

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

Hisayuki Tsukuma<sup>a</sup> Yoshihiko Konno<sup>a</sup>

<sup>a</sup>*Graduate School of Science and Technology, Chiba University, 1-33 Yayoi-cho, Chiba-shi 263-8522, 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. Under the normality assumption, we extensively use the techniques of Haff, Stein, and Loh to derive an unbiased estimate of risk function for a subclass of equivariant estimators, from which we give alternative combined estimators to the Graybill-Deal type estimator. We also show that some of the results obtained under the normality assumption remain robust when the error matrices follow the elliptically contoured distributions. Finally, we conduct the Monte-Carlo simulation to show that our proposed estimators perform better than 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

---

## 1 Introduction

There has been a lot of literature on estimating the common mean of normal distributions, which includes Graybill and Deal (1959), Brown and Cohen (1974), Khatri and Shah (1974), and Loh (1991). Of these, Graybill and Deal (1959) first showed that the Graybill and Deal estimator, a combined estimator for the common mean of two univariate normal distributions, has

---

*Email address:* [konno@math.s.chiba-u.ac.jp](mailto:konno@math.s.chiba-u.ac.jp) (Yoshihiko Konno).

*URL:* <http://www.math.s.chiba-u.ac.jp/~konno/> (Yoshihiko Konno).

smaller variance than either of each sample mean when the sample size is at least eleven.

The paper is mainly concerned with estimating the common regression matrix of two GMANOVA models with different covariance matrices. Sugiura and Kubokawa (1988) first considered this problem and proposed the Graybill-Deal type estimator of the common regression matrix of two GMANOVA models. Our purpose of the present paper is to propose an alternative estimator which performs better than the estimator of Sugiura and Kubokawa in a decision-theoretic point of view. The precise formulation of this problem is as follows.

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

$$\mathbf{Y}_1 = \mathbf{A}_{11}\Xi\mathbf{A}_{12} + \boldsymbol{\epsilon}_1 \quad \text{and} \quad \mathbf{Y}_2 = \mathbf{A}_{21}\Xi\mathbf{A}_{22} + \boldsymbol{\epsilon}_2, \quad (1)$$

where  $\mathbf{A}_{i1}$  and  $\mathbf{A}_{i2}$  are, respectively,  $N_i \times m$  and  $q \times p_i$  known full-rank matrices with  $N_i > m$  and  $p_i \geq q$ ,  $\Xi$  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 two cases of error distributions: (i) The error matrices  $\boldsymbol{\epsilon}_1$  and  $\boldsymbol{\epsilon}_2$  are independently distributed as the multivariate normal distributions with the covariance matrices  $\mathbf{I}_{N_1} \otimes \boldsymbol{\Omega}_1$  and  $\mathbf{I}_{N_2} \otimes \boldsymbol{\Omega}_2$ , respectively, i.e., the rows of the matrix  $\boldsymbol{\epsilon}_i$  are independently and identically distributed as the multivariate normal distributions with the mean zero and the covariance matrix  $\boldsymbol{\Omega}_i$ . (ii) The error matrices  $\boldsymbol{\epsilon}_1$  and  $\boldsymbol{\epsilon}_2$  are jointly distributed as the elliptically contoured distribution with the density function

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

where  $g$  is a nonnegative unknown function and  $\boldsymbol{\Omega}_i$ ,  $i = 1, 2$ , are  $p_i \times p_i$  scale matrices. In both cases (i) and (ii), we assume that  $\boldsymbol{\Omega}_i$  are unknown positive definite  $p_i \times p_i$  matrices. Here we denote by  $\mathbf{B}'$ ,  $|\mathbf{B}|$ , and  $\text{tr}(\mathbf{B})$  the transpose, determinant, and trace of a squared matrix  $\mathbf{B}$ . We consider the problem of estimating  $\Xi$  under the loss function

$$\begin{aligned} \tilde{\mathbf{L}}((\Xi, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2), \hat{\Xi}) = & \text{tr} \{ \mathbf{A}_{11}(\hat{\Xi} - \Xi) \mathbf{A}_{12} \boldsymbol{\Omega}_1^{-1} \mathbf{A}'_{12} (\hat{\Xi} - \Xi)' \mathbf{A}'_{11} \} \\ & + \text{tr} \{ \tilde{\mathbf{C}}(\hat{\Xi} - \Xi) \mathbf{A}_{22} \boldsymbol{\Omega}_2^{-1} \mathbf{A}'_{22} (\hat{\Xi} - \Xi)' \tilde{\mathbf{C}}' \}, \end{aligned} \quad (3)$$

where  $\hat{\Xi}$  is an estimator of  $\Xi$  and  $\tilde{\mathbf{C}}$  is an  $N_2 \times m$  known matrix of full rank. When  $\tilde{\mathbf{C}} = \mathbf{A}_{21}$ , 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  $\hat{\Xi}$  is measured by the risk function  $\mathbb{E}[\tilde{\mathbf{L}}((\Xi, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2), \hat{\Xi})]$ . 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  $\text{tr}\{(\hat{\Xi} - \Xi)\mathbf{Q}(\hat{\Xi} - \Xi)'\}$  for a  $q \times q$  known positive definite matrix  $\mathbf{Q}$ .

In Section 2, we consider the estimation problem of the common regression matrix of the model (1) where the distributions of two error matrices  $\epsilon_1$  and  $\epsilon_2$  are distributed independently as the multivariate normal distributions. First we derive a canonical form of two sample problem of estimating the common regression matrix of the GMANOVA models. Next we derive a family of fully equivariant estimators for this problem. Using the methods of Stein-Haff-Loh, we obtain an unbiased estimate of the risk for a subclass of equivariant estimators. In the view of the unbiased estimate of the risk, we give an alternative estimator to the Graybill-Deal type estimator. In Section 3, we consider the estimation problem of the common regression matrix of the model (1) where the distributions of two error matrices  $\epsilon_1$  and  $\epsilon_2$  are distributed jointly and uncorrelatedly as the elliptically contoured distribution with the density function (2). We also derive a canonical form of the problems when the error distributions are elliptically contoured. Using the extended Haff-Stein identity due to Kubokawa and Srivastava (1999, 2001), we derive the risk representation for the subclass of equivariant estimators, which is an extension of the results obtained under the normal assumption to the results under the elliptically contoured model. Since complex nature of the risk representation under elliptically contoured distributions, we restrict ourselves to the problem of estimating the common mean of the elliptically contoured distributions, i.e., the case when  $N_1 = N_2$ ,  $p_1 = p_2$  in (1) and derive an alternative estimator from the our risk representation. In Section 4, we first carry out Monte-Carlo simulation to show that our proposed estimators reduce the risk substantially over the Graybill-Deal type estimator when we observe the data  $(\mathbf{Y}_1, \mathbf{Y}_2)$  from the model (1). Next we carry out simulation related to the results in Section 3. 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 Under normal errors

### 2.1 A canonical form

Assume that the errors  $\boldsymbol{\epsilon}_1$  and  $\boldsymbol{\epsilon}_2$  are independently and identically distributed as matrix-variate normal distributions. Hence we observe random matrices  $\mathbf{Y}_1$  and  $\mathbf{Y}_2$  which are independently distributed as

$$\mathbf{Y}_i \sim N_{N_i \times p_i}(\mathbf{A}_{i1} \boldsymbol{\Xi} \mathbf{A}_{i2}, \mathbf{I}_{N_i} \otimes \boldsymbol{\Omega}_i), \quad i = 1, 2. \quad (4)$$

To derive a canonical form of (4), let  $\boldsymbol{\Gamma}_i$  be  $N_i \times N_i$  orthogonal matrices such that  $\boldsymbol{\Gamma}_i \mathbf{A}_{i1} = [(\mathbf{A}'_{i1} \mathbf{A}_{i1})^{1/2}, \mathbf{0}_{m \times (N_i - m)}]'$  and also let  $\boldsymbol{\Upsilon}_i$  be  $p_i \times p_i$  orthogonal matrices such that  $\mathbf{A}_{i2} \boldsymbol{\Upsilon}_i = [(\mathbf{A}_{i2} \mathbf{A}'_{i2})^{1/2}, \mathbf{0}_{q \times (p_i - q)}]$ . Here we denote by  $\mathbf{B}^{1/2}$  a non-negative definite square root of a squared matrix  $\mathbf{B}$ . Furthermore we write

$$\boldsymbol{\Theta} = (\mathbf{A}'_{11} \mathbf{A}_{11})^{1/2} \boldsymbol{\Xi} (\mathbf{A}_{12} \mathbf{A}'_{12})^{1/2}, \quad (5a)$$

$$\mathbf{A} = (\mathbf{A}'_{21} \mathbf{A}_{21})^{1/2} (\mathbf{A}'_{11} \mathbf{A}_{11})^{-1/2}, \quad (5b)$$

$$\boldsymbol{\Lambda} = \begin{pmatrix} (\mathbf{A}_{22} \mathbf{A}'_{22})^{-1/2} (\mathbf{A}_{12} \mathbf{A}'_{12})^{1/2} & \mathbf{0}_{q \times (p_2 - q)} \\ \mathbf{0}_{(p_2 - q) \times q} & \mathbf{I}_{p_2 - q} \end{pmatrix}, \quad (5c)$$

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Upsilon}'_1 \boldsymbol{\Omega}_1 \boldsymbol{\Upsilon}_1 = \begin{pmatrix} \boldsymbol{\Sigma}_{11}^{(1)} & \boldsymbol{\Sigma}_{12}^{(1)} \\ \boldsymbol{\Sigma}_{21}^{(1)} & \boldsymbol{\Sigma}_{22}^{(1)} \end{pmatrix}, \quad (5d)$$

$$\boldsymbol{\Sigma}_2 = \boldsymbol{\Lambda}' \boldsymbol{\Upsilon}'_2 \boldsymbol{\Omega}_2 \boldsymbol{\Upsilon}_2 \boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{\Sigma}_{11}^{(2)} & \boldsymbol{\Sigma}_{12}^{(2)} \\ \boldsymbol{\Sigma}_{21}^{(2)} & \boldsymbol{\Sigma}_{22}^{(2)} \end{pmatrix}, \quad (5e)$$

where  $\boldsymbol{\Sigma}_{11}^{(i)}$ ,  $i = 1, 2$ , are  $q \times q$  positive definite matrices. Then the transformations of both  $\mathbf{Y}_1 \rightarrow \boldsymbol{\Gamma}_1 \mathbf{Y}_1 \boldsymbol{\Upsilon}_1$  and  $\mathbf{Y}_2 \rightarrow \boldsymbol{\Gamma}_2 \mathbf{Y}_2 \boldsymbol{\Upsilon}_2 \boldsymbol{\Lambda}$  yield the following form: We observe that each  $\mathbf{Y}_i$ ,  $i = 1, 2$ , yields a set of random matrices  $(\mathbf{X}_i, \mathbf{Z}_i, \mathbf{S}_i, \hat{\boldsymbol{\gamma}}_i, \mathbf{W}_i)$ , where

$$\mathbf{X}_1 | \mathbf{Z}_1 \sim N_{m \times q}(\boldsymbol{\Theta} + \mathbf{Z}_1 \boldsymbol{\gamma}_1, \mathbf{I}_m \otimes \boldsymbol{\Sigma}_{11.2}^{(1)}), \quad (6a)$$

$$\mathbf{X}_2 | \mathbf{Z}_2 \sim N_{m \times q}(\mathbf{A} \boldsymbol{\Theta} + \mathbf{Z}_2 \boldsymbol{\gamma}_2, \mathbf{I}_m \otimes \boldsymbol{\Sigma}_{11.2}^{(2)}) \quad (6b)$$

and, for  $i = 1, 2$ ,

$$\mathbf{Z}_i \sim N_{m \times (p_i - q)}(\mathbf{0}, \mathbf{I}_m \otimes \Sigma_{22}^{(i)}), \quad (7a)$$

$$\mathbf{S}_i \sim W_q(\Sigma_{11.2}^{(i)}, n_i), \quad n_i = N_i - m - p_i + q, \quad (7b)$$

$$\hat{\gamma}_i | \mathbf{W}_i \sim N_{(p_i - q) \times q}(\gamma_i, \mathbf{W}_i^{-1} \otimes \Sigma_{11.2}^{(i)}), \quad (7c)$$

$$\mathbf{W}_i \sim W_{p_i - q}(\Sigma_{22}^{(i)}, n_i + p_i - q), \quad (7d)$$

where  $\Sigma_{11.2}^{(i)} = \Sigma_{11}^{(i)} - \Sigma_{12}^{(i)}(\Sigma_{22}^{(i)})^{-1}\Sigma_{21}^{(i)}$  and  $\gamma_i = (\Sigma_{22}^{(i)})^{-1}\Sigma_{21}^{(i)}$ . Here, note that  $(\mathbf{X}_i, \mathbf{Z}_i)$ ,  $(\mathbf{W}_i, \hat{\gamma}_i)$  and  $\mathbf{S}_i$  are independent and that  $\mathbf{A}$  is an  $m \times m$  known nonsingular matrix. Furthermore, the loss function (3) turns into

$$\begin{aligned} L((\Theta, \Sigma_1, \Sigma_2), \widehat{\Theta}) &= \text{tr}[(\widehat{\Theta} - \Theta)(\Sigma_{11.2}^{(1)})^{-1}(\widehat{\Theta} - \Theta)'] \\ &\quad + \text{tr}[\mathbf{C}'\mathbf{C}(\widehat{\Theta} - \Theta)(\Sigma_{11.2}^{(2)})^{-1}(\widehat{\Theta} - \Theta)'], \end{aligned} \quad (8)$$

where  $\widehat{\Theta}$  is an estimator of  $\Theta$  and  $\mathbf{C}$  is an  $N_2 \times m$  known matrix of full rank. Under this canonical form, the problem of estimating  $\Xi$  in (1) changes into that of estimating  $\Theta$  based on  $(\mathbf{X}_i, \mathbf{Z}_i, \mathbf{S}_i, \hat{\gamma}_i, \mathbf{W}_i | i = 1, 2)$  under the loss function (8). Then the risk function is defined by

$$\mathbf{R}((\Theta, \Sigma_1, \Sigma_2), \widehat{\Theta}) = \mathbb{E}[L((\Theta, \Sigma_1, \Sigma_2), \widehat{\Theta})], \quad (9)$$

where the expectation is taken with respect to  $(\mathbf{X}_i, \mathbf{Z}_i, \mathbf{S}_i, \hat{\gamma}_i, \mathbf{W}_i | i = 1, 2)$ .

## 2.2 An equivariant estimator of $\Theta$

Next, we derive a class of estimators of  $\Theta$ . To this end, let  $G$  be a group of transformations on the sample space. Each element of  $G$  consists of triples  $(\mathbf{D}, \mathbf{P}_1, \mathbf{P}_2)$ , where  $\mathbf{D}$  is  $m \times q$  matrix and

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

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

$$\begin{aligned}
[\mathbf{X}_1, \mathbf{Z}_1] &\rightarrow [\mathbf{X}_1, \mathbf{Z}_1]\mathbf{P}'_1 + [\mathbf{D}, \mathbf{0}_{m \times (p_1 - q)}], \\
[\mathbf{X}_2, \mathbf{Z}_2] &\rightarrow [\mathbf{X}_2, \mathbf{Z}_2]\mathbf{P}'_2 + [\mathbf{AD}, \mathbf{0}_{m \times (p_2 - q)}], \\
\begin{pmatrix} \mathbf{S}_i + \hat{\gamma}'_i \mathbf{W}_i \hat{\gamma}_i & \hat{\gamma}'_i \mathbf{W}_i \\ \mathbf{W}_i \hat{\gamma}_i & \mathbf{W}_i \end{pmatrix} &\rightarrow \mathbf{P}_i \begin{pmatrix} \mathbf{S}_i + \hat{\gamma}'_i \mathbf{W}_i \hat{\gamma}_i & \hat{\gamma}'_i \mathbf{W}_i \\ \mathbf{W}_i \hat{\gamma}_i & \mathbf{W}_i \end{pmatrix} \mathbf{P}'_i,
\end{aligned}$$

and we denote by  $g \circ (\mathbf{X}_i, \mathbf{Z}_i, \mathbf{S}_i, \hat{\gamma}_i, \mathbf{W}_i | i = 1, 2)$  the action of  $g$  on this sample where  $g$  is an element of  $G$ , i.e.,  $g = (\mathbf{D}, \mathbf{P}_1, \mathbf{P}_2)$ . Furthermore, the action of  $g$  on the parameter is defined as  $\Theta \rightarrow \Theta \mathbf{P}'_{11} + \mathbf{D}$ , and  $\Sigma^{(i)} \rightarrow \mathbf{P}_i \Sigma^{(i)} \mathbf{P}'_i$ ,  $i = 1, 2$ . Then the model is easily shown to be invariant under the group of transformations. Furthermore, let

$$\widehat{\Theta}_i = \mathbf{X}_i - \mathbf{Z}_i \hat{\gamma}_i, \quad i = 1, 2. \quad (10)$$

Note that  $\widehat{\Theta}_1$  and  $\widehat{\Theta}_2$  are the maximum likelihood estimators of  $\Theta$  and  $\mathbf{A}\Theta$  for one-sample problem, respectively. Then the action of  $g$  on sample and parameters is rewritten as

$$\begin{aligned}
&\Theta \rightarrow \Theta \mathbf{P}'_{11} + \mathbf{D} \\
&(\Sigma_{11.2}^{(i)}, \Sigma_{22}^{(i)}, (\Sigma_{22}^{(i)})^{-1} \Sigma_{21}^{(i)}) \\
&\rightarrow (\mathbf{P}_{11} \Sigma_{11.2}^{(i)} \mathbf{P}'_{11}, \mathbf{P}_{i.22} \Sigma_{22}^{(i)} \mathbf{P}'_{i.22}, (\mathbf{P}'_{i.22})^{-1} (\Sigma_{22}^{(i)})^{-1} \Sigma_{21}^{(i)} \mathbf{P}'_{11} + (\mathbf{P}'_{i.22})^{-1} \mathbf{P}'_{i.12}), \\
&(\widehat{\Theta}_1, \mathbf{Z}_1, \widehat{\Theta}_2, \mathbf{Z}_2) \\
&\rightarrow (\widehat{\Theta}_1 \mathbf{P}'_{11} + \mathbf{D}, \mathbf{Z}_1 \mathbf{P}'_{1.22}, \widehat{\Theta}_2 \mathbf{P}'_{11} + \mathbf{AD}, \mathbf{Z}_2 \mathbf{P}'_{2.22}), \\
&(\mathbf{S}_i, \mathbf{W}_i, \hat{\gamma}_i) \\
&\rightarrow (\mathbf{P}_{11} \mathbf{S}_i \mathbf{P}'_{11}, \mathbf{P}_{i.22} \mathbf{W}_i \mathbf{P}'_{i.22}, (\mathbf{P}'_{i.22})^{-1} \hat{\gamma}_i \mathbf{P}'_{11} + (\mathbf{P}'_{i.22})^{-1} \mathbf{P}'_{i.12}),
\end{aligned}$$

for  $i = 1, 2$ . It is reasonable to require that an equivariant estimator  $\widehat{\Theta}^{EQI}$  should satisfy

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

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

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

$$\widehat{\Theta}^{EQI} = \widehat{\Theta}_1 \mathbf{B}' \tilde{\Phi} (\mathbf{B}')^{-1} + \mathbf{A}^{-1} \widehat{\Theta}_2 \mathbf{B}' (\mathbf{I}_q - \tilde{\Phi}) (\mathbf{B}')^{-1}, \quad (11)$$

where  $\tilde{\Phi} \equiv \tilde{\Phi}((\widehat{\Theta}_1 - \mathbf{A}^{-1}\widehat{\Theta}_2)\mathbf{B}', \mathbf{F}, \mathbf{Z}_1\mathbf{W}_1^{-1/2}, \mathbf{Z}_2\mathbf{W}_2^{-1/2})$  is a  $q \times q$  matrix and  $\widehat{\Theta}_i, i = 1, 2$ , are given by (10).

Since the class of the equivariant estimators (11) is too large to evaluate their risk systematically, we restrict ourselves to an equivariant estimator (11) where  $\tilde{\Phi}$  is a diagonal matrix and depends only on  $\mathbf{F}$ , i.e.,

$$\widehat{\Theta}^{EQ} = \widehat{\Theta}_1\mathbf{B}'\Phi(\mathbf{B}')^{-1} + \mathbf{A}^{-1}\widehat{\Theta}_2\mathbf{B}'(\mathbf{I}_q - \Phi)(\mathbf{B}')^{-1}, \quad (12)$$

where  $\widehat{\Theta}_i, i = 1, 2$ , is given by (10) and  $\Phi = \Phi(\mathbf{F})$  is a  $q \times q$  diagonal matrix with diagonal elements  $\phi_i(\mathbf{F}), i = 1, 2, \dots, p$ . Here we assume that  $\phi_i(\mathbf{F})$  depends only on  $\mathbf{F} = \mathbf{diag}(f_1, f_2, \dots, f_q)$  with  $f_1 \geq f_2 \geq \dots \geq f_q$ , the eigenvalues of  $\mathbf{S}_2(\mathbf{S}_1 + \mathbf{S}_2)^{-1}$ .

### 2.3 Graybill-Deal type estimator

In this subsection, we look over the connection between our proposed class of estimators and the Graybill-Deal type estimator given by Sugiura and Kubokawa (1988). Furthermore, we state our scenario to obtain an alternative estimator. Using the transformation (5a) – (5e), we can see that the estimator of Sugiura and Kubokawa is rewritten as

$$\begin{aligned} \text{vec}(\widehat{\Theta}^{SK}) &= \{\mathbf{I}_m \otimes (\mathbf{S}_1/n_1)^{-1} + (\mathbf{A}'\mathbf{A}) \otimes (\mathbf{S}_2/n_2)^{-1}\}^{-1} \\ &\times \{\mathbf{I}_m \otimes (\mathbf{S}_1/n_1)^{-1}\text{vec}(\widehat{\Theta}_1) + (\mathbf{A}'\mathbf{A}) \otimes (\mathbf{S}_2/n_2)^{-1}\text{vec}(\mathbf{A}^{-1}\widehat{\Theta}_2)\}, \end{aligned} \quad (13)$$

where we denote by  $\text{vec}(\mathbf{U})$  an  $mq \times 1$  vector consisting of  $(u_1, u_2, \dots, u_m)'$  for  $\mathbf{U} = (u'_1, u'_2, \dots, u'_m)'$  and  $\mathbf{G} \otimes \mathbf{H}$  stands for the Kronecker product of matrices  $\mathbf{G}$  and  $\mathbf{H}$  defined by  $(g_{ij}\mathbf{H})$  for  $\mathbf{G} = (g_{ij})$ . On the other hand, we can rewritten the estimator (12) as

$$\begin{aligned} \text{vec}(\widehat{\Theta}^{EQ}) &= \{\mathbf{I}_m \otimes (\mathbf{B}'\mathbf{diag}(\beta_j)\mathbf{B}) + \mathbf{I}_m \otimes (\mathbf{B}'\mathbf{diag}(\alpha_j)\mathbf{B})\}^{-1} \\ &\times \{\mathbf{I}_m \otimes (\mathbf{B}'\mathbf{diag}(\beta_j)\mathbf{B}) \text{vec}(\widehat{\Theta}_1) \\ &+ \mathbf{I}_m \otimes (\mathbf{B}'\mathbf{diag}(\alpha_j)\mathbf{B})\text{vec}(\mathbf{A}^{-1}\widehat{\Theta}_2)\}, \end{aligned} \quad (14)$$

if we put  $\phi_j = \beta_j/(\alpha_j + \beta_j), j = 1, 2, \dots, q$ , where  $\alpha_j$  and  $\beta_j$  are real-valued functions of  $\mathbf{F}$ . Here we denote by  $\mathbf{diag}(\beta_j)$  a  $q \times q$  diagonal matrix whose  $j$ -th diagonal elements are given by  $\beta_j$ . Furthermore, putting  $\alpha_j = n_2/f_j$  and  $\beta_j = n_1/(1 - f_j)$ , we can see that the equivariant estimator of the form (12) reduces to

$$\begin{aligned} \text{vec}(\widehat{\Theta}^{EQ1}) &= \{\mathbf{I}_m \otimes (\mathbf{S}_1/n_1)^{-1} + \mathbf{I}_m \otimes (\mathbf{S}_2/n_2)^{-1}\}^{-1} \\ &\quad \times \{\mathbf{I}_m \otimes (\mathbf{S}_1/n_1)^{-1} \text{vec}(\widehat{\Theta}_1) + \mathbf{I}_m \otimes (\mathbf{S}_2/n_2)^{-1} \text{vec}(\mathbf{A}^{-1} \widehat{\Theta}_2)\}, \end{aligned} \quad (15)$$

equivalently

$$\widehat{\Theta}^{EQ1} = \{\widehat{\Theta}_1(\mathbf{S}_1/n_1)^{-1} + \mathbf{A}^{-1} \widehat{\Theta}_2(\mathbf{S}_2/n_2)^{-1}\} \left\{ \sum_{i=1}^2 (\mathbf{S}_i/n_i)^{-1} \right\}^{-1}.$$

The estimator (15) can be regarded as a counterpart of the Graybill-Deal type estimator (13) inside the class of equivariant estimators of the form (12). It is well known that the eigenvalues of  $\mathbf{S}_2(\mathbf{S}_1 + \mathbf{S}_2)^{-1}$  are more spread than the eigenvalues of expected value of  $\mathbf{S}_2(\mathbf{S}_1 + \mathbf{S}_2)^{-1}$ . Hence we look for alternative estimators for  $\Theta$  by correcting the eigenvalues of  $\mathbf{S}_2(\mathbf{S}_1 + \mathbf{S}_2)^{-1}$ . Through these consideration, we use the following scenario to obtain an alternative estimator to (13). First we look into the class of equivariant estimators of the form (12) and obtain better estimator which has the form (12). Then we change the term  $\mathbf{I}_m \otimes (\mathbf{B}' \text{diag}(\alpha_j) \mathbf{B})$  in (14) into  $(\mathbf{A}' \mathbf{A}) \otimes (\mathbf{B}' \text{diag}(\alpha_j) \mathbf{B})$  to get an alternative estimator so as to (13).

#### 2.4 A subclass of the equivariant estimator and its risk

To obtain alternative estimator of the form (12), we evaluate its risk in terms of unbiased risk method due to Stein-Haff-Loh. The risk of the above estimator can be written as

$$\begin{aligned} &R((\Theta, \Sigma_1, \Sigma_2), \widehat{\Theta}) \\ &= \mathbb{E} \left[ \text{tr} \{ (\widehat{\Theta}_1 - \Theta)(\Sigma_{11.2}^{(1)})^{-1} (\widehat{\Theta}_1 - \Theta)' \} \right. \\ &\quad + 2 \text{tr} \{ (\widehat{\Theta}_1 - \Theta)(\Sigma_{11.2}^{(1)})^{-1} \mathbf{B}^{-1} (\mathbf{I}_q - \Phi) \mathbf{B} (\mathbf{A}^{-1} \widehat{\Theta}_2 - \widehat{\Theta}_1)' \} \\ &\quad + \text{tr} \{ (\Sigma_{11.2}^{(1)})^{-1} \mathbf{B}^{-1} (\mathbf{I}_q - \Phi) \mathbf{H}_1 (\mathbf{I}_q - \Phi) (\mathbf{B}')^{-1} \} \\ &\quad + \text{tr} \{ (\mathbf{C} \mathbf{A}^{-1})' (\mathbf{C} \mathbf{A}^{-1}) (\widehat{\Theta}_2 - \mathbf{A} \Theta) (\Sigma_{11.2}^{(2)})^{-1} (\widehat{\Theta}_2 - \mathbf{A} \Theta)' \} \\ &\quad + 2 \text{tr} \{ (\mathbf{C} \mathbf{A}^{-1})' (\mathbf{C} \mathbf{A}^{-1}) (\widehat{\Theta}_2 - \mathbf{A} \Theta) (\Sigma_{11.2}^{(2)})^{-1} \mathbf{B}^{-1} \Phi \mathbf{B} (\mathbf{A} \widehat{\Theta}_1 - \widehat{\Theta}_2)' \} \\ &\quad \left. + \text{tr} \{ (\Sigma_{11.2}^{(2)})^{-1} \mathbf{B}^{-1} \Phi \mathbf{H}_2 \Phi (\mathbf{B}')^{-1} \} \right], \end{aligned} \quad (16)$$

where

$$\mathbf{H}_1 = \mathbf{B} (\widehat{\Theta}_1 - \mathbf{A}^{-1} \widehat{\Theta}_2)' (\widehat{\Theta}_1 - \mathbf{A}^{-1} \widehat{\Theta}_2) \mathbf{B}', \quad (17a)$$

$$\mathbf{H}_2 = \mathbf{B} (\widehat{\Theta}_1 - \mathbf{A}^{-1} \widehat{\Theta}_2)' (\mathbf{C}' \mathbf{C}) (\widehat{\Theta}_1 - \mathbf{A}^{-1} \widehat{\Theta}_2) \mathbf{B}'. \quad (17b)$$

Now we use the Haff-Stein identity for Wishart distribution and calculation on eigenstructure technique due to Stein (1975, 1977), Haff (1991), and Loh (1988) to evaluate the third and sixth terms in right-hand side of (16) while we use formula for the second moments of the maximum likelihood estimator of the GMANOVA model to evaluate the other terms in right-hand side of (16). Then we obtain an unbiased estimate of risk for the equivariant estimators (12). The proof of the theorem is given in Section 5.

**Theorem 2.** *The risk of  $\widehat{\Theta}^{EQ}$  is given by*

$$\begin{aligned}
& \mathbf{R}((\Theta, \Sigma_1, \Sigma_2), \widehat{\Theta}^{EQ}) \\
&= \mathbb{E} \left[ q(r_2 - r_1) + \sum_{j=1}^q \left\{ 2(r_1 - r_2)\phi_j + (n_1 - q - 1) \frac{(1 - \phi_j)^2}{1 - f_j} \{\mathbf{H}_1\}_{jj} \right. \right. \\
&+ 4\{\mathbf{H}_1\}_{jj}(1 - \phi_j)f_j \frac{\partial \phi_j}{\partial f_j} + 2 \sum_{k \neq j} \{\mathbf{H}_1\}_{jj}(1 - \phi_j)(\phi_j - \phi_k) \frac{f_k}{f_j - f_k} \\
&+ (n_2 - q - 1) \frac{\phi_j^2}{f_j} \{\mathbf{H}_2\}_{jj} + 4\{\mathbf{H}_2\}_{jj}\phi_j(1 - f_j) \frac{\partial \phi_j}{\partial f_j} \\
&\left. \left. + 2 \sum_{k \neq j} \{\mathbf{H}_2\}_{jj}\phi_j(\phi_j - \phi_k) \frac{1 - f_k}{f_j - f_k} \right\} \right], \tag{18}
\end{aligned}$$

where  $r_1 = m(n_1 + p_1 - q - 1)/(n_1 - 1)$ ,  $r_2 = \{(n_2 + p_2 - q - 1)/(n_2 - 1)\} \text{tr} \{(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}$ , and  $\{\mathbf{H}_1\}_{jj}$  and  $\{\mathbf{H}_2\}_{jj}$  are  $j$ -th diagonal elements of the matrices given by (17a) and (17b), respectively.

## 2.5 Choice of $\Phi$

From Theorem 2, we obtain the unbiased estimate of the risk of the subclass of equivariant estimators given by (12). We denote by  $\widehat{\mathbf{R}}$  the unbiased estimate of the risk, i.e., the terms inside large bracket in the right-hand side of (18). Although we obtain the unbiased estimate of risk for the class of estimators given by (12), 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 estimate of the risk from the unbiased estimate of the risk. First we replace  $\mathbf{H}_1$  and  $\mathbf{H}_2$  in (18) by their approximation. To this end, we observe that

$$\begin{aligned}
& \mathbb{E}[(\widehat{\Theta}_1 - \mathbf{A}^{-1}\widehat{\Theta}_2)'(\widehat{\Theta}_1 - \mathbf{A}^{-1}\widehat{\Theta}_2)] \\
& \quad = m\tilde{r}_1 \Sigma_{11.2}^{(1)} + \tilde{r}_2 \operatorname{tr} \{(\mathbf{A}')^{-1} \mathbf{A}^{-1}\} \Sigma_{11.2}^{(2)}, \\
& \mathbb{E}[(\mathbf{A}\widehat{\Theta}_1 - \widehat{\Theta}_2)'(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})(\mathbf{A}\widehat{\Theta}_1 - \widehat{\Theta}_2)] \\
& \quad = \tilde{r}_1 \operatorname{tr}(\mathbf{C}'\mathbf{C}) \Sigma_{11.2}^{(1)} + \tilde{r}_2 \operatorname{tr} \{(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\} \Sigma_{11.2}^{(2)},
\end{aligned}$$

where  $\tilde{r}_i = (n_i + p_i - q - 1)/(n_i - 1)$ . Replacing  $\Sigma_{11.2}^{(i)}$  in right-hand side of the above equations with their maximum likelihood estimators  $\mathbf{S}_i/n_i$ ,  $i = 1, 2$ , we approximate  $\{\mathbf{H}_1\}_{jj}$  and  $\{\mathbf{H}_2\}_{jj}$ ,  $j = 1, 2, \dots, q$ , by

$$\begin{aligned}
\{\mathbf{H}_1\}_{jj} & \approx \{\mathbf{B}(m\tilde{r}_1 \mathbf{S}_1/n_1 + \tilde{r}_2 \operatorname{tr} \{(\mathbf{A}')^{-1} \mathbf{A}^{-1}\} \mathbf{S}_2/n_2) \mathbf{B}'\}_{jj}, \\
& = m\tilde{r}_1(1 - f_j)/n_1 + \tilde{r}_2 \operatorname{tr} \{(\mathbf{A}')^{-1} \mathbf{A}^{-1}\} f_j/n_2 \\
& \equiv h_{1j}, \\
\{\mathbf{H}_2\}_{jj} & \approx \{\mathbf{B}(\tilde{r}_1 \operatorname{tr}(\mathbf{C}'\mathbf{C}) \mathbf{S}_1/n_1 + \tilde{r}_2 \operatorname{tr} \{(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\} \mathbf{S}_2/n_2) \mathbf{B}'\}_{jj} \\
& = \tilde{r}_1 \operatorname{tr}(\mathbf{C}'\mathbf{C})(1 - f_j)/n_1 + \tilde{r}_2 \operatorname{tr} \{(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\} f_j/n_2 \\
& \equiv h_{2j}.
\end{aligned}$$

We extensively use notation  $\{\mathbf{A}\}_{jj}$ ,  $j = 1, 2, \dots, q$ , to denote the  $j$ -th diagonal element of a  $q \times q$  squared matrix  $\mathbf{A}$ . Furthermore, using the fact that

$$\frac{\partial \phi_j}{\partial f_j} = f_j \frac{\partial}{\partial f_j} \left( \frac{\phi_j}{f_j} \right) + \frac{\phi_j}{f_j} = (1 - f_j) \frac{\partial}{\partial (1 - f_j)} \left( \frac{1 - \phi_j}{1 - f_j} \right) + \frac{1 - \phi_j}{1 - f_j}, \quad (19)$$

we can see that the unbiased estimate for risk of  $\widehat{\Theta}^{EQ}$  given by (18) is approximated by

$$\begin{aligned}
\widehat{\mathbf{R}} & \approx q(r_2 - r_1) + \sum_{j=1}^q \left\{ 2(r_1 - r_2)\phi_j + (n_1 - q - 1) \frac{(1 - \phi_j)^2}{1 - f_j} h_{1j} \right. \\
& \quad + 4h_{1j}(1 - \phi_j)f_j \left[ (1 - f_j) \frac{\partial}{\partial (1 - f_j)} \left( \frac{1 - \phi_j}{1 - f_j} \right) + \frac{1 - \phi_j}{1 - f_j} \right] \\
& \quad + 2 \sum_{k \neq j} h_{1j}(1 - \phi_j)(\phi_j - \phi_k) \frac{f_k}{f_j - f_k} \\
& \quad + (n_2 - q - 1) \frac{\phi_j^2}{f_j} h_{2j} + 4h_{2j}\phi_j(1 - f_j) \left[ f_j \frac{\partial}{\partial f_j} \left( \frac{\phi_j}{f_j} \right) + \frac{\phi_j}{f_j} \right] \\
& \quad \left. + 2 \sum_{k \neq j} h_{2j}\phi_j(\phi_j - \phi_k) \frac{1 - f_k}{f_j - f_k} \right\}.
\end{aligned}$$

Ignoring the derivative terms, we get

$$\begin{aligned}
\widehat{\mathbf{R}} &\approx q(r_2 - r_1) + \sum_{j=1}^q \left\{ 2(r_1 - r_2)\phi_j \right. \\
&\quad + (n_1 - q - 1) \frac{(1 - \phi_j)^2}{1 - f_j} h_{1j} + 4h_{1j}(1 - \phi_j)^2 \frac{f_j}{1 - f_j} \\
&\quad + 2 \sum_{k \neq j} h_{1j}(1 - \phi_j)(\phi_j - \phi_k) \frac{f_k}{f_j - f_k} \\
&\quad \left. + (n_2 - q - 1) \frac{\phi_j^2}{f_j} h_{2j} + 4h_{2j}\phi_j^2 \frac{1 - f_j}{f_j} + 2 \sum_{k \neq j} h_{2j}\phi_j(\phi_j - \phi_k) \frac{1 - f_k}{f_j - f_k} \right\} \\
&= q(r_2 - r_1) + \sum_{j=1}^q \left\{ 2(r_1 - r_2)\phi_j \right. \\
&\quad + (n_1 - q - 1) \frac{(1 - \phi_j)^2}{1 - f_j} h_{1j} + 4h_{1j}(1 - \phi_j)^2 \frac{f_j}{1 - f_j} \\
&\quad - 2 \sum_{k \neq j} h_{1j}(1 - \phi_j)^2 \frac{f_k}{f_j - f_k} + 2 \sum_{k \neq j} h_{1j}(1 - \phi_j)(1 - \phi_k) \frac{f_k}{f_j - f_k} \\
&\quad \left. + (n_2 - q - 1) \frac{\phi_j^2}{f_j} h_{2j} + 4h_{2j}\phi_j^2 \frac{1 - f_j}{f_j} + 2 \sum_{k \neq j} h_{2j}\phi_j(\phi_j - \phi_k) \frac{1 - f_k}{f_j - f_k} \right\} \\
&= \widetilde{\mathbf{R}}, \quad \text{say.}
\end{aligned}$$

Although the estimate of the risk  $\widetilde{\mathbf{R}}$  is no longer unbiased, it is feasible to obtain alternative estimators of  $\Theta$ . Then we minimize  $\widetilde{\mathbf{R}}$  with respect to  $\phi_j$  ( $j = 1, \dots, q$ ), which gives

$$\begin{aligned}
0 = \frac{\partial \widetilde{\mathbf{R}}}{\partial \phi_j} &= r_1 - r_2 - (n_1 - q - 1) \frac{1 - \phi_j}{1 - f_j} h_{1j} - 4h_{1j}(1 - \phi_j) \frac{f_j}{1 - f_j} \\
&\quad + 2 \sum_{k \neq j} h_{1j}(1 - \phi_j) \frac{f_k}{f_j - f_k} - \sum_{k \neq j} h_{1j}(1 - \phi_k) \frac{f_k}{f_j - f_k} \\
&\quad + (n_2 - q - 1) \frac{\phi_j}{f_j} h_{2j} + 4h_{2j} \frac{1 - f_j}{f_j} \phi_j \\
&\quad + 2h_{2j}\phi_j \sum_{k \neq j} \frac{1 - f_k}{f_j - f_k} - h_{2j} \sum_{k \neq j} \phi_k \frac{1 - f_k}{f_j - f_k}.
\end{aligned}$$

Hence, solving for  $\phi_j$  with ignoring the sixth and the tenth terms in the last right-hand side above, we finally get

$$\phi_j^{ST} = \frac{\hat{\beta}_j^{ST}/(1 - f_j)}{\hat{\beta}_j^{ST}/(1 - f_j) + \hat{\alpha}_j^{ST}/f_j}, \quad (20)$$

where

$$\begin{aligned}
\hat{\alpha}_j^{ST} &= (n_2 - q - 1)h_{2j} + (r_1 - r_2)f_j + 4h_{2j}(1 - f_j) + 2h_{2j} \sum_{k \neq j} \frac{f_j(1 - f_k)}{f_j - f_k}, \\
\hat{\beta}_j^{ST} &= (n_1 - q - 1)h_{1j} + (r_2 - r_1)(1 - f_j) + 4h_{1j}f_j - 2h_{1j} \sum_{k \neq j} \frac{(1 - f_j)f_k}{f_j - f_k}, \\
h_{1j} &= m\tilde{r}_1(1 - f_j)/n_1 + \tilde{r}_2(\text{tr}(\mathbf{A}')^{-1}\mathbf{A}^{-1})f_j/n_2, \\
h_{2j} &= \tilde{r}_1(\text{tr}\mathbf{C}'\mathbf{C})(1 - f_j)/n_1 + \tilde{r}_2(\text{tr}(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1}))f_j/n_2, \\
\tilde{r}_i &= \frac{n_i + p_i - q - 1}{n_i - 1} = \frac{N_i - m - 1}{N_i - m - p_i + q - 1} \quad (i = 1, 2), \\
r_1 &= m\tilde{r}_1, \\
r_2 &= \tilde{r}_2 \text{tr}(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1}).
\end{aligned}$$

Consequently we propose an estimator of the form (14) with (20). Because of complex nature of the estimation problem, we can not carry out analytic comparison between Graybill-Deal type estimator (13) and our proposed estimator. However, we justify our proposed estimator via simulation study in Section 4.

**Remark 1.** For the special case, the estimator (20) reduces a simple form. When  $\mathbf{C}'\mathbf{C} = \mathbf{A}'\mathbf{A}$ ,  $N_1 = N_2$  and  $p_1 = p_2$ , we have  $r_1 = r_2$ . This case generalizes the results obtained by Loh (1991). When  $\mathbf{C}'\mathbf{C} = \mathbf{I}_m$ , we have  $h_{1j} = h_{2j}$ ,  $j = 1, \dots, q$ .

### 3 Under elliptical errors

Consider the GMANOVA model (1) and suppose that the error  $(\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2)$  is distributed as an elliptically contoured distribution and has the density function (2).

#### 3.1 A canonical form

To construct a canonical form of (1), let  $\boldsymbol{\Gamma}_i$  be  $N_i \times N_i$  orthogonal matrices such that  $\boldsymbol{\Gamma}_i \mathbf{A}_{i1} = [(\mathbf{A}'_{i1} \mathbf{A}_{i1})^{1/2}, \mathbf{0}_{m \times (N_i - m)}]'$  and also let  $\boldsymbol{\Upsilon}_i$  be  $p_i \times p_i$  orthogonal matrices such that  $\mathbf{A}_{i2} \boldsymbol{\Upsilon}_i = [(\mathbf{A}_{i2} \mathbf{A}'_{i2})^{1/2}, \mathbf{0}_{q \times (p_i - q)}]$  for  $i = 1, 2$ . Recall that  $\boldsymbol{\Theta}$ ,  $\mathbf{A}$ ,  $\boldsymbol{\Lambda}$ ,  $\boldsymbol{\Sigma}_1$ , and  $\boldsymbol{\Sigma}_2$  are given by (5b)-(5e), respectively. Also recall that  $\boldsymbol{\Sigma}_{11,2}^{(i)} = \boldsymbol{\Sigma}_{11}^{(i)} - \boldsymbol{\Sigma}_{12}^{(i)} \boldsymbol{\Sigma}_{22}^{(i)-1} \boldsymbol{\Sigma}_{21}^{(i)}$  and that  $\boldsymbol{\gamma}_i = \boldsymbol{\Sigma}_{22}^{(i)-1} \boldsymbol{\Sigma}_{21}^{(i)}$ . Then the transformations with  $\boldsymbol{\Gamma}_i$  and  $\boldsymbol{\Upsilon}_i$  yield the following lemma:

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

$$\begin{aligned}
& |\boldsymbol{\Sigma}_1|^{-N_1/2} |\boldsymbol{\Sigma}_2|^{-N_2/2} g \left\{ \text{tr} \left[ (\boldsymbol{\Sigma}_{11.2}^{(1)})^{-1} (\mathbf{X}_1 - \mathbf{Z}_1 \boldsymbol{\gamma}_1 - \boldsymbol{\Theta})' (\mathbf{X}_1 - \mathbf{Z}_1 \boldsymbol{\gamma}_1 - \boldsymbol{\Theta}) \right. \right. \\
& \quad \left. \left. + \mathbf{s}'_1 \mathbf{s}_1 + (\mathbf{u}_1 - (\mathbf{w}'_1 \mathbf{w}_1)^{1/2} \boldsymbol{\gamma}_1)' (\mathbf{u}_1 - (\mathbf{w}'_1 \mathbf{w}_1)^{1/2} \boldsymbol{\gamma}_1) \right] \right. \\
& \quad \left. + \text{tr} \left[ (\boldsymbol{\Sigma}_{22}^{(1)})^{-1} \{ \mathbf{Z}'_1 \mathbf{Z}_1 + \mathbf{w}'_1 \mathbf{w}_1 \} \right] + \text{tr} \left[ (\boldsymbol{\Sigma}_{11.2}^{(2)})^{-1} (\mathbf{X}_2 - \mathbf{Z}_2 \boldsymbol{\gamma}_2 - \mathbf{A} \boldsymbol{\Theta})' \right. \right. \\
& \quad \left. \left. \times (\mathbf{X}_2 - \mathbf{Z}_2 \boldsymbol{\gamma}_2 - \mathbf{A} \boldsymbol{\Theta}) + \mathbf{s}'_2 \mathbf{s}_2 + (\mathbf{u}_2 - (\mathbf{w}'_2 \mathbf{w}_2)^{1/2} \boldsymbol{\gamma}_2)' \right. \right. \\
& \quad \left. \left. \times (\mathbf{u}_2 - (\mathbf{w}'_2 \mathbf{w}_2)^{1/2} \boldsymbol{\gamma}_2) \right] + \text{tr} \left[ (\boldsymbol{\Sigma}_{22}^{(2)})^{-1} \{ \mathbf{Z}'_2 \mathbf{Z}_2 + \mathbf{w}'_2 \mathbf{w}_2 \} \right] \right\}, \quad (21)
\end{aligned}$$

where  $\mathbf{X}_i$  are  $m \times q$  matrices,  $\mathbf{Z}_i$  are  $m \times (p_i - q)$  matrices,  $\mathbf{s}_i$  are  $(N_i - m - p_i + q) \times q$  matrices,  $\mathbf{u}_i$  are  $(p_i - q) \times q$  matrices, and  $\mathbf{w}_i$  are  $(N_i - m) \times (p_i - q)$  matrices for  $i = 1, 2$ .

**Proof.** Let

$$\boldsymbol{\Gamma}_1 \mathbf{Y}_1 \boldsymbol{\Upsilon}_1 = \begin{pmatrix} \mathbf{X}_1 & \mathbf{Z}_1 \\ \mathbf{y}_1 & \mathbf{w}_1 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Gamma}_2 \mathbf{Y}_2 \boldsymbol{\Upsilon}_2 \boldsymbol{\Lambda} = \begin{pmatrix} \mathbf{X}_2 & \mathbf{Z}_2 \\ \mathbf{y}_2 & \mathbf{w}_2 \end{pmatrix}.$$

Then the Jacobian of the above transformations is given by

$$J[(\mathbf{Y}_i; i = 1, 2) \rightarrow (\mathbf{X}_i, \mathbf{Z}_i, \mathbf{y}_i, \mathbf{w}_i; i = 1, 2)] = |\boldsymbol{\Lambda}|^{-N_2}.$$

Note that

$$\boldsymbol{\Sigma}_i^{-1} = \begin{pmatrix} \mathbf{I}_q & \mathbf{0}_{q \times (p_i - q)} \\ -\boldsymbol{\gamma}_i & \mathbf{I}_{p_i - q} \end{pmatrix} \begin{pmatrix} (\boldsymbol{\Sigma}_{11.2}^{(i)})^{-1} & \mathbf{0}_{q \times (p_i - q)} \\ \mathbf{0}_{(p_i - q) \times q} & (\boldsymbol{\Sigma}_{22}^{(i)})^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I}_q & -\boldsymbol{\gamma}'_i \\ \mathbf{0}_{(p_i - q) \times q} & \mathbf{I}_{p_i - q} \end{pmatrix}.$$

Thus we can write the density (2) as

$$\begin{aligned}
& |\boldsymbol{\Sigma}_1|^{-N_1/2} |\boldsymbol{\Sigma}_2|^{-N_2/2} g \left\{ \text{tr} \left[ (\boldsymbol{\Sigma}_{11.2}^{(1)})^{-1} (\mathbf{X}_1 - \mathbf{Z}_1 \boldsymbol{\gamma}_1 - \boldsymbol{\Theta})' (\mathbf{X}_1 - \mathbf{Z}_1 \boldsymbol{\gamma}_1 - \boldsymbol{\Theta}) \right. \right. \\
& \quad \left. \left. + (\mathbf{y}_1 - \mathbf{w}_1 \boldsymbol{\gamma}_1)' (\mathbf{y}_1 - \mathbf{w}_1 \boldsymbol{\gamma}_1) \right] + \text{tr} \left[ (\boldsymbol{\Sigma}_{22}^{(1)})^{-1} \{ \mathbf{Z}'_1 \mathbf{Z}_1 + \mathbf{w}'_1 \mathbf{w}_1 \} \right] \right. \\
& \quad \left. + \text{tr} \left[ (\boldsymbol{\Sigma}_{11.2}^{(2)})^{-1} (\mathbf{X}_2 - \mathbf{Z}_2 \boldsymbol{\gamma}_2 - \mathbf{A} \boldsymbol{\Theta})' (\mathbf{X}_2 - \mathbf{Z}_2 \boldsymbol{\gamma}_2 - \mathbf{A} \boldsymbol{\Theta}) \right. \right. \\
& \quad \left. \left. + (\mathbf{y}_2 - \mathbf{w}_2 \boldsymbol{\gamma}_2)' (\mathbf{y}_2 - \mathbf{w}_2 \boldsymbol{\gamma}_2) \right] + \text{tr} \left[ (\boldsymbol{\Sigma}_{22}^{(2)})^{-1} \{ \mathbf{Z}'_2 \mathbf{Z}_2 + \mathbf{w}'_2 \mathbf{w}_2 \} \right] \right\}.
\end{aligned}$$

Furthermore, let  $\boldsymbol{\Gamma}^i$  be  $(N_i - m) \times (N_i - m)$  orthogonal matrices such that  $\boldsymbol{\Gamma}^i \mathbf{w}_i = [(\mathbf{w}'_i \mathbf{w}_i)^{1/2}, \mathbf{0}_{(p_i - q) \times (N_i - m - p_i + q)}]'$  and let  $\boldsymbol{\Gamma}^i \mathbf{y}_i = (\mathbf{u}'_i, \mathbf{s}'_i)'$ . Hence, from this orthogonal transformations  $\mathbf{y}_i \rightarrow \boldsymbol{\Gamma}^i \mathbf{y}_i$ , we complete the proof.  $\square$

For  $i = 1, 2$ , put  $\mathbf{S}_i = \mathbf{s}'_i \mathbf{s}_i$ ,  $\mathbf{W}_i = \mathbf{w}'_i \mathbf{w}_i$ ,  $\hat{\boldsymbol{\gamma}}_i = \mathbf{W}_i^{-1/2} \mathbf{u}_i$ ,  $\hat{\boldsymbol{\Theta}}_i = \mathbf{X}_i - \mathbf{Z}_i \hat{\boldsymbol{\gamma}}_i$ , and  $n_i = N_i - m - p_i + q$ . Now we consider the problem of estimating  $\boldsymbol{\Theta}$

based on  $(\widehat{\Theta}_i, \mathbf{S}_i)$ ,  $i = 1, 2$ , under the loss function (8). Its risk function is  $\mathbf{R}((\Theta, \Sigma_1, \Sigma_2), \widehat{\Theta}) = \mathbb{E}[L((\Theta, \Sigma_1, \Sigma_2), \widehat{\Theta})]$ , where the expectation is taken with respect to the density function given by (21). We consider a class of combined estimators of the form

$$\widehat{\Theta}^{EQ} = \widehat{\Theta}_1 \mathbf{B}' \Phi (\mathbf{B}')^{-1} + \mathbf{A}^{-1} \widehat{\Theta}_2 \mathbf{B}' (\mathbf{I}_q - \Phi) (\mathbf{B}')^{-1}, \quad (22)$$

where  $\mathbf{B}$  is a  $q \times q$  nonsingular matrix such that  $\mathbf{B}(\mathbf{S}_1 + \mathbf{S}_2)\mathbf{B}' = \mathbf{I}_q$ ,  $\mathbf{B}\mathbf{S}_2\mathbf{B}' = \mathbf{F}$ ,  $\mathbf{F} = \text{diag}(f_1, \dots, f_q)$  with  $f_1 \geq \dots \geq f_q$  and  $\Phi$  is a diagonal matrix whose  $i$ -th elements  $\phi_i$  ( $i = 1, 2, \dots, q$ ) are functions of  $\mathbf{F}$ .

To evaluate the risk of the estimators (22), 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  $(\mathbf{X}_i, \mathbf{Z}_i, \mathbf{s}_i, \mathbf{u}_i, \mathbf{w}_i | i = 1, 2)$  and define

$$\mathbb{E}_G[U] = \int U \times |\Sigma_1|^{-N_1/2} |\Sigma_2|^{-N_2/2} G(d) \prod_{i=1}^2 d\mathbf{X}_i d\mathbf{Z}_i d\mathbf{s}_i d\mathbf{u}_i d\mathbf{w}_i, \quad (23)$$

where  $G(x) = \frac{1}{2} \int_x^{+\infty} g(t) dt$  and  $d$  is given by the terms inside large curly bracket of (21).

**Theorem 3** *The risk of the estimator (22) is written as*

$$\begin{aligned} & \mathbf{R}((\Theta, \Sigma_1, \Sigma_2), \widehat{\Theta}^{EQ}) \\ &= \mathbb{E}_G \left[ q(\hat{r}_2 - \hat{r}_1) + \sum_{j=1}^q \left\{ 2(\hat{r}_1 - \hat{r}_2)\phi_j + (n_1 - q - 1) \frac{(1 - \phi_j)^2}{1 - f_j} \{\mathbf{H}_1\}_{jj} \right. \right. \\ & \quad + 4\{\mathbf{H}_1\}_{jj}(1 - \phi_j)f_j \frac{\partial \phi_j}{\partial f_j} + 2 \sum_{k \neq j} \{\mathbf{H}_1\}_{jj}(1 - \phi_j)(\phi_j - \phi_k) \frac{f_k}{f_j - f_k} \\ & \quad + (n_2 - q - 1) \frac{\phi_j^2}{f_j} \{\mathbf{H}_2\}_{jj} + 4\{\mathbf{H}_2\}_{jj}\phi_j(1 - f_j) \frac{\partial \phi_j}{\partial f_j} \\ & \quad \left. \left. + 2 \sum_{k \neq j} \{\mathbf{H}_2\}_{jj}\phi_j(\phi_j - \phi_k) \frac{1 - f_k}{f_j - f_k} \right\} \right], \quad (24) \end{aligned}$$

where  $\{\mathbf{H}_1\}_{jj}$  and  $\{\mathbf{H}_2\}_{jj}$  are  $j$ -th diagonal elements of the matrices given by (17a) and (17b), respectively, and

$$\hat{r}_1 = \text{tr}(\mathbf{I}_m + \mathbf{Z}_1 \mathbf{W}_1^{-1} \mathbf{Z}'_1), \quad (25a)$$

$$\hat{r}_2 = \text{tr} \{ (\mathbf{I}_m + \mathbf{Z}_2 \mathbf{W}_2^{-1} \mathbf{Z}'_2) (\mathbf{C}\mathbf{A}^{-1})' (\mathbf{C}\mathbf{A}^{-1}) \}. \quad (25b)$$

**Proof.** The proof of the theorem is put into Section 6.  $\square$

### 3.2 Choice of $\Phi$

Unlike the unbiased estimate of the risk for the multivariate normal error, it seems difficult to obtain to approximate  $\{\mathbf{H}_1\}_{jj}$  and  $\{\mathbf{H}_2\}_{jj}$  in (24) since the formula (24) involves in integration defined by (23). So we assume that  $N_1 = N_2 = N$ ,  $p_1 = p_2 = p$  and  $\mathbf{C}'\mathbf{C} = \mathbf{A}'\mathbf{A} = \mathbf{I}_m$ , which is the same case where Loh (1991) treated the problem of estimating the common mean of the multivariate normal distribution. From this additional assumption and using symmetry of the distributions, we can see that  $\mathbb{E}_G[\text{tr}(\mathbf{Z}_1\mathbf{W}_1^{-1}\mathbf{Z}'_1)] = \mathbb{E}_G[\text{tr}(\mathbf{Z}_2\mathbf{W}_2^{-1}\mathbf{Z}'_2)]$  and that  $\mathbb{E}_G[\phi_j \text{tr}(\mathbf{Z}_1\mathbf{W}_1^{-1}\mathbf{Z}'_1)] = \mathbb{E}_G[\phi_j \text{tr}(\mathbf{Z}_2\mathbf{W}_2^{-1}\mathbf{Z}'_2)]$  for  $j = 1, 2, \dots, q$ . These imply that  $\mathbb{E}_G[\hat{r}_1 - \hat{r}_2] = 0$  and  $\mathbb{E}_G[(\hat{r}_1 - \hat{r}_2)\phi_j] = 0$  for  $j = 1, 2, \dots, q$ . Also note that  $\mathbf{H}_1 = \mathbf{H}_2$ . Thus the risk can be written as

$$\begin{aligned}
& \mathbf{R}((\Theta, \Sigma_1, \Sigma_2), \widehat{\Theta}^{EQ}) \\
&= \mathbb{E}_G \left[ \sum_{j=1}^q \{\mathbf{H}_1\}_{jj} \left\{ (n_0 - q - 1) \frac{(1 - \phi_j)^2}{1 - f_j} + 4(1 - \phi_j) f_j \frac{\partial \phi_j}{\partial f_j} \right. \right. \\
&+ 2 \sum_{k \neq j} (1 - \phi_j)(\phi_j - \phi_k) \frac{f_k}{f_j - f_k} + (n_0 - q - 1) \frac{\phi_j^2}{f_j} + 4\phi_j(1 - f_j) \frac{\partial \phi_j}{\partial f_j} \\
&\left. \left. + 2 \sum_{k \neq j} \phi_j(\phi_j - \phi_k) \frac{1 - f_k}{f_j - f_k} \right\} \right] \equiv \mathbb{E}_G[\widetilde{\mathbf{R}}_0], \tag{26}
\end{aligned}$$

where  $n_0 = N - m - p + q$  ( $N = N_1 = N_2$ ,  $p = p_1 = p_2$ ).

Now we use the relation (19) and ignore the derivative terms, then we derivate  $\widetilde{\mathbf{R}}_0$  with respect to  $\phi_j$  separately, to get

$$\begin{aligned}
0 = \frac{\partial \widetilde{\mathbf{R}}_0}{\partial \phi_j} &= \{\mathbf{H}_1\}_{jj} \times \left\{ -(n_0 - q - 1) \frac{1 - \phi_j}{1 - f_j} - 4(1 - \phi_j) \frac{f_j}{1 - f_j} \right. \\
&+ 2 \sum_{k \neq j} (1 - \phi_j) \frac{f_k}{f_j - f_k} - \sum_{k \neq j} (1 - \phi_k) \frac{f_k}{f_j - f_k} \\
&+ (n_0 - q - 1) \frac{\phi_j}{f_j} + 4 \frac{1 - f_j}{f_j} \phi_j \\
&\left. + 2\phi_j \sum_{k \neq j} \frac{1 - f_k}{f_j - f_k} - \sum_{k \neq j} \phi_k \frac{1 - f_k}{f_j - f_k} \right\}.
\end{aligned}$$

Hence, solving for  $\phi_j$  with ignoring the fourth and the eighth terms in the large curly bracket of the last right-hand side above, we get

$$\hat{\phi}_j^{ST} = \frac{\hat{\beta}_j^{ST}/(1 - f_j)}{\hat{\beta}_j^{ST}/(1 - f_j) + \hat{\alpha}_j^{ST}/f_j}, \tag{27}$$

where

$$\hat{\alpha}_j^{ST} = n_0 - q - 1 + 4(1 - f_j) + 2 \sum_{k \neq j} \frac{f_j(1 - f_k)}{f_j - f_k},$$

$$\hat{\beta}_j^{ST} = n_0 - q - 1 + 4f_j - 2 \sum_{k \neq j} \frac{(1 - f_j)f_k}{f_j - f_k}.$$

Finally, we reach to an alternative estimator of the form (22) with (27).

## 4 Numerical studies

### 4.1 Numerical study for GMANOVA under normal errors

Since the risk of the Stein type estimator 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 the canonical form (6a)–(7d) with special cases for  $(N_1, N_2, p_1, p_2, m, q)$ . These results are given in Table 1.

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

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

and

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

For  $(\boldsymbol{\Sigma}_{11.2}^{(1)}, \boldsymbol{\Sigma}_{11.2}^{(2)})$ , we assume that the eigenvalues of  $\boldsymbol{\Sigma}_{11.2}^{(2)}(\boldsymbol{\Sigma}_{11.2}^{(1)})^{-1}$  are close together and that these eigenvalues are widely spread out. Furthermore, we put  $\Theta = \mathbf{0}$ ,  $\boldsymbol{\Sigma}_{22}^{(1)} = \boldsymbol{\Sigma}_{22}^{(2)} = \mathbf{I}_2$ , and  $\boldsymbol{\Sigma}_{12}^{(1)} = \boldsymbol{\Sigma}_{12}^{(2)} = \mathbf{0}$ .

Recall that, when  $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$  is known, the maximum likelihood estimator of  $\Theta$  in (6a) and (6b) is given by

$$\begin{aligned} \text{vec}(\widetilde{\Theta}^{ML}) &= [\mathbf{I}_m \otimes (\boldsymbol{\Sigma}_{11.2}^{(1)})^{-1} + \mathbf{A}'\mathbf{A} \otimes (\boldsymbol{\Sigma}_{11.2}^{(2)})^{-1}]^{-1} \\ &\times \{[\mathbf{I}_m \otimes (\boldsymbol{\Sigma}_{11.2}^{(1)})^{-1}] \text{vec}(\widetilde{\Theta}_1) + \{\mathbf{A}'\mathbf{A} \otimes (\boldsymbol{\Sigma}_{11.2}^{(2)})^{-1}\} \text{vec}(\mathbf{A}^{-1}\widetilde{\Theta}_2)\}, \end{aligned} \quad (28)$$

where  $\widetilde{\Theta}_i = \mathbf{X}_i - \boldsymbol{\gamma}_i \mathbf{Z}_i$  ( $i = 1, 2$ ). Here the risk of  $\text{vec}(\widetilde{\Theta}^{ML})$  is evaluated as follows:

**Lemma 2**

$$\begin{aligned} \mathbf{R}((\Theta, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2), \widetilde{\Theta}^{ML}) \\ = \text{tr} \{[\mathbf{I}_m \otimes (\boldsymbol{\Sigma}_{11.2}^{(1)})^{-1} + (\mathbf{C}'\mathbf{C}) \otimes (\boldsymbol{\Sigma}_{11.2}^{(2)})^{-1}] \\ \times [\mathbf{I}_m \otimes (\boldsymbol{\Sigma}_{11.2}^{(1)})^{-1} + (\mathbf{A}'\mathbf{A}) \otimes (\boldsymbol{\Sigma}_{11.2}^{(2)})^{-1}]^{-1}\}. \end{aligned}$$

Furthermore, if  $\mathbf{A}'\mathbf{A} = \mathbf{C}'\mathbf{C}$ , then  $\mathbf{R}((\Theta, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2), \widetilde{\Theta}^{ML}) = mq$ .

In Table 1, ‘‘ML’’ indicates the maximum likelihood estimator (28) and its risk value was calculated by Lemma 2. Moreover, ‘‘SK’’ and ‘‘ST’’ denote the Graybill-Deal type estimator (13) by Sugiura and Kubokawa (1988) and the Stein type estimator, respectively, and estimated standard errors are in parentheses. Here, the Stein type estimator is of the form

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

where  $\{\bar{\alpha}_j^{ST}\}_{j=1}^q$  and  $\{\bar{\beta}_j^{ST}\}_{j=1}^q$  are made from Stein’s isotonic regressions on  $\{\hat{\alpha}_j^{ST}/f_j\}_{j=1}^q$  and on  $\{\hat{\beta}_j^{ST}/(1-f_j)\}_{j=1}^q$ , respectively, and  $(\hat{\alpha}_j^{ST}$  and  $\hat{\beta}_j^{ST})$  are given by

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

Note that we modify  $\hat{\alpha}_j$  and  $\hat{\beta}_j$  in (20) as above by ignoring the second terms  $(r_1 - r_2)f_j$  in  $\hat{\alpha}_j$  and  $(r_2 - r_1)(1 - f_j)$  in  $\hat{\beta}_j$ . For a detailed description of Stein’s isotonic regression, see Lin and Perlman (1985). Furthermore, ‘‘AV’’ in Table 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:

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

Eigenvalues of $\Sigma_{11.2}^{(2)}(\Sigma_{11.2}^{(1)})^{-1}$	ML	SK	ST	AV
<b><math>A'A = C'C = \text{diag}(1, 1)</math></b>				
<b><math>N_1 = N_2 = 12, p_1 = p_2 = 7, m = 2, q = 5</math></b>				
(1, 1, 1, 1, 1)	10	19.39 (0.106)	16.61 (0.090)	14.4 %
(10, 0.1, 0.1, 0.1, 0.1)	10	20.62 (0.132)	18.60 (0.116)	9.8 %
( $10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$ )	10	18.00 (0.156)	18.00 (0.156)	0.0 %
( $10^8, 10^4, 1, 10^{-4}, 10^{-8}$ )	10	20.27 (0.128)	20.44 (0.129)	-0.8 %
<b><math>N_1 = N_2 = 20, p_1 = p_2 = 12, m = 2, q = 10</math></b>				
(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	20	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)	20	36.07 (0.157)	28.75 (0.108)	20.3 %
( $10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$ )	20	27.52 (0.106)	27.52 (0.106)	0.0 %
( $10^5, 10^4, 10^3, 10^2, 10, 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$ )	20	33.97 (0.128)	34.01 (0.128)	-0.1 %
<b><math>A'A = \text{diag}(3, 1/3), C'C = \text{diag}(1, 1)</math></b>				
<b><math>N_1 = N_2 = 12, p_1 = p_2 = 7, m = 2, q = 5</math></b>				
(1, 1, 1, 1, 1)	10.00	20.39 (0.132)	16.87 (0.106)	17.3 %
(10, 0.1, 0.1, 0.1, 0.1)	13.48	27.31 (0.189)	24.50 (0.165)	10.3 %
( $10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$ )	15.33	25.94 (0.203)	25.94 (0.203)	0.0 %
( $10^8, 10^4, 1, 10^{-4}, 10^{-8}$ )	12.67	26.18 (0.203)	26.31 (0.203)	-0.5 %
<b><math>N_1 = N_2 = 20, p_1 = p_2 = 12, m = 2, q = 10</math></b>				
(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	20.00	35.45 (0.150)	26.19 (0.103)	26.1 %
(10, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)	27.95	48.96 (0.232)	39.43 (0.171)	19.5 %
( $10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$ )	32.00	42.11 (0.183)	42.11 (0.183)	0.0 %
( $10^5, 10^4, 10^3, 10^2, 10, 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$ )	24.73	42.97 (0.200)	43.03 (0.201)	-0.1 %

1. In Table 1, when the eigenvalues of  $\Sigma_{11.2}^{(2)}(\Sigma_{11.2}^{(1)})^{-1}$  are close together, the AVs are large. Specially, in cases when  $\mathbf{A}'\mathbf{A} = \mathbf{diag}(3, 1/3)$ ,  $\mathbf{C}'\mathbf{C} = \mathbf{diag}(1, 1)$ ,  $N_1 = N_2 = 20$ ,  $p_1 = p_2 = 12$ ,  $m = 2$ ,  $q = 10$ , and these eigenvalues are equal to 1, the AV is 26.1%.
2. On the contrary, when the eigenvalues of  $\Sigma_{11.2}^{(2)}(\Sigma_{11.2}^{(1)})^{-1}$  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  $p$  and fixed sample-size  $N$ .

**Remark 2.** Under another assumptions for  $\Sigma_{11.2}^{(2)}(\Sigma_{11.2}^{(1)})^{-1}$  as examined by Loh (1991), we simulated the risk values of GD and ST and obtained the results that ST performs better than GD.

#### 4.2 Numerical study for estimating the common mean under elliptical errors

First we illustrate the model (1) with the density (2) and estimators when  $N_1 = N_2 = N$ ,  $m = 1$ ,  $p_1 = p_2 = q_1 = q_2 = p$ ,  $\mathbf{A}_{11} = \mathbf{A}_{21} = \mathbf{1}_N$  and  $\mathbf{A}_{12} = \mathbf{A}_{22} = \mathbf{I}_p$ . From an orthogonal transformation in the similar way as in Section 3, we obtain a canonical form of density (2) as

$$|\Sigma_1|^{-N/2}|\Sigma_2|^{-N/2}g\left(\sum_{i=1}^2[\text{tr}\{\Sigma_i^{-1}(\mathbf{X}_i - \boldsymbol{\theta})(\mathbf{X}_i - \boldsymbol{\theta})' + \Sigma_i^{-1}\mathbf{S}_i\}]\right), \quad (29)$$

where  $\boldsymbol{\theta} = \sqrt{N}\boldsymbol{\xi}$ ,  $\Omega_i = \Sigma_i$ ,  $\mathbf{X}_i = \mathbf{Y}'_i\mathbf{1}_N/\sqrt{N}$ , and  $\mathbf{S}_i = \mathbf{Y}'_i(\mathbf{I}_N - \mathbf{1}_N\mathbf{1}'_N/N)\mathbf{Y}_i$  for  $i = 1, 2$ . Therefore, the problem of estimating  $\boldsymbol{\xi}$  in (2) turns into that of estimating the common mean vector  $\boldsymbol{\theta}$  in (29). Then, if  $g$  is decreasing and  $(\Sigma_1, \Sigma_2)$  is known, we can see that the maximum likelihood estimator is of the form

$$\hat{\boldsymbol{\theta}}^{ML} = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}(\Sigma_1^{-1}\mathbf{X}_1 + \Sigma_2^{-1}\mathbf{X}_2).$$

Furthermore, the Graybill-Deal type estimator can be written as

$$\hat{\boldsymbol{\theta}}^{GD} = (\mathbf{S}_1^{-1} + \mathbf{S}_2^{-1})^{-1}(\mathbf{S}_1^{-1}\mathbf{X}_1 + \mathbf{S}_2^{-1}\mathbf{X}_2) \quad (30)$$

and also the Stein type estimator as

$$\hat{\boldsymbol{\theta}} = \mathbf{B}^{-1}\Phi^{ST}\mathbf{B}\mathbf{X}_1 + \mathbf{B}^{-1}(\mathbf{I}_p - \Phi^{ST})\mathbf{B}\mathbf{X}_2, \quad (31)$$

where  $\mathbf{B}(\mathbf{S}_1 + \mathbf{S}_2)\mathbf{B}' = \mathbf{I}_p$ ,  $\mathbf{B}\mathbf{S}_2\mathbf{B}' = \mathbf{F} = \text{diag}(f_1, \dots, f_p)$  with  $f_1 \geq \dots \geq f_p$  and

$$\begin{aligned}
\Phi^{ST} &= \mathbf{diag}(\hat{\phi}_1^{ST}, \dots, \hat{\phi}_p^{ST}), \\
\hat{\phi}_j^{ST} &= \frac{\hat{\beta}_j^{ST}/(1-f_j)}{\hat{\beta}_j^{ST}/(1-f_j) + \hat{\alpha}_j^{ST}/f_j}, \\
\hat{\alpha}_j^{ST} &= (N-1) - p - 1 + 4(1-f_j) + 2 \sum_{k \neq j} \frac{f_j(1-f_k)}{f_j - f_k}, \\
\hat{\beta}_j^{ST} &= (N-1) - p - 1 + 4f_j - 2 \sum_{k \neq j} \frac{(1-f_j)f_k}{f_j - f_k}.
\end{aligned}$$

Since the model (29) is not i.i.d. sampling set-up of two sample problems, we carry out Monte-Carlo simulation to show that our proposed estimator (31) reduces the risk over the Graybill-Deal estimator (30) under the i.i.d. sampling from two independent multivariate elliptically contoured distributions instead of sampling from the model (29). Hence, we carry out Monte-Carlo simulation when we sample  $(\mathbf{Y}_1, \mathbf{Y}_2)$  which can be represented as

$$\mathbf{Y}_1 = \mathbf{1}_N \boldsymbol{\xi}' + \boldsymbol{\epsilon}_1 \quad \text{and} \quad \mathbf{Y}_2 = \mathbf{1}_N \boldsymbol{\xi}' + \boldsymbol{\epsilon}_2, \quad (32)$$

where  $\mathbf{Y}_1, \mathbf{Y}_2, \boldsymbol{\epsilon}_1,$  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

$$|\boldsymbol{\Sigma}_i|^{-N/2} h(\mathbf{e}'_{ij} \boldsymbol{\Sigma}_i^{-1} \mathbf{e}_{ij}), \quad i = 1, 2, j = 1, \dots, N, \quad (33)$$

where  $\boldsymbol{\epsilon}_i = (\mathbf{e}_{i1}, \mathbf{e}_{i2}, \dots, \mathbf{e}_{iN})'$  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 (33), we consider an improvement under density (29). However, our simulation results justify our derivation of alternative estimator under the model (29).

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

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

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

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

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  $\xi = \mathbf{0}$  and that  $\Sigma_2 \Sigma_1^{-1}$  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. For the Stein type estimator, we modified  $\Phi^{ST}$  by means of the Stein isotonic regression. These simulation results are given in Tables 2 and 3, respectively. In tables, “ML”, “GD”, and “ST” denote  $\hat{\theta}^{ML}$ ,  $\hat{\theta}^{GD}$ , and  $\hat{\theta}^{ST}$ , respectively, and “AV” is the average of improvement in risk of ST against GD.

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

Eigenvalues of $\Sigma_2 \Sigma_1^{-1}$	ML	GD	ST	AV
<b><math>N = 8, \quad p = 5</math></b>				
(1, 1, 1, 1, 1)	14.186 (0.504)	26.927 (1.272)	24.271 (1.423)	9.86 %
(10, 0.1, 0.1, 0.1, 0.1)	14.250 (0.551)	32.441 (2.726)	28.716 (2.292)	11.48 %
( $10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$ )	15.651 (1.094)	29.349 (2.148)	29.349 (2.148)	0.00 %
( $10^8, 10^4, 1, 10^{-4}, 10^{-8}$ )	14.912 (0.542)	29.434 (0.992)	29.784 (1.017)	-1.19 %
<b><math>N = 13, \quad p = 10</math></b>				
(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	30.478 (0.992)	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)	27.553 (0.970)	84.003 (2.670)	60.471 (1.803)	28.01 %
( $10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$ )	28.401 (0.863)	47.988 (1.340)	47.988 (1.340)	0.00 %
( $10^5, 10^4, 10^3, 10^2, 10, 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$ )	28.307 (0.907)	63.451 (2.292)	63.348 (2.267)	0.16 %

We summarize these results as follows:

1. In almost cases, the AVs are positive. These are large when the eigenvalues of  $\Sigma_2 \Sigma_1^{-1}$  are close together, and particularly, when only one of these eigenvalues is 10 with  $(N, p) = (13, 10)$ , the AVs are more than 27%.
2. On the contrary, when the eigenvalues of  $\Sigma_2 \Sigma_1^{-1}$  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  $\Sigma_2 \Sigma_1^{-1}$  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  $\Sigma_2 \Sigma_1^{-1}$  are the same, it is expected that the AVs increase with increasing dimension  $p$  and small sample-size  $N$ .

Table 3: Estimated risks under Kotz-type distributions with

$$(u, r, s) = (5, 0.5, 2)$$

(Estimated standard errors are in parentheses)

Eigenvalues of $\Sigma_2 \Sigma_1^{-1}$	ML	GD	ST	AV
<b><math>N = 8, \quad p = 5</math></b>				
(1, 1, 1, 1, 1)	2.465 (0.012)	3.755 (0.021)	3.143 (0.017)	16.32 %
(10, 0.1, 0.1, 0.1, 0.1)	2.420 (0.010)	3.970 (0.027)	3.572 (0.023)	10.01 %
( $10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$ )	2.452 (0.010)	3.509 (0.029)	3.509 (0.029)	0.00 %
( $10^8, 10^4, 1, 10^{-4}, 10^{-8}$ )	2.435 (0.011)	3.846 (0.024)	3.878 (0.024)	-0.83 %
<b><math>N = 13, \quad p = 10</math></b>				
(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	2.922 (0.010)	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)	2.927 (0.009)	6.151 (0.038)	4.484 (0.024)	27.11 %
( $10^{10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}, 10^{-10}$ )	2.928 (0.008)	4.460 (0.064)	4.460 (0.064)	0.00 %
( $10^5, 10^4, 10^3, 10^2, 10, 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$ )	2.900 (0.010)	5.107 (0.024)	5.107 (0.024)	0.00 %

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 (29) remains robust even if the rows of errors are i.i.d.

## 5 Proof of Theorem 2

In this section, we state lemmas which are useful in proving the main theorems. These include some computational lemmas on moments of the maximum likelihood estimators, integration-by-parts formulae, and calculus lemmas on eigenstructures. Once we introduce the lemmas, it is straightforward to give the proof of Theorem 2.

**Lemma 3** *Let  $r_1 = m\tilde{r}_1$ ,  $r_2 = \tilde{r}_2 \text{tr}\{(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}$ ,  $\tilde{r}_i = (n_i + p_i - q - 1)/(n_i - 1)$ ,  $i = 1, 2$ . Then we have*

$$\mathbb{E}[\text{tr}\{(\widehat{\Theta}_1 - \Theta)(\Sigma_{11.2}^{(1)})^{-1}(\widehat{\Theta}_1 - \Theta)'\}] = qr_1, \quad (34a)$$

$$\mathbb{E}[\text{tr}\{(\widehat{\Theta}_2 - \mathbf{A}\Theta)(\Sigma_{11.2}^{(2)})^{-1}(\widehat{\Theta}_2 - \mathbf{A}\Theta)'(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}] = qr_2, \quad (34b)$$

$$\begin{aligned} \mathbb{E}[\text{tr}\{(\widehat{\Theta}_1 - \Theta)(\Sigma_{11.2}^{(1)})^{-1}\mathbf{B}^{-1}(\mathbf{I}_q - \Phi)\mathbf{B}(\mathbf{A}^{-1}\widehat{\Theta}_2 - \widehat{\Theta}_1)'\}] \\ = -\mathbb{E}\left[\left(q - \sum_{i=1}^q \phi_i\right)r_1\right], \end{aligned} \quad (34c)$$

$$\begin{aligned} \mathbb{E}[\text{tr}\{(\widehat{\Theta}_2 - \mathbf{A}\Theta)(\Sigma_{11.2}^{(2)})^{-1}\mathbf{B}^{-1}\Phi\mathbf{B}(\mathbf{A}\widehat{\Theta}_1 - \widehat{\Theta}_2)'(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}] \\ = -\mathbb{E}\left[\left(\sum_{i=1}^q \phi_i\right)r_2\right]. \end{aligned} \quad (34d)$$

**Proof.** Note that

$$\begin{aligned} \widehat{\Theta}_1 | \mathbf{Z}_1, \mathbf{W}_1 &\sim N_{m \times q}(\Theta, (\mathbf{I}_m + \mathbf{Z}_1\mathbf{W}_1^{-1}\mathbf{Z}_1') \otimes \Sigma_{11.2}^{(1)}), \\ \widehat{\Theta}_2 | \mathbf{Z}_2, \mathbf{W}_2 &\sim N_{m \times q}(\mathbf{A}\Theta, (\mathbf{I}_m + \mathbf{Z}_2\mathbf{W}_2^{-1}\mathbf{Z}_2') \otimes \Sigma_{11.2}^{(2)}), \end{aligned}$$

and that  $\widehat{\Theta}_1$  and  $\widehat{\Theta}_2$  are independent. Use the fact that  $\mathbb{E}[XQX'] = \text{tr}(Q'\Sigma)\Psi + MQM'$  when  $X \sim N_{m \times n}(M, \Psi \otimes \Sigma)$  to get

$$\begin{aligned} \mathbb{E}[\text{tr}\{(\widehat{\Theta}_1 - \Theta)(\Sigma_{11.2}^{(1)})^{-1}(\widehat{\Theta}_1 - \Theta)'\}] &= \mathbb{E}[q \text{tr}(\mathbf{I}_m + \mathbf{Z}_1\mathbf{W}_1^{-1}\mathbf{Z}_1')], \\ \mathbb{E}[\text{tr}\{(\widehat{\Theta}_2 - \mathbf{A}\Theta)(\Sigma_{11.2}^{(2)})^{-1}(\widehat{\Theta}_2 - \mathbf{A}\Theta)'(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}] \\ &= \mathbb{E}[q \text{tr}\{(\mathbf{I}_m + \mathbf{Z}_2\mathbf{W}_2^{-1}\mathbf{Z}_2')(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}], \\ \mathbb{E}[\text{tr}\{(\widehat{\Theta}_1 - \Theta)(\Sigma_{11.2}^{(1)})^{-1}\mathbf{B}^{-1}(\mathbf{I}_q - \Phi)\mathbf{B}(\mathbf{A}^{-1}\widehat{\Theta}_2 - \widehat{\Theta}_1)'\}] \\ &= -\mathbb{E}[\text{tr}\{(\widehat{\Theta}_1 - \Theta)(\Sigma_{11.2}^{(1)})^{-1}\mathbf{B}^{-1}(\mathbf{I}_q - \Phi)\mathbf{B}(\widehat{\Theta}_1 - \Theta)'\}] \\ &= -\mathbb{E}[\text{tr}\{\mathbf{B}^{-1}(\mathbf{I}_q - \Phi)\mathbf{B}\} \times \text{tr}(\mathbf{I}_m + \mathbf{Z}_1\mathbf{W}_1^{-1}\mathbf{Z}_1')], \\ \mathbb{E}[\text{tr}\{(\widehat{\Theta}_2 - \mathbf{A}\Theta)(\Sigma_{11.2}^{(2)})^{-1}\mathbf{B}^{-1}\Phi\mathbf{B}(\mathbf{A}\widehat{\Theta}_1 - \widehat{\Theta}_2)'(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}] \\ &= -\mathbb{E}[\text{tr}\{(\widehat{\Theta}_2 - \mathbf{A}\Theta)(\Sigma_{11.2}^{(2)})^{-1}\mathbf{B}^{-1}\Phi\mathbf{B}(\widehat{\Theta}_2 - \mathbf{A}\Theta)'(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}] \\ &= -\mathbb{E}[\text{tr}\{\mathbf{B}^{-1}\Phi\mathbf{B}\} \times \text{tr}\{(\mathbf{I}_m + \mathbf{Z}_2\mathbf{W}_2^{-1}\mathbf{Z}_2')(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}]. \end{aligned}$$

Finally, from (7a) and (7d), we get (34a)–(34d).  $\square$

**Lemma 4 (Stein-Haff identity)** *Assume that a  $q \times q$  positive definite matrix  $\mathbf{S}$  follows the Wishart distribution  $W_q(\Sigma, a)$ . Also let*

$$\mathcal{D} = \left( \frac{1}{2}(1 + \delta_{ij}) \frac{\partial}{\partial s_{ij}} \right), \quad (35)$$

where  $s_{ij}$  are the  $(i, j)$ -th elements of  $\mathbf{S}$  and  $\delta_{ij}$  is the Kronecker delta. For a suitable  $q \times q$  matrix  $\mathbf{V}$  we have

$$\mathbb{E}[\text{tr}(\mathbf{V}\Sigma^{-1})] = \mathbb{E}[2 \text{tr}(\mathcal{D}\mathbf{V}) + (a - q - 1) \text{tr}(\mathbf{S}^{-1}\mathbf{V})].$$

**Lemma 5 (Loh, 1988 and 1991)** For  $i = 1, 2$ , let  $\mathcal{D}_i$  be  $q \times q$  differential operators which are define by (35) with replacing  $\mathbf{S}$  by  $\mathbf{S}_i$ . Also let  $\mathbf{x}$  be a  $q \times 1$  vector which is independent of  $\mathbf{S}_1$  and  $\mathbf{S}_2$ . Then

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

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

Note here that  $\{ \mathbf{B} \mathbf{x} \}_j^2 = \{ \mathbf{B} \mathbf{x} \}_j \{ \mathbf{x}' \mathbf{B}' \}_j = \{ \mathbf{B} \mathbf{x} \mathbf{x}' \mathbf{B}' \}_{jj}$ , where  $\{ \mathbf{B} \mathbf{x} \mathbf{x}' \mathbf{B}' \}_{jj}$  denote the  $(j, j)$ -elements of  $\mathbf{B} \mathbf{x} \mathbf{x}' \mathbf{B}'$ . Hence we have

**Lemma 6**

$$\begin{aligned} & \text{tr} \{ \mathcal{D}_1 [ \mathbf{B}^{-1} (\mathbf{I}_q - \Phi) \mathbf{H}_1 (\mathbf{I}_q - \Phi) (\mathbf{B}')^{-1} ] \} \\ &= \sum_{j=1}^q \left[ \{ \mathbf{H}_1 \}_{jj} (1 - \phi_j)^2 \sum_{k \neq j} \frac{f_k}{f_k - f_j} + 2 \{ \mathbf{H}_1 \}_{jj} (1 - \phi_j) f_j \frac{\partial \phi_j}{\partial f_j} \right. \\ & \quad \left. - \sum_{k \neq j} \{ \mathbf{H}_1 \}_{kk} (1 - \phi_j) (1 - \phi_k) \frac{f_j}{f_j - f_k} \right], \\ & \text{tr} \{ \mathcal{D}_2 [ \mathbf{B}^{-1} \Phi \mathbf{H}_2 \Phi (\mathbf{B}')^{-1} ] \} \\ &= \sum_{j=1}^q \left[ \{ \mathbf{H}_2 \}_{jj} \phi_j^2 \sum_{k \neq j} \frac{1 - f_k}{f_j - f_k} + 2 \{ \mathbf{H}_2 \}_{jj} \phi_j (1 - f_j) \frac{\partial \phi_j}{\partial f_j} \right. \\ & \quad \left. - \sum_{k \neq j} \{ \mathbf{H}_2 \}_{kk} \phi_j \phi_k \frac{1 - f_j}{f_k - f_j} \right], \end{aligned}$$

where  $\mathbf{H}_1$  and  $\mathbf{H}_2$  are given by (17a) and (17b), respectively.

**Proof.** If we put  $\mathbf{A}^{-1} \widehat{\Theta}_2 - \widehat{\Theta}_1 = (\mathbf{x}_1, \dots, \mathbf{x}_m)'$ , we can see that

$$\mathbf{H}_1 = \mathbf{B} (\mathbf{A}^{-1} \widehat{\Theta}_2 - \widehat{\Theta}_1)' (\mathbf{A}^{-1} \widehat{\Theta}_2 - \widehat{\Theta}_1) \mathbf{B}' = \sum_{l=1}^m \mathbf{B} \mathbf{x}_l \mathbf{x}_l' \mathbf{B}'.$$

Hence, from this equation and Lemma 5, we get the first expression. The second expression can be obtained from the similar argument.  $\square$

**Proof of Theorem 2.** First apply Lemma 4 to the third and sixth terms in right-hand side of (16) and then use Lemma 3 to the other terms in right-hand side of (16) to get that the risk  $\mathbf{R}((\Theta, \Sigma_1, \Sigma_2), \widehat{\Theta}^{EQ})$  is rewritten as

$$\begin{aligned} & q(r_2 - r_1) + \mathbb{E} \left[ 2(r_1 - r_2) \sum_{j=1}^q \phi_j + \text{tr} \left\{ (n_1 - q - 1) \mathbf{S}_1^{-1} \mathbf{B}^{-1} (\mathbf{I}_q - \Phi) \mathbf{H}_1 \right. \right. \\ & \times (\mathbf{I}_q - \Phi) (\mathbf{B}')^{-1} + 2\mathcal{D}_1 [\mathbf{B}^{-1} (\mathbf{I}_q - \Phi) \mathbf{H}_1 (\mathbf{I}_q - \Phi) (\mathbf{B}')^{-1}] \\ & \left. \left. + (n_2 - q - 1) \mathbf{S}_2^{-1} \mathbf{B}^{-1} \Phi \mathbf{H}_2 \Phi (\mathbf{B}')^{-1} + 2\mathcal{D}_2 [\mathbf{B}^{-1} \Phi \mathbf{H}_2 \Phi (\mathbf{B}')^{-1}] \right\} \right]. \end{aligned}$$

Finally apply Lemma 6 to the third and fourth terms inside the expectation of the above equation to complete the theorem.  $\square$

## 6 Proof of Theorem 3

In this section, we state lemmas which are useful in proving Theorem 3. These lemmas are counterparts of the lemmas given in the previous section, which is extended under the elliptically contoured distributions. Then we give the proof of Theorem 3. For  $i = 1, 2$ , let  $\mathbf{Q}_i \equiv \mathbf{Q}_i(\mathbf{X}_i)$  be  $q \times m$  matrix-valued functions of  $\mathbf{X}_i = (x_{i.jk})$  and let  $\mathbf{K}_i \equiv \mathbf{K}_i(\mathbf{u}_i)$  be  $q \times (p_i - q)$  matrix-valued functions of  $\mathbf{u}_i = (u_{i.jk})$ .

Denote differential operators in terms of  $\mathbf{X}_i$  and  $\mathbf{u}_i$  by

$$\nabla_{\mathbf{X}_i} = \left( \frac{\partial}{\partial x_{i.jk}} \right) \quad \text{and} \quad \nabla_{\mathbf{u}_i} = \left( \frac{\partial}{\partial u_{i.jk}} \right).$$

Here, the actions of  $\nabla_{\mathbf{X}_i}$  on  $\mathbf{Q}_i$  and of  $\nabla_{\mathbf{u}_i}$  on  $\mathbf{Q}_i$  and  $\mathbf{K}_i$  are defined as

$$\nabla_{\mathbf{X}_i} \mathbf{Q}_i = \left( \sum_{a=1}^q \frac{\partial \mathbf{Q}_{i.ak}}{\partial x_{i.ja}} \right), \quad \nabla_{\mathbf{u}_i} \mathbf{K}_i = \left( \sum_{a=1}^q \frac{\partial \mathbf{K}_{i.ak}}{\partial u_{i.ja}} \right).$$

**Lemma 7** *Let  $\Psi$  and  $\mathbf{C}$  be, respectively,  $q \times q$  and  $m \times m$  matrices. Then*

$$\text{tr}(\nabla_{\mathbf{X}_i} \Psi \mathbf{X}_i' \mathbf{C}) = (\text{tr} \Psi)(\text{tr} \mathbf{C}).$$

**Lemma 8 (Kubokawa and Srivastava, 2001)** *For  $i = 1, 2$ ,  $j = 1, \dots, q$ ,  $k = 1, \dots, m$ , suppose that each element of  $\mathbf{Q}_i \equiv \mathbf{Q}_i(\mathbf{X}_i)$  is differentiable with respect to  $x_{i.jk}$  and also, for  $i = 1, 2$ ,  $j = 1, \dots, p_i - q$ ,  $k = 1, \dots, m$ , that elements of  $\mathbf{K}_i \equiv \mathbf{K}_i(\mathbf{u}_i)$  are differentiable with respect to  $u_{i.jk}$ . Furthermore, assume that*

- (i) *there exists finite expectation of the absolute value of each element of the following matrices:*

$$\begin{aligned} & (\mathbf{X}_1 - \mathbf{Z}_1\boldsymbol{\gamma}_1 - \boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11.2}^{(1)})^{-1}\mathbf{Q}_1, \\ & (\mathbf{X}_2 - \mathbf{Z}_2\boldsymbol{\gamma}_1 - \mathbf{A}\boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11.2}^{(2)})^{-1}\mathbf{Q}_2, \\ & (u_i - \mathbf{W}_i^{1/2}\boldsymbol{\gamma}_i)(\boldsymbol{\Sigma}_{11.2}^{(i)})^{-1}\mathbf{K}_i; \end{aligned}$$

- (ii)  $\lim_{x_{i,jk} \rightarrow \pm\infty} \mathbf{Q}_i(\mathbf{X}_i)G(x_{i,jk}^2 + a^2) = \mathbf{0}$  for  $i = 1, 2, j = 1, \dots, q, k = 1, \dots, m$ ;  
(iii)  $\lim_{u_{i,jk} \rightarrow \pm\infty} \mathbf{K}_i(\mathbf{u}_i)G(u_{i,jk}^2 + a^2) = \mathbf{0}$  for  $i = 1, 2, j = 1, \dots, p_i - q, k = 1, \dots, m$ .

Then, for  $i = 1, 2$ , we have

$$\mathbb{E}[\text{tr}\{(\mathbf{X}_1 - \mathbf{Z}_1\boldsymbol{\gamma}_1 - \boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11.2}^{(1)})^{-1}\mathbf{Q}_1\}] = \mathbf{E}_G[\text{tr}(\nabla_{\mathbf{X}_1}\mathbf{Q}_1)], \quad (36a)$$

$$\mathbb{E}[\text{tr}\{(\mathbf{X}_2 - \mathbf{Z}_2\boldsymbol{\gamma}_1 - \mathbf{A}\boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11.2}^{(2)})^{-1}\mathbf{Q}_2\}] = \mathbf{E}_G[\text{tr}(\nabla_{\mathbf{X}_2}\mathbf{Q}_2)], \quad (36b)$$

$$\mathbb{E}[\text{tr}\{(u_i - \mathbf{W}_i^{1/2}\boldsymbol{\gamma}_i)(\boldsymbol{\Sigma}_{11.2}^{(i)})^{-1}\mathbf{K}_i\}] = \mathbf{E}_G[\text{tr}(\nabla_{u_i}\mathbf{K}_i)]. \quad (36c)$$

From Lemmas 7 and 8, we immediately have the followings:

### Lemma 9

$$\mathbb{E}[\text{tr}\{(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11.2}^{(1)})^{-1}(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})'\}] = \mathbb{E}_G[q \text{tr}(\mathbf{I}_m + \mathbf{Z}_1\mathbf{W}_1^{-1}\mathbf{Z}_1')], \quad (37a)$$

$$\begin{aligned} & \mathbb{E}[\text{tr}\{(\widehat{\boldsymbol{\Theta}}_2 - \mathbf{A}\boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11.2}^{(2)})^{-1}(\widehat{\boldsymbol{\Theta}}_2 - \mathbf{A}\boldsymbol{\Theta})'(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}] \\ & = \mathbb{E}_G[q \text{tr}\{(\mathbf{I}_m + \mathbf{Z}_2\mathbf{W}_2^{-1}\mathbf{Z}_2')(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}], \end{aligned} \quad (37b)$$

$$\begin{aligned} & \mathbb{E}[\text{tr}\{(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11.2}^{(1)})^{-1}\mathbf{B}^{-1}(\mathbf{I}_q - \boldsymbol{\Phi})\mathbf{B}(\mathbf{A}^{-1}\widehat{\boldsymbol{\Theta}}_2 - \widehat{\boldsymbol{\Theta}}_1)'\}] \\ & = \mathbb{E}_G\left[-\left(\sum_{j=1}^q(1 - \phi_j)\right) \text{tr}(\mathbf{I}_m + \mathbf{Z}_1\mathbf{W}_1^{-1}\mathbf{Z}_1')\right], \end{aligned} \quad (37c)$$

$$\begin{aligned} & \mathbb{E}[\text{tr}\{(\widehat{\boldsymbol{\Theta}}_2 - \mathbf{A}\boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11.2}^{(2)})^{-1}\mathbf{B}^{-1}\boldsymbol{\Phi}\mathbf{B}(\mathbf{A}\widehat{\boldsymbol{\Theta}}_1 - \widehat{\boldsymbol{\Theta}}_2)'(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}] \\ & = \mathbb{E}_G\left[-\left(\sum_{j=1}^q\phi_j\right) \text{tr}\{(\mathbf{I}_m + \mathbf{Z}_2\mathbf{W}_2^{-1}\mathbf{Z}_2')(\mathbf{C}\mathbf{A}^{-1})'(\mathbf{C}\mathbf{A}^{-1})\}\right]. \end{aligned} \quad (37d)$$

**Proof.** Note that the density function (21) is symmetric at  $\mathbf{X}_1 - \mathbf{Z}_1\boldsymbol{\gamma}_1 - \boldsymbol{\Theta} = \mathbf{0}$ ,  $\mathbf{X}_2 - \mathbf{Z}_2\boldsymbol{\gamma}_2 - \mathbf{A}\boldsymbol{\Theta} = \mathbf{0}$ , and  $u_i - \mathbf{W}_i^{1/2}\boldsymbol{\gamma}_i = \mathbf{0}$  ( $i = 1, 2$ ).

For (37a), we observe that

$$\begin{aligned}
& \mathbb{E}[\text{tr}\{(\widehat{\Theta}_1 - \Theta)(\Sigma_{11.2}^{(1)})^{-1}(\widehat{\Theta}_1 - \Theta)'\}] \\
&= \mathbb{E}[\text{tr}\{(\mathbf{X}_1 - \mathbf{Z}_1\gamma_1 - \Theta)(\Sigma_{11.2}^{(1)})^{-1}(\mathbf{X}_1 - \mathbf{Z}_1\gamma_1 - \Theta)'\}] \\
&\quad - 2\mathbb{E}[\text{tr}\{\mathbf{Z}_1(\widehat{\gamma}_1 - \gamma_1)(\Sigma_{11.2}^{(1)})^{-1}(\mathbf{X}_1 - \mathbf{Z}_1\gamma_1 - \Theta)'\}] \\
&\quad + \mathbb{E}[\text{tr}\{\mathbf{Z}_1(\widehat{\gamma}_1 - \gamma_1)(\Sigma_{11.2}^{(1)})^{-1}(\widehat{\gamma}_1 - \gamma_1)'\mathbf{Z}_1'\}].
\end{aligned}$$

Here the second term of the right-hand side in the above equation is zero. Hence, from the fact that  $\widehat{\gamma}_1 = \mathbf{W}_1^{-1/2}\mathbf{u}_1$  and Lemma 8, we get the right-hand side of (37a).

By the similar way, we have (37b). For (37c), we can see from symmetry of density function that

$$\begin{aligned}
& \mathbb{E}[\text{tr}\{(\widehat{\Theta}_1 - \Theta)(\Sigma_{11.2}^{(1)})^{-1}\mathbf{B}^{-1}(\mathbf{I}_q - \Phi)\mathbf{B}(\mathbf{A}^{-1}\widehat{\Theta}_2 - \widehat{\Theta}_1)'\}] \\
&= -\mathbb{E}[\text{tr}\{(\widehat{\Theta}_1 - \Theta)(\Sigma_{11.2}^{(1)})^{-1}\mathbf{B}^{-1}(\mathbf{I}_q - \Phi)\mathbf{B}(\widehat{\Theta}_1 - \Theta)'\}] \\
&= -\mathbb{E}[\text{tr}\{(\mathbf{X}_1 - \mathbf{Z}_1\gamma_1 - \Theta)(\Sigma_{11.2}^{(1)})^{-1}\mathbf{B}^{-1}(\mathbf{I}_q - \Phi)\mathbf{B} \\
&\quad \times (\mathbf{X}_1 - \mathbf{Z}_1\gamma_1 - \Theta)'\}] \\
&\quad - \mathbb{E}[\text{tr}\{(\mathbf{u}_1 - \mathbf{W}_1^{1/2}\gamma_1)(\Sigma_{11.2}^{(1)})^{-1}\mathbf{B}^{-1}(\mathbf{I}_q - \Phi)\mathbf{B}(\mathbf{u}_1 - \mathbf{W}_1^{1/2}\gamma_1)' \\
&\quad \times \mathbf{W}_1^{-1/2}\mathbf{Z}_1'\mathbf{Z}_1\mathbf{W}_1^{-1/2}\}].
\end{aligned}$$

Thus, from Lemmas 7 and 8, we get the right-hand side of (37c). The derivation of (37d) is similar to that of (37c).  $\square$

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

$$\{\mathbf{D}_i \mathbf{V}_i\}_{jk} = \sum_{a=1}^p d_{i,ja} v_{i,ak}, \quad i = 1, 2, \quad (38)$$

where

$$d_{i,ja} = \frac{1}{2}(1 + \delta_{ja}) \frac{\partial}{\partial s_{i,ja}}$$

with  $\delta_{ja} = 1$  for  $j = a$  and  $\delta_{ja} = 0$  for  $j \neq a$ . Also put  $\mathbf{S}_i = (\mathbf{s}'_{i1}, \dots, \mathbf{s}'_{in_i})'$  and  $\mathbf{s}_{ij} = (s_{i,j1}, \dots, s_{i,jp})$  for  $i = 1, 2$  and  $j = 1, 2, \dots, n_i$ . Hence we have  $\mathbf{S}_i = \mathbf{s}'_i \mathbf{s}_i = \sum_{j=1}^{n_i} \mathbf{s}'_{ij} \mathbf{s}_{ij}$  for  $i = 1, 2$ .

**Lemma 10 (Kubokawa and Srivastava, 2000)** *Let*

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

*be  $p \times p$  matrices whose elements are differentiable with respect to  $s_{i,jk}$  ( $j = 1, 2, \dots, n_i, k = 1, 2, \dots, p$ ). Furthermore, assume that*

- (a)  $\mathbb{E} \left[ \left| \text{tr} \mathbf{V}_i \boldsymbol{\Sigma}_{11.2}^{(i)-1} \right| \right]$  ( $i = 1, 2$ ) is finite;
- (b)  $\lim_{s_{i,jk} \rightarrow \pm\infty} |s_{i,jk}| \mathbf{V}_i \cdot \left( \sum_{j=1}^{n_i} \mathbf{s}'_{1j_i} \mathbf{s}_{1j_i} \right)^{-1} G(s_{i,jk}^2 + a) = \mathbf{0}$  for any real  $a$ .

Then we have

$$\begin{aligned} & \mathbb{E} \left[ \sum_{i=1}^2 \text{tr} (\{\boldsymbol{\Sigma}_{11.2}^{(i)}\}^{-1} \mathbf{V}_i) \right] \\ &= \mathbb{E}_G \left[ \sum_{i=1}^2 \left\{ (n_i - q - 1) \text{tr} (\mathbf{S}_i^{-1} \mathbf{V}_i) + 2 \text{tr} (\mathcal{D}_i \mathbf{V}_i) \right\} \right], \end{aligned}$$

where  $n_i = N_i - m - p_i + q$ .

**Proof of Theorem 3.** The proof proceeds much the same way as in that of Theorem 2. Recall that the risk of the estimators of the form (22) can be written as (16) where the expectation is taken with respect to the density (21). Now first apply Lemmas 9 and 10 to the risk (16) and next use Lemma 6 to get the desired result.  $\square$

## 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] Brown, L.D. and Cohen, A. (1974). Point and confidence estimation of a common mean and recovery of interblock information. *Ann. Statist.* **2**, 963–976.
- [4] Chiou W. and Cohen, A. (1985). On estimating a common multivariate normal mean vector. *Ann. Inst. Statist. Math.* **37**, 499–506.
- [5] Cohen, A. and Sackrowitz, H.B. (1974). On estimating the common mean of two normal distributions. *Ann. Statist.* **2**, 1274–1282.
- [6] Fang, K.T., Kotz, S., and Ng, K.W. (1990). *Symmetric multivariate and related distributions*. Chapman and Hall, New York.
- [7] Fang, K.T. and Zhang, Y.T. (1990). *Generalized Multivariate Analysis*. Springer-Verlag, Berlin and Science Press, Beijing.

- [8] Graybill, F.A. and Deal, R.B. (1959). Combined unbiased estimator. *Biometrics* **15**, 543–550.
- [9] Haff, L.R. (1991). The variational forms of certain Bayes estimators. *Ann. Statist.* **18**, 1163–1190.
- [10] James, W. and Stein, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1**, 361–380. Univ. of California Press.
- [11] Johnson, M. (1987). *Multivariate Statistical Simulation*. John Wiley and Sons, New York.
- [12] Kariya, T. (1989). Equivariant estimation in a model with an ancillary statistic. *Ann. Statist.* **17** 920–928.
- [13] Kariya, T., Konno, Y., and Strawderman, W.E. (1996). Double shrinkage estimators in the GMANOVA model. *J. Multivariate Anal.* **56**, 245–258.
- [14] 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.
- [15] Khatri, C.G. and Shah, K.R. (1974). Estimation of location parameters from two linear models under normality. *Comm. Statist. A-Theory and Methods*, **3**, 647–663.
- [16] Kubokawa, T. (1989). *Estimating common parameters of growth curve models under a quadratic loss*. *Comm. Statist. A-Theory and Methods* **18**, 3149–3155.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] Loh, W.L. (1988). Estimating covariance matrices. Ph. D. dissertation, Dept. Statist., Stanford Univ.
- [21] Loh, W.L. (1991). Estimating the common mean of two multivariate normal distributions. *Ann. Statist.* **19**, 297–313.
- [22] Sugiura, N. and Kubokawa, T. (1988). *Estimating common parameters of growth curve models*. *Ann. Inst. Statist. Math.* **40**, 119–135.
- [23] Stein, C. (1975). Estimation of a covariance matrix. Rietz Lecture, 39th Annual IMS Meeting, Atlanta. Unpublished manuscript.

- [24] 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.).