$			
(1, 1, 1, 1, 1)	3.755 (0.021)	3.143 (0.017)	16.32 %
(10, 0.1, 0.1, 0.1, 0.1)	3.970 (0.027)	3.572 (0.023)	10.01 %
$(10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10})$	3.509 (0.029)	3.509 (0.029)	0.00 %
$(10^{8}, 10^{4}, 1, 10^{-4}, 10^{-8})$	3.846 (0.024)	3.878 (0.024)	−0.83 %
$N = 13, \quad p = 10$			
$(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$	5.027 (0.022)	3.749 (0.015)	25.43 %
$(10, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)$	6.151 (0.038)	4.484 (0.024)	27.11 %
$(10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10})$	4.460 (0.064)	4.460 (0.064)	0.00 %
$(10^{5}, 10^{4}, 10^{3}, 10^{2}, 10, 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4})$	5.107 (0.024)	5.107 (0.024)	0.00 %

eigenvalues is 10 with $(N, p) = (13, 10)$, the AVs are more than 27%.

2. On the contrary, when the eigenvalues of $\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{-1}$ are spread out, the AVs are small.
3. Furthermore, the AVs are negative when these eigenvalues are extremely spread out. However, since the negative AVs are about -1% and $\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{-1}$ are extreme, the use of ST is more effective than that of GD in a sense.
4. From Tables 2–3, so long as the eigenvalues of $\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{-1}$ are the same, it is expected that the AVs increase with increasing dimension p and small sample-size N.
5. Tables 2 and 3 indicate that the AVs are substantial under independently and identically sampling set-up from non-normal distribution, although we cannot derive ST under this situation. Hence, these results suggest that the improvement under density (16) remains robust even if the rows of errors are i.i.d.

Remark 4. Another simulations were conducted for sample size $N = 50$ and then, under both t and Kotz-type distributions, the maximum of AV's was about 4 % when $p = 5$ and about 8 % when $p = 10$.

5 Proof of Theorem 2

We first state useful lemmas to prove Theorem 2. Put $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m)'$, $\widehat{\boldsymbol{\Theta}}_i = (\hat{\boldsymbol{\theta}}_{i1}, \ldots, \hat{\boldsymbol{\theta}}_{im})'$, $\boldsymbol{X}_i = (\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{im})'$, and $\boldsymbol{Z}_i = (\boldsymbol{z}_{i1}, \ldots, \boldsymbol{z}_{im})'$ for $i =$ 1, 2. Here, $\boldsymbol{\theta}_l$, $\hat{\boldsymbol{\theta}}_{il}$, and $\boldsymbol{x}_{il}$ are $q \times 1$ vectors and $\boldsymbol{z}_{il}$ are $(p_i - q) \times 1$ vectors $(i = 1, 2,\ l = 1, \ldots, m)$. Let $\boldsymbol{q}_{il}$ be $q \times 1$ vector-valued functions of $\boldsymbol{x}_{il} = (x_{il \cdot j})$. Also, for $i = 1, 2$, let $\boldsymbol{K}_i \equiv \boldsymbol{K}_i(\boldsymbol{u}_i)$ be $q \times (p_i - q)$ matrix-valued functions of $\boldsymbol{u}_i = (u_{i \cdot j_i k})$. Denote differential operators in terms of $\boldsymbol{x}_{il}$ and $\boldsymbol{u}_i$ by

$$\nabla_{x_{il}} = \left(\frac{\partial}{\partial x_{il \cdot j}} \right) \qquad \text{and} \qquad \nabla_{u_i} = \left(\frac{\partial}{\partial u_{i \cdot j_i k}} \right).$$

Here, the actions of $\nabla_{x_{il}}$ on $\boldsymbol{q}_{il} = (\boldsymbol{q}_{il \cdot b})$ and of ∇_{u_i} on $\boldsymbol{K}_i = (\boldsymbol{K}_{i \cdot kd})$ are defined as

$$\{\nabla_{x_{il}} \boldsymbol{q}_{il}'\}_{ab} = \frac{\partial \boldsymbol{q}_{il \cdot b}}{\partial x_{il \cdot a}} \qquad \text{and} \qquad \{\nabla_{u_i} \boldsymbol{K}_i\}_{cd} = \left(\sum_{k=1}^{q} \frac{\partial \boldsymbol{K}_{i \cdot kd}}{\partial u_{i \cdot ck}} \right).$$

Lemma 2. *Let $\boldsymbol{G}$ and $\boldsymbol{H}_i$ be, respectively, $q \times q$ and $(p_i - q) \times (p_i - q)$ matrices of constants for $i = 1$, 2. Then* $\mathrm{tr}\,\{\nabla_{u_i} \boldsymbol{G} \boldsymbol{u}_i' \boldsymbol{H}_i\} = (\,\mathrm{tr}\,\{\boldsymbol{G}\})(\,\mathrm{tr}\,\{\boldsymbol{H}_i\})$.

Lemma 3 (Kubokawa and Srivastava, 2001). *For $i = 1, 2$, $l = 1, \ldots, m$, and $j = 1, \ldots, q$, suppose that each element of $\boldsymbol{q}_{il} \equiv \boldsymbol{q}_{il}(\boldsymbol{x}_{il})$ is differentiable with respect to $x_{il \cdot j}$. Also, for $i = 1, 2$, $j_i = 1, \ldots, p_i - q$, $k = 1, \ldots, q$, suppose that elements of $\boldsymbol{K}_i \equiv \boldsymbol{K}_i(\boldsymbol{u}_i)$ are differentiable with respect to $u_{i \cdot j_i k}$. Furthermore, assume that*

(i) *there exist finite expectations of the absolute values of* $(\boldsymbol{x}_{1l} - \boldsymbol{\gamma}_1' \boldsymbol{z}_{1l} - \boldsymbol{\theta}_l)' \boldsymbol{\Sigma}_1^{-1} \boldsymbol{q}_{1l}$, $(\boldsymbol{x}_{2l} - \boldsymbol{\gamma}_2' \boldsymbol{z}_{2l} - a_l \boldsymbol{\theta}_l)' \boldsymbol{\Sigma}_2^{-1} \boldsymbol{q}_{2l}$, *and each element of* $(\boldsymbol{u}_i - \boldsymbol{W}_i^{1/2} \gamma_i) \boldsymbol{\Sigma}_i^{-1} \boldsymbol{K}_i$ *;*
(ii) $\lim_{x_{il \cdot j} \to \pm\infty} \boldsymbol{q}_{il}(\boldsymbol{x}_{il}) G(x_{il \cdot j}^2 + a^2) = \boldsymbol{0}$ *for $i = 1, 2$, $j = 1, \ldots, q$, $l = 1, \ldots, m$ and any real a;*
(iii) $\lim_{u_{i \cdot j_i k} \to \pm\infty} \boldsymbol{K}_i(\boldsymbol{u}_i) G(u_{i \cdot j_i k}^2 + a^2) = \boldsymbol{0}$ *for $i = 1, 2$, $j_i = 1, \ldots, p_i - q$, $k = 1, \ldots, q$ and any real a.*

Then, for $i = 1, 2$ and $l = 1, \ldots, m$, we have

$$\begin{aligned}
&\mathbb{E}[(\boldsymbol{x}_{1l} - \boldsymbol{\gamma}_1' \boldsymbol{z}_{1l} - \boldsymbol{\theta}_l)' \boldsymbol{\Sigma}_1^{-1} \boldsymbol{q}_{1l}] = \mathbb{E}_G[\,\mathrm{tr}\,(\nabla_{x_{1l}} \boldsymbol{q}_{1l}')],\\
&\mathbb{E}[(\boldsymbol{x}_{2l} - \boldsymbol{\gamma}_2' \boldsymbol{z}_{2l} - a_l \boldsymbol{\theta}_l)' \boldsymbol{\Sigma}_2^{-1} \boldsymbol{q}_{2l}] = \mathbb{E}_G[\,\mathrm{tr}\,(\nabla_{x_{2l}} \boldsymbol{q}_{2l}')],\\
&\mathbb{E}[\,\mathrm{tr}\,\{(\boldsymbol{u}_i - \boldsymbol{W}_i^{1/2} \gamma_i) \boldsymbol{\Sigma}_i^{-1} \boldsymbol{K}_i\}] = \mathbb{E}_G[\,\mathrm{tr}\,(\nabla_{u_i} \boldsymbol{K}_i)].
\end{aligned}$$

From Lemmas 2 and 3, we immediately have the followings:

Lemma 4. *For $i = 1, 2$ and $l = 1, \ldots, q$, let $\hat{r}_{il} = 1 + \boldsymbol{z}_{il}'\boldsymbol{W}_i^{-1}\boldsymbol{z}_{il}$ and $\boldsymbol{\Phi}_l = \mathbf{diag}\,(\phi_{l1}, \ldots, \phi_{lq})$. Then*

$$\mathbb{E}[(\hat{\boldsymbol{\theta}}_{1l} - \boldsymbol{\theta}_l)'\boldsymbol{\Sigma}_1^{-1}(\hat{\boldsymbol{\theta}}_{1l} - \boldsymbol{\theta}_l)] = \mathbb{E}_G[q\hat{r}_{1l}], \tag{20a}$$

$$\mathbb{E}[(a_l^{-1}\hat{\boldsymbol{\theta}}_{2l} - \boldsymbol{\theta}_l)'\boldsymbol{\Sigma}_2^{-1}(a_l^{-1}\hat{\boldsymbol{\theta}}_{2l} - \boldsymbol{\theta}_l)] = \mathbb{E}_G[a_l^{-2}q\hat{r}_{2l}], \tag{20b}$$

$$\mathbb{E}[(\hat{\boldsymbol{\theta}}_{1l} - \boldsymbol{\theta}_l)'\boldsymbol{\Sigma}_1^{-1}\boldsymbol{B}^{-1}(\boldsymbol{I}_q - \boldsymbol{\Phi}_l)\boldsymbol{B}(a_l^{-1}\hat{\boldsymbol{\theta}}_{2l} - \hat{\boldsymbol{\theta}}_{1l})] = \mathbb{E}_G\Big[-\hat{r}_{1l}\sum_{j=1}^{q}(1 - \phi_{lj})\Big], \tag{20c}$$

$$\mathbb{E}[(a_l^{-1}\hat{\boldsymbol{\theta}}_{2l} - \boldsymbol{\theta}_l)'\boldsymbol{\Sigma}_2^{-1}\boldsymbol{B}^{-1}\boldsymbol{\Phi}_l\boldsymbol{B}(\hat{\boldsymbol{\theta}}_{1l} - a_l^{-1}\hat{\boldsymbol{\theta}}_{2l})'] = \mathbb{E}_G\Big[-a_l^{-2}\hat{r}_{2l}\sum_{j=1}^{q}\phi_{lj}\Big]. \tag{20d}$$

Proof. Note that the density function (4) is symmetric at $\boldsymbol{x}_{1l} - \boldsymbol{\gamma}_1'\boldsymbol{z}_{1l} - \boldsymbol{\theta}_l = \boldsymbol{0}$ and $\boldsymbol{x}_{2l} - \boldsymbol{\gamma}_2'\boldsymbol{z}_{2l} - a_l\boldsymbol{\theta}_l = \boldsymbol{0}$ $(l = 1, \ldots, q)$ and at $\boldsymbol{u}_i - \boldsymbol{W}_i^{1/2}\gamma_i = \boldsymbol{0}$ $(i = 1, 2)$.

For (20a), we observe that

$$\begin{aligned}
&\mathbb{E}[(\hat{\boldsymbol{\theta}}_{1l} - \boldsymbol{\theta}_l)'\boldsymbol{\Sigma}_1^{-1}(\hat{\boldsymbol{\theta}}_{1l} - \boldsymbol{\theta}_l)]\\
&= \mathbb{E}[(\boldsymbol{x}_{1l} - \widehat{\gamma}_1'\boldsymbol{z}_{1l} - \boldsymbol{\theta}_l)'\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{x}_{1l} - \widehat{\gamma}_1'\boldsymbol{z}_{1l} - \boldsymbol{\theta}_l)]\\
&= \mathbb{E}[(\boldsymbol{x}_{1l} - \boldsymbol{\gamma}_1'\boldsymbol{z}_{1l} - \boldsymbol{\theta}_l)'\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{x}_{1l} - \boldsymbol{\gamma}_1'\boldsymbol{z}_{1l} - \boldsymbol{\theta}_l)]\\
&\qquad +2\mathbb{E}[(\boldsymbol{x}_{1l} - \boldsymbol{\gamma}_1'\boldsymbol{z}_{1l} - \boldsymbol{\theta}_l)'\boldsymbol{\Sigma}_1^{-1}(\gamma_1 - \widehat{\gamma}_1)'\boldsymbol{z}_{1l}]\\
&\qquad +\mathbb{E}[\,\mathrm{tr}\,\{(\boldsymbol{u}_1 - \boldsymbol{W}_1^{1/2}\gamma_1)\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{u}_1 - \boldsymbol{W}_1^{1/2}\gamma_1)'\boldsymbol{W}_1^{-1/2}\boldsymbol{z}_{1l}\boldsymbol{z}_{1l}'\boldsymbol{W}_1^{-1/2}\}].
\end{aligned}$$

Here the second term of the right-hand side in the above equation is zero. Hence, from Lemma 3, we get the right-hand side of (20a).

By the similar way, we have (20b). For (20c), we can see from symmetry of density function that

$$\begin{aligned}
&\mathbb{E}[(\hat{\boldsymbol{\theta}}_{1l} - \boldsymbol{\theta}_l)'\boldsymbol{\Sigma}_1^{-1}\boldsymbol{B}^{-1}(\boldsymbol{I}_q - \boldsymbol{\Phi}_l)\boldsymbol{B}(a_l^{-1}\hat{\boldsymbol{\theta}}_{2l} - \hat{\boldsymbol{\theta}}_{1l})]\\
&= -\mathbb{E}[(\boldsymbol{x}_{1l} - \boldsymbol{\gamma}_1'\boldsymbol{z}_{1l} - \boldsymbol{\theta}_l)'\boldsymbol{\Sigma}_1^{-1}\boldsymbol{B}^{-1}(\boldsymbol{I}_q - \boldsymbol{\Phi}_l)\boldsymbol{B}(\boldsymbol{x}_{1l} - \boldsymbol{\gamma}_1'\boldsymbol{z}_{1l} - \boldsymbol{\theta}_l)]\\
&\quad -\mathbb{E}[\,\mathrm{tr}\,\{(\boldsymbol{u}_1 - \boldsymbol{W}_1^{1/2}\gamma_1)\boldsymbol{\Sigma}_1^{-1}\boldsymbol{B}^{-1}(\boldsymbol{I}_q - \boldsymbol{\Phi}_l)\boldsymbol{B}\\
&\qquad\qquad \times(\boldsymbol{u}_1 - \boldsymbol{W}_1^{1/2}\gamma_1)'\boldsymbol{W}_1^{-1/2}\boldsymbol{z}_{1l}\boldsymbol{z}_{1l}'\boldsymbol{W}_1^{-1/2}\}].
\end{aligned}$$

Thus, from Lemmas 2 and 3, we get the right-hand side of (20c). The derivation of (20d) is similar to that of (20c). □

For $i = 1, 2$, let $\boldsymbol{V}_i \equiv \boldsymbol{V}_i(\boldsymbol{S}_1,\, \boldsymbol{S}_2) = (v_{i\cdot jk})$ be $q \times q$ matrices such that the (j, k)-elements $v_{i\cdot jk}$ are functions of $\boldsymbol{S}_1 = (s_{1\cdot jk})$ and $\boldsymbol{S}_2 = (s_{2\cdot jk})$. For $i = 1, 2$, let

$$\{\mathcal{D}_i \boldsymbol{V}_i\}_{jk} = \sum_{a=1}^{q} d_{i\cdot ja} v_{i\cdot ak}, \qquad i = 1, 2, \tag{21}$$

where $d_{i\cdot ja} = (1/2)(1 + \delta_{ja})(\partial/\partial s_{i\cdot ja})$ with $\delta_{ja} = 1$ for $j = a$ and $\delta_{ja} = 0$ for $j \neq a$. Also put $\boldsymbol{s}_i = (\boldsymbol{s}'_{i1}, \ldots, \boldsymbol{s}'_{in_i})'$ and $\boldsymbol{s}_{ij_i} = (s_{i\cdot j_i 1}, \ldots, s_{i\cdot j_i q})$ for $i = 1$, 2 and $j_i = 1, 2, \ldots, n_i$. Hence we have $\boldsymbol{S}_i = \boldsymbol{s}'_i \boldsymbol{s}_i = \sum_{j_i=1}^{n_i} \boldsymbol{s}'_{ij_i} \boldsymbol{s}_{ij_i}$ for $i = 1$, 2.

Lemma 5 (Kubokawa and Srivastava, 1999). *Let*

$$\boldsymbol{V}_i \equiv \boldsymbol{V}_i \left(\sum_{j_1=1}^{n_1} \boldsymbol{s}'_{1j_1} \boldsymbol{s}_{1j_1}, \sum_{j_2=1}^{n_2} \boldsymbol{s}'_{2j_2} \boldsymbol{s}_{2j_2} \right), \qquad i = 1, 2,$$

be $q \times q$ matrices whose elements are differentiable with respect to $s_{i\cdot j_i k}$ ($j_i = 1, 2, \ldots, n_i$, $k = 1, 2, \ldots, q$). Furthermore, assume that

(a) $\mathbb{E}\Big[\big|\operatorname{tr}(\boldsymbol{V}_i \boldsymbol{\Sigma}_i^{-1})\big|\Big]$ *($i = 1, 2$) is finite;*

(b) $\displaystyle\lim_{s_{i\cdot j_i k} \to \pm\infty} |s_{i\cdot j_i k}| \boldsymbol{V}_i \cdot \left(\sum_{j_i=1}^{n_i} \boldsymbol{s}'_{1j_i} \boldsymbol{s}_{1j_i} \right)^{-1} G(s^2_{i\cdot j_i k} + a) = \boldsymbol{0}$ *for any real a.*

Then we have

$$\mathbb{E}\left[\sum_{i=1}^{2} \operatorname{tr}(\boldsymbol{\Sigma}_i^{-1} \boldsymbol{V}_i) \right] = \mathbb{E}_G \left[\sum_{i=1}^{2} \Big\{ (n_i - q - 1) \operatorname{tr}(\boldsymbol{S}_i^{-1} \boldsymbol{V}_i) + 2 \operatorname{tr}(\mathcal{D}_i \boldsymbol{V}_i) \Big\} \right].$$

Lemma 6 (Loh, 1988 and 1991). *For $i = 1, 2$, let $\mathcal{D}_i$ be $q \times q$ differential operators which are defined by (21). Also let $\boldsymbol{x}$ be a $q \times 1$ vector which is independent of $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ and let $\boldsymbol{\Phi} =$ **diag** $(\phi_1, \ldots, \phi_q)$ whose elements are functions of $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$. Then*

$$\begin{aligned}
&\operatorname{tr}\{\mathcal{D}_1[\boldsymbol{B}^{-1}(\boldsymbol{I}_q - \boldsymbol{\Phi})\boldsymbol{B}\boldsymbol{x}\boldsymbol{x}'\boldsymbol{B}'(\boldsymbol{I}_q - \boldsymbol{\Phi})(\boldsymbol{B}')^{-1}]\} \\
&\quad = \sum_{j=1}^{q} \Bigg[\{\boldsymbol{B}\boldsymbol{x}\}_j^2 (1 - \phi_j)^2 \sum_{k \neq j} \frac{f_k}{f_k - f_j} + 2\{\boldsymbol{B}\boldsymbol{x}\}_j^2 (1 - \phi_j) f_j \frac{\partial \phi_j}{\partial f_j} \\
&\qquad\qquad - \sum_{k \neq j} \{\boldsymbol{B}\boldsymbol{x}\}_k^2 (1 - \phi_j)(1 - \phi_k) \frac{f_j}{f_j - f_k} \Bigg], \\
&\operatorname{tr}\{\mathcal{D}_2[\boldsymbol{B}^{-1}\boldsymbol{\Phi}\boldsymbol{B}\boldsymbol{x}\boldsymbol{x}'\boldsymbol{B}'\boldsymbol{\Phi}(\boldsymbol{B}')^{-1}]\} \\
&\quad = \sum_{j=1}^{q} \Bigg[\{\boldsymbol{B}\boldsymbol{x}\}_j^2 \phi_j^2 \sum_{k \neq j} \frac{1 - f_k}{f_j - f_k} + 2\{\boldsymbol{B}\boldsymbol{x}\}_j^2 \phi_j (1 - f_j) \frac{\partial \phi_j}{\partial f_j} \\
&\qquad\qquad - \sum_{k \neq j} \{\boldsymbol{B}\boldsymbol{x}\}_k^2 \phi_j \phi_k \frac{1 - f_j}{f_k - f_j} \Bigg],
\end{aligned}$$

where $\{\boldsymbol{B}\boldsymbol{x}\}_j$ denote the j-th elements of $\boldsymbol{B}\boldsymbol{x}$.

Proof of Theorem 2. From define of $\widehat{\mathbf{\Theta}}^{RE} = (\hat{\boldsymbol{\theta}}_1^{RE}, \ldots, \hat{\boldsymbol{\theta}}_m^{RE})'$ given by (11), we can see that

$$\hat{\boldsymbol{\theta}}_l^{RE} = \boldsymbol{B}^{-1}\mathbf{\Phi}_l\boldsymbol{B}\hat{\boldsymbol{\theta}}_{1l} + \boldsymbol{B}^{-1}(\boldsymbol{I}_q - \mathbf{\Phi}_l)\boldsymbol{B}(a_l^{-1}\hat{\boldsymbol{\theta}}_{2l}),$$

where $\mathbf{\Phi}_l = \mathbf{diag}\,(\phi_{l1}, \ldots, \phi_{lq})$ with $\phi_{lj} = \beta_j/(\beta_j + t_l\alpha_j)$ for $l = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, q$ and $\boldsymbol{T} = \mathbf{diag}\,(t_1, \ldots, t_m)$ with $t_l > 0$, $l = 1, 2, \ldots, m$. Then the risk of $\widehat{\mathbf{\Theta}}^{RE}$ can be written as

$$\begin{aligned}
&R((\mathbf{\Theta}, \mathbf{\Sigma}_1, \mathbf{\Sigma}_2), \widehat{\mathbf{\Theta}}^{RE}) \\
&= \sum_{l=1}^{m} \mathbb{E}[(\hat{\boldsymbol{\theta}}_l^{RE} - \boldsymbol{\theta}_l)'(\mathbf{\Sigma}_1^{-1} + c_l^2\mathbf{\Sigma}_2^{-1})(\hat{\boldsymbol{\theta}}_l^{RE} - \boldsymbol{\theta}_l)] \\
&= \sum_{l=1}^{m} \mathbb{E}[(\hat{\boldsymbol{\theta}}_{1l} - \boldsymbol{\theta}_l)'\mathbf{\Sigma}_1^{-1}(\hat{\boldsymbol{\theta}}_{1l} - \boldsymbol{\theta}_l) \\
&\qquad +2(\hat{\boldsymbol{\theta}}_{1l} - \boldsymbol{\theta}_l)'\mathbf{\Sigma}_1^{-1}\boldsymbol{B}^{-1}(\boldsymbol{I}_q - \mathbf{\Phi}_l)\boldsymbol{B}(a_l^{-1}\hat{\boldsymbol{\theta}}_{2l} - \hat{\boldsymbol{\theta}}_{1l}) \\
&\qquad + \mathrm{tr}\,\{\mathbf{\Sigma}_1^{-1}\boldsymbol{B}^{-1}(\boldsymbol{I}_q - \mathbf{\Phi}_l)\boldsymbol{B}(a_l^{-1}\hat{\boldsymbol{\theta}}_{2l} - \hat{\boldsymbol{\theta}}_{1l})(a_l^{-1}\hat{\boldsymbol{\theta}}_{2l} - \hat{\boldsymbol{\theta}}_{1l})'\boldsymbol{B}'(\boldsymbol{I}_q - \mathbf{\Phi}_l)\boldsymbol{B}'^{-1}\} \\
&\qquad +c_l^2(a_l^{-1}\hat{\boldsymbol{\theta}}_{2l} - \boldsymbol{\theta}_l)'\mathbf{\Sigma}_2^{-1}(a_l^{-1}\hat{\boldsymbol{\theta}}_{2l} - \boldsymbol{\theta}_l) \\
&\qquad +2c_l^2(a_l^{-1}\hat{\boldsymbol{\theta}}_{2l} - \boldsymbol{\theta}_l)'\mathbf{\Sigma}_2^{-1}\boldsymbol{B}^{-1}\mathbf{\Phi}_l\boldsymbol{B}(\hat{\boldsymbol{\theta}}_{1l} - a_l^{-1}\hat{\boldsymbol{\theta}}_{2l})' \\
&\qquad +c_l^2\,\mathrm{tr}\,\{\mathbf{\Sigma}_2^{-1}\boldsymbol{B}^{-1}\mathbf{\Phi}_l\boldsymbol{B}(\hat{\boldsymbol{\theta}}_{1l} - a_l^{-1}\hat{\boldsymbol{\theta}}_{2l})(\hat{\boldsymbol{\theta}}_{1l} - a_l^{-1}\hat{\boldsymbol{\theta}}_{2l})'\boldsymbol{B}'\mathbf{\Phi}_l\boldsymbol{B}'^{-1}\}] \qquad (22)
\end{aligned}$$

First apply Lemmas 4 and 5 to the risk (22) and next use Lemma 6 to get the desired result. □

Acknowledgement

The authors thank two anonymous referees for their careful reading and valuable comments.

References

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

[2] Bilodeau, M. and Kariya, T. (1989). Minimax estimators in the MANOVA models. *J. Multivariate Anal.* **28**, 260–270.

[3] Fang, K.T., Kotz, S., and Ng, K.W. (1990). *Symmetric multivariate and related distributions*. Chapman and Hall, New York.

[4] Graybill, F.A. and Deal, R.B. (1959). Combined unbiased estimator. *Biometrics* **15**, 543–550.

[5] Haff, L.R. (1991). The variational forms of certain Bayes estimators. *Ann. Statist.* **18**, 1163–1190.

[6] James, W. and Stein, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1**, 361–380. Univ. of California Press.

[7] Kariya, T. (1989). Equivariant estimation in a model with an ancillary statistic. *Ann. Statist.* **17** 920–928.

[8] Kariya, T., Konno, Y., and Strawderman, W.E. (1996). Double shrinkage estimators in the GMANOVA model. *J. Multivariate Anal.* **56**, 245–258.

[9] Kariya, T., Konno, Y., and Strawderman, W.E. (1999). Construction of shrinkage estimators for the regression coefficient matrix in the GMANOVA model. *Comm. Statist. A-Theory and Methods* **28**, 597–611.

[10] Kubokawa, T. (1989). *Estimating common parameters of growth curve models under a quadratic loss. Comm. Statist. A-Theory and Methods* **18**, 3149–3155.

[11] Kubokawa, T. and Srivastava, M.S. (1999). Robust improvement in estimation of a covariance matrix in an elliptically contoured distribution. *Ann. Statist.* **27**, 600–609.

[12] Kubokawa, T. and Srivastava, M.S. (2001). Robust improvement in estimation of a mean matrix in an elliptically contoured distribution. *J. Multivariate Anal.* **76**, 138–152.

[13] Lin, S.P. and Perlman, M.D. (1985). *A Monte Carlo comparisons of four estimators for a covariance matrix.* In *Multivariate Analysis VI* (P.K. Krishnaiah, ed.) 411–429.

[14] Loh, W.L. (1988). Estimating covariance matrices. Ph. D. dissertation, Dept. Statist., Stanford Univ.

[15] Loh, W.L. (1991). Estimating the common mean of two multivariate normal distributions. *Ann. Statist.* **19**, 297–313.

[16] Stein, C. (1975). Estimation of a covariance matrix. Rietz Lecture, 39th Annual IMS Meeting, Atlanta. Unpublished manuscript.

[17] 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.

[18] Sugiura, N. and Kubokawa, T. (1988). *Estimating common parameters of growth curve models. Ann. Inst. Statist. Math.* **40**, 119–135.

[19] Tsukuma, H. and Konno, Y. (2002). Alternative estimators of the common regression matrix in two GMANOVA models under weighted quadratic losses. Technical Reports of Mathematical Sciences, Chiba University, Vol.18, No. 12. (See also `http://www.math.s.chiba-u.ac.jp/report/content.html`)