数理解析研究所講究録 1334 巻 2003 年 95-111

Modifying the Graybill-Deal estimator of the common regression matrix in two growth curve models

千葉大学大学院自然科学研究科　津熊　久幸 (Hisayuki Tsukuma)
Graduate School of Science and Technology
Chiba University
千葉大学大学院自然科学研究科　今野　良彦 (Yoshihiko Konno)
Graduate School of Science and Technology
Chiba University

Abstract

We consider the problem of estimating the common regression matrix of two growth curve 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 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 an alternative combined estimator to the Graybill-Deal type estimator. Finally, we conduct the Monte-Carlo simulation to show that our proposed estimator performs better than the Graybill-Deal type estimator.

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), Shinozaki (1978), Chiou and Cohen (1985), and Loh (1991). Of these, Graybill and Deal (1959) first showed that the Graybill-Deal estimator, a combined estimator for the common mean of two univariate normal distributions, has smaller variance than either of each sample mean when the sample size is at least eleven.

This paper is mainly concerned with estimating the common regression matrix of two growth curve 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 growth curve models. In the present paper we propose an alternative to the estimator of Sugiura and Kubokawa in a decision-theoretic point of view. 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 two growth curve models

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

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 $p_i \geq q$, $\boldsymbol{\Xi}$ is an $m \times q$ matrix of unknown parameters, and $\boldsymbol{\epsilon}_i$ are $N_i \times p_i$ error matrices which are independently distributed as the multivariate normal distributions with the covariance matrices $\boldsymbol{I}_{N_1} \otimes \boldsymbol{\Omega}_1$ and $\boldsymbol{I}_{N_2} \otimes \boldsymbol{\Omega}_2$, 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 $\boldsymbol{\Omega}_i$. Here we assume that the $\boldsymbol{\Omega}_i$'s are unknown positive definite $p_i \times p_i$ matrices. In the sequel of the paper we use notation $\boldsymbol{B}'$, $|\boldsymbol{B}|$, and $\mathrm{tr}\,(\boldsymbol{B})$, and $(\boldsymbol{B})^{1/2}$ which stand for the transpose, determinant, trace, and a non-negative definite square root 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 and multivariate mixed linear models treated in Kubokawa and Srivastava (2002).

We consider the problem of estimating Ξ under the loss function

$$\begin{aligned}\widetilde{\boldsymbol{L}}((\Xi,\,\boldsymbol{\Omega}_1,\,\boldsymbol{\Omega}_2),\,\widehat{\Xi}) &= \mathrm{tr}\,\{\boldsymbol{A}_{11}(\widehat{\Xi}-\Xi)\boldsymbol{A}_{12}\boldsymbol{\Omega}_1^{-1}\boldsymbol{A}_{12}'(\widehat{\Xi}-\Xi)'\boldsymbol{A}_{11}'\} \\ &\quad + \mathrm{tr}\,\{\widetilde{\boldsymbol{C}}(\widehat{\Xi}-\Xi)\boldsymbol{A}_{22}\boldsymbol{\Omega}_2^{-1}\boldsymbol{A}_{22}'(\widehat{\Xi}-\Xi)'\widetilde{\boldsymbol{C}}'\},\end{aligned} \tag{2}$$

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 growth curve 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 growth curve 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 derive a canonical form of two sample problem of estimating the common regression matrix of the growth curve models and we give 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 carry out Monte-Carlo simulation to show that our proposed estimator reduces the risk substantially over the Graybill-Deal type estimator when we observe the data $(\boldsymbol{Y}_1,\,\boldsymbol{Y}_2)$ from the model (1). In Section 4, we give technical lemmas and the proof of the main result.

2 Derivation of alternative estimators

2.1 A canonical form

Recall that we observe random matrices $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ which are independently distributed as

$$\boldsymbol{Y}_i \sim N_{N_i \times p_i}(\boldsymbol{A}_{i1}\Xi\boldsymbol{A}_{i2},\boldsymbol{I}_{N_i} \otimes \boldsymbol{\Omega}_i), \quad i = 1, 2, \tag{3}$$

that is, the rows of $\boldsymbol{Y}_i - \boldsymbol{A}_{i1}\Xi\boldsymbol{A}_{i2}$ are distributed as p_i-variate normal distribution with the zero mean and the covariance matrix $\boldsymbol{\Omega}_i$. To derive a canonical form of (3), let $\boldsymbol{\Gamma}_i$, i = 1, 2, be $N_i \times N_i$ orthogonal matrices such that $\boldsymbol{\Gamma}_i\boldsymbol{A}_{i1} = [(\boldsymbol{A}_{i1}'\boldsymbol{A}_{i1})^{1/2}, \boldsymbol{0}_{m\times(N_i-m)}]'$ and also let $\boldsymbol{\Upsilon}_i$ be $p_i \times p_i$ orthogonal matrices such that $\boldsymbol{A}_{i2}\boldsymbol{\Upsilon}_i = [(\boldsymbol{A}_{i2}\boldsymbol{A}_{i2}')^{1/2}, \boldsymbol{0}_{q\times(p_i-q)}]$.

Furthermore we write

$$
\begin{aligned}
\boldsymbol{\Theta} &= (\boldsymbol{A}_{11}'\boldsymbol{A}_{11})^{1/2}\boldsymbol{\Xi}(\boldsymbol{A}_{12}\boldsymbol{A}_{12}')^{1/2}, && \text{(4a)}\\
\boldsymbol{A} &= (\boldsymbol{A}_{21}'\boldsymbol{A}_{21})^{1/2}(\boldsymbol{A}_{11}'\boldsymbol{A}_{11})^{-1/2}, && \text{(4b)}\\
\boldsymbol{\Lambda} &= \begin{pmatrix} (\boldsymbol{A}_{22}\boldsymbol{A}_{22}')^{-1/2}(\boldsymbol{A}_{12}\boldsymbol{A}_{12}')^{1/2} & \boldsymbol{0}_{q\times(p_2-q)} \\ \boldsymbol{0}_{(p_2-q)\times q} & \boldsymbol{I}_{p_2-q} \end{pmatrix}, && \text{(4c)}\\
\boldsymbol{\Sigma}_1 &= \boldsymbol{\Upsilon}_1'\boldsymbol{\Omega}_1\boldsymbol{\Upsilon}_1 = \begin{pmatrix} \boldsymbol{\Sigma}_{11}^{(1)} & \boldsymbol{\Sigma}_{12}^{(1)} \\ \boldsymbol{\Sigma}_{21}^{(1)} & \boldsymbol{\Sigma}_{22}^{(1)} \end{pmatrix}, && \text{(4d)}\\
\boldsymbol{\Sigma}_2 &= \boldsymbol{\Lambda}'\boldsymbol{\Upsilon}_2'\boldsymbol{\Omega}_2\boldsymbol{\Upsilon}_2\boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{\Sigma}_{11}^{(2)} & \boldsymbol{\Sigma}_{12}^{(2)} \\ \boldsymbol{\Sigma}_{21}^{(2)} & \boldsymbol{\Sigma}_{22}^{(2)} \end{pmatrix}, && \text{(4e)}
\end{aligned}
$$

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

$$
\begin{aligned}
\boldsymbol{X}_1 \mid \boldsymbol{Z}_1 &\sim N_{m\times q}(\boldsymbol{\Theta} + \boldsymbol{Z}_1\boldsymbol{\gamma}_1, \boldsymbol{I}_m \otimes \boldsymbol{\Sigma}_{11\cdot 2}^{(1)}), && \text{(5a)}\\
\boldsymbol{X}_2 \mid \boldsymbol{Z}_2 &\sim N_{m\times q}(\boldsymbol{A}\boldsymbol{\Theta} + \boldsymbol{Z}_2\boldsymbol{\gamma}_2, \boldsymbol{I}_m \otimes \boldsymbol{\Sigma}_{11\cdot 2}^{(2)}) && \text{(5b)}
\end{aligned}
$$

and, for $i = 1, 2$,

$$
\begin{aligned}
\boldsymbol{Z}_i &\sim N_{m\times(p_i-q)}(\boldsymbol{0}, \boldsymbol{I}_m \otimes \boldsymbol{\Sigma}_{22}^{(i)}), && \text{(6a)}\\
\boldsymbol{S}_i &\sim W_q(\boldsymbol{\Sigma}_{11\cdot 2}^{(i)},\ n_i),\ n_i = N_i - m - p_i + q, && \text{(6b)}\\
\widehat{\boldsymbol{\gamma}}_i \mid \boldsymbol{W}_i &\sim N_{(p_i-q)\times q}(\boldsymbol{\gamma}_i,\ \boldsymbol{W}_i^{-1} \otimes \boldsymbol{\Sigma}_{11\cdot 2}^{(i)}), && \text{(6c)}\\
\boldsymbol{W}_i &\sim W_{p_i-q}(\boldsymbol{\Sigma}_{22}^{(i)}, n_i + p_i - q), && \text{(6d)}
\end{aligned}
$$

where $\boldsymbol{\Sigma}_{11\cdot 2}^{(i)} = \boldsymbol{\Sigma}_{11}^{(i)} - \boldsymbol{\Sigma}_{12}^{(i)}(\boldsymbol{\Sigma}_{22}^{(i)})^{-1}\boldsymbol{\Sigma}_{21}^{(i)}$ and $\boldsymbol{\gamma}_i = (\boldsymbol{\Sigma}_{22}^{(i)})^{-1}\boldsymbol{\Sigma}_{21}^{(i)}$. Here, note that $\boldsymbol{A}$ is an $m \times m$ known nonsingular matrix and that $(\boldsymbol{X}_i,\ \boldsymbol{Z}_i)$, $(\boldsymbol{W}_i,\ \widehat{\boldsymbol{\gamma}}_i)$, and $\boldsymbol{S}_i$ are independent.

Furthermore, the loss function (2) turns into

$$
\begin{aligned}
\boldsymbol{L}((\boldsymbol{\Theta},\ \boldsymbol{\Sigma}_1,\ \boldsymbol{\Sigma}_2),\ \widehat{\boldsymbol{\Theta}}) &= \operatorname{tr}[(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot 2}^{(1)})^{-1}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})'] \\
&\quad + \operatorname{tr}[\boldsymbol{C}'\boldsymbol{C}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot 2}^{(2)})^{-1}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})'], \qquad (7)
\end{aligned}
$$

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 $(\boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{S}_i,\ \widehat{\boldsymbol{\gamma}}_i,\ \boldsymbol{W}_i \mid i = 1, 2)$ under the loss function (7). Then the risk function is defined by

$$
\boldsymbol{R}((\boldsymbol{\Theta},\ \boldsymbol{\Sigma}_1,\ \boldsymbol{\Sigma}_2),\ \widehat{\boldsymbol{\Theta}}) = \mathbb{E}[\boldsymbol{L}((\boldsymbol{\Theta},\ \boldsymbol{\Sigma}_1,\ \boldsymbol{\Sigma}_2),\ \widehat{\boldsymbol{\Theta}})], \qquad (8)
$$

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

2.2 An equivariant estimator of Θ

Next we derive a class of estimators of $\boldsymbol{\Theta}$. 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 $m\times q$ 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 the $\boldsymbol{P}_{i\cdot 22}$'s are $q\times q$ and $(p_i-q)\times(p_i-q)$ nonsingular matrices, respectively, and the $\boldsymbol{P}_{i\cdot 12}$'s 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 estimation problem of the common regression matrix in two growth curve 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. The action of $(\boldsymbol{D},\ \boldsymbol{P}_1,\ \boldsymbol{P}_2)$ on $(\boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{S}_i,\ \widehat{\boldsymbol{\gamma}}_i,\ \boldsymbol{W}_i|\, i=1,\ 2)$ is defined as

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

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

$$\widehat{\boldsymbol{\Theta}}_i = \boldsymbol{X}_i - \boldsymbol{Z}_i\widehat{\boldsymbol{\gamma}}_i, \qquad i=1,\ 2. \tag{9}$$

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

$$\begin{aligned}
&\boldsymbol{\Theta}\to\boldsymbol{\Theta}\boldsymbol{P}_{11}'+\boldsymbol{D},\\
&(\boldsymbol{\Sigma}_{11\cdot 2}^{(i)},\ \boldsymbol{\Sigma}_{22}^{(i)},\ (\boldsymbol{\Sigma}_{22}^{(i)})^{-1}\boldsymbol{\Sigma}_{21}^{(i)})\\
&\to(\boldsymbol{P}_{11}\boldsymbol{\Sigma}_{11\cdot 2}^{(i)}\boldsymbol{P}_{11}',\ \boldsymbol{P}_{i\cdot 22}\boldsymbol{\Sigma}_{22}^{(i)}\boldsymbol{P}_{i\cdot 22}',\ (\boldsymbol{P}_{i\cdot 22}')^{-1}(\boldsymbol{\Sigma}_{22}^{(i)})^{-1}\boldsymbol{\Sigma}_{21}^{(i)}\boldsymbol{P}_{11}'+(\boldsymbol{P}_{i\cdot 22}')^{-1}\boldsymbol{P}_{i\cdot 12}'),\\
&(\widehat{\boldsymbol{\Theta}}_1,\ \boldsymbol{Z}_1,\ \widehat{\boldsymbol{\Theta}}_2,\ \boldsymbol{Z}_2)\\
&\to(\widehat{\boldsymbol{\Theta}}_1\boldsymbol{P}_{11}'+\boldsymbol{D},\ \boldsymbol{Z}_1\boldsymbol{P}_{1\cdot 22}',\ \widehat{\boldsymbol{\Theta}}_2\boldsymbol{P}_{11}'+\boldsymbol{A}\boldsymbol{D},\ \boldsymbol{Z}_2\boldsymbol{P}_{2\cdot 22}'),\\
&(\boldsymbol{S}_i,\ \boldsymbol{W}_i,\ \widehat{\boldsymbol{\gamma}}_i)\\
&\to(\boldsymbol{P}_{11}\boldsymbol{S}_i\boldsymbol{P}_{11}',\ \boldsymbol{P}_{i\cdot 22}\boldsymbol{W}_i\boldsymbol{P}_{i\cdot 22}',\ (\boldsymbol{P}_{i\cdot 22}')^{-1}\widehat{\boldsymbol{\gamma}}_i\boldsymbol{P}_{11}'+(\boldsymbol{P}_{i\cdot 22}')^{-1}\boldsymbol{P}_{i\cdot 12}')
\end{aligned}$$

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

$$\widehat{\boldsymbol{\Theta}}^{EQI}(g\circ(\boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{S}_i,\ \widehat{\boldsymbol{\gamma}}_i,\ \boldsymbol{W}_i|\, i=1,\ 2)) = \widehat{\boldsymbol{\Theta}}^{EQI}(\boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{S}_i,\ \boldsymbol{W}_i,\ \widehat{\boldsymbol{\gamma}}_i|\, i=1,\ 2)\boldsymbol{P}_{11}'+\boldsymbol{D},$$

so that $\widehat{\boldsymbol{\Theta}}^{EQI}(g \circ (\boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{S}_i,\ \widehat{\gamma}_i |\, i = 1,\ 2))$ estimates the parameter $\boldsymbol{\Theta P}'_{11} + \boldsymbol{D}$ as does $\widehat{\boldsymbol{\Theta}}^{EQI}(\boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{S}_i,\ \boldsymbol{W}_i,\ \widehat{\gamma}_i |\, i = 1,\ 2)\boldsymbol{P}'_{11} + \boldsymbol{D}$. 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 $\boldsymbol{BS}_2\boldsymbol{B}' = \boldsymbol{F}$ and $f_1 \geq \cdots \geq f_q \geq 0$. Then under the group of transformations, an equivariant estimator of $\boldsymbol{\Theta}^{EQI}$ 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 and $\widehat{\boldsymbol{\Theta}}_i$, $i = 1,\ 2$, are given by (9).

Proof. The proof of this theorem can be obtained similarly as in that of Theorem 4.1 in Loh (1988). □

Since the class of the equivariant estimators (10) is too large to evaluate their risk systematically, we restrict ourselves to an equivariant estimator (10) where $\tilde{\boldsymbol{\Phi}}$ is a diagonal matrix and depends only on $\boldsymbol{F}$, i.e.,

$$\widehat{\boldsymbol{\Theta}}^{EQ} = \widehat{\boldsymbol{\Theta}}_1\boldsymbol{B}'\boldsymbol{\Phi}(\boldsymbol{B}')^{-1} + \boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2\boldsymbol{B}'(\boldsymbol{I}_q - \boldsymbol{\Phi})(\boldsymbol{B}')^{-1}, \tag{11}$$

where $\widehat{\boldsymbol{\Theta}}_i$, $i = 1,\ 2$, are given by (9) and $\boldsymbol{\Phi} = \boldsymbol{\Phi}(\boldsymbol{F})$ is a $q \times q$ diagonal matrix with diagonal elements $\phi_i(\boldsymbol{F})$, $i = 1,\ 2,\ \ldots,\ q$. Here we assume that $\phi_i(\boldsymbol{F})$ depends 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 $\boldsymbol{S}_2(\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1}$.

Remark 1. Since $\boldsymbol{S}_i$ is independent of $\boldsymbol{X}_i$, $\boldsymbol{Z}_i$ and $\widehat{\gamma}_i$ for i $= 1, 2$, we can see that $\boldsymbol{B}$ and $\boldsymbol{F}$ are independent of $\widehat{\boldsymbol{\Theta}}_1$ and $\widehat{\boldsymbol{\Theta}}_2$. Therefore we have

$$\begin{aligned}\mathbb{E}[\widehat{\boldsymbol{\Theta}}^{EQ}] &= \mathbb{E}[\mathbb{E}[\widehat{\boldsymbol{\Theta}}_1\boldsymbol{B}'\boldsymbol{\Phi}(\boldsymbol{B}')^{-1} + \boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2\boldsymbol{B}'(\boldsymbol{I}_q - \boldsymbol{\Phi})(\boldsymbol{B}')^{-1} |\, (\boldsymbol{B},\ \boldsymbol{F})]] \\ &= \mathbb{E}[\boldsymbol{\Theta B}'\boldsymbol{\Phi}(\boldsymbol{B}')^{-1} + \boldsymbol{\Theta B}'(\boldsymbol{I}_q - \boldsymbol{\Phi})(\boldsymbol{B}')^{-1}] = \boldsymbol{\Theta},\end{aligned}$$

which shows that $\widehat{\boldsymbol{\Theta}}^{EQ}$ is an unbiased estimator of $\boldsymbol{\Theta}$.

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 $(4a) - (4e)$, we can see that the estimator of Sugiura and Kubokawa is rewritten as

$$\begin{aligned}\mathrm{vec}(\widehat{\boldsymbol{\Theta}}^{GDI}) &= \{\boldsymbol{I}_m \otimes (\boldsymbol{S}_1/n_1)^{-1} + (\boldsymbol{A}'\boldsymbol{A}) \otimes (\boldsymbol{S}_2/n_2)^{-1}\}^{-1} \\ &\quad \times\{\boldsymbol{I}_m \otimes (\boldsymbol{S}_1/n_1)^{-1}\mathrm{vec}(\widehat{\boldsymbol{\Theta}}_1) + (\boldsymbol{A}'\boldsymbol{A}) \otimes (\boldsymbol{S}_2/n_2)^{-1}\mathrm{vec}(\boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2)\},\end{aligned} \tag{12}$$

where we denote by $\mathrm{vec}(\boldsymbol{U})$ an $mq \times 1$ vector consisting of $(u_1, u_2, \ldots, u_m)'$ for $\boldsymbol{U} = (u_1', u_2', \ldots, u_m')'$ 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 $\boldsymbol{G} = (g_{ij})$. On the other hand, we can rewrite the estimator (11) as

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

if we put $\phi_j = \beta_j/(\alpha_j + \beta_j)$, $j = 1, 2, \ldots, q$, where α_j and β_j are real-valued functions of $\boldsymbol{F}$. Here we denote by $\mathbf{diag}(\beta_j)$ a $q \times q$ diagonal matrix whose j-th diagonal elements are given by β_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 (11) reduces to

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

equivalently

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

The estimator (13) can be regarded as a counterpart of the Graybill-Deal type estimator (12) inside the class of equivariant estimators of the form (11). It is well known that the eigenvalues of $\boldsymbol{S}_2(\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1}$ are more spread than the eigenvalues of 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 $\boldsymbol{S}_2(\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1}$.

Through these consideration, we use the following scenario to obtain an alternative estimator to (12). First we look into the class of equivariant estimators of the form (11) and obtain an alternative estimator which has the form (11). Then we change the term $I_m \otimes (\boldsymbol{B}'\mathbf{diag}(\alpha_j)\boldsymbol{B})$ in (13) into $(\boldsymbol{A}'\boldsymbol{A}) \otimes (\boldsymbol{B}'\mathbf{diag}(\alpha_j)\boldsymbol{B})$ to get an alternative estimator which is in the class of the estimators (12).

2.4 A subclass of equivariant estimators and its risk

To obtain alternative estimator of the form (11), we evaluate its risk in terms of unbiased risk method due to Stein-Haff-Loh. The risk of the estimator of the form (11) can be written as

$$\begin{aligned} &\boldsymbol{R}((\boldsymbol{\Theta}, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2), \widehat{\boldsymbol{\Theta}}) \\ &= \mathbb{E}\Big[\mathrm{tr}\,\{(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot2}^{(1)})^{-1}(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})'\} \\ &\quad +2\,\mathrm{tr}\,\{(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot2}^{(1)})^{-1}\boldsymbol{B}^{-1}(I_q - \boldsymbol{\Phi})\boldsymbol{B}(\boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2 - \widehat{\boldsymbol{\Theta}}_1)'\} \\ &\quad +\mathrm{tr}\,\{(\boldsymbol{\Sigma}_{11\cdot2}^{(1)})^{-1}\boldsymbol{B}^{-1}(\boldsymbol{I}_q - \boldsymbol{\Phi})\boldsymbol{H}_1(\boldsymbol{I}_q - \boldsymbol{\Phi})(\boldsymbol{B}')^{-1}\} \\ &\quad +\mathrm{tr}\,\{(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})(\widehat{\boldsymbol{\Theta}}_2 - \boldsymbol{A}\boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot2}^{(2)})^{-1}(\widehat{\boldsymbol{\Theta}}_2 - \boldsymbol{A}\boldsymbol{\Theta})'\} \end{aligned}$$

$$+2\,\mathrm{tr}\,\{(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})(\widehat{\boldsymbol{\Theta}}_2 - \boldsymbol{A}\boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot 2}^{(2)})^{-1}\boldsymbol{B}^{-1}\boldsymbol{\Phi}\boldsymbol{B}(\boldsymbol{A}\widehat{\boldsymbol{\Theta}}_1 - \widehat{\boldsymbol{\Theta}}_2)'\}$$
$$+\,\mathrm{tr}\,\{(\boldsymbol{\Sigma}_{11\cdot 2}^{(2)})^{-1}\boldsymbol{B}^{-1}\boldsymbol{\Phi}\boldsymbol{H}_2\boldsymbol{\Phi}(\boldsymbol{B}')^{-1}\}\Big], \tag{14}$$

where

$$\boldsymbol{H}_1 = \boldsymbol{B}(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2)'(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2)\boldsymbol{B}', \tag{15a}$$
$$\boldsymbol{H}_2 = \boldsymbol{B}(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2)'(\boldsymbol{C}'\boldsymbol{C})(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2)\boldsymbol{B}'. \tag{15b}$$

Now we use the Haff-Stein identity for Wishart distribution and calculation on eigen-structure technique due to Stein (1975, 1977), Haff (1991), and Loh (1988) to evaluate the third and sixth terms in right-hand side of (14) while we use formula for the second moments of the maximum likelihood estimator of the growth curve model to evaluate the other terms in right-hand side of (14). Then we obtain an unbiased estimate of risk for the equivariant estimators (11). Since the proof of Theorem 2, the main result of the paper, involves in technical argument, we postpone it until Section 4.

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

$$\boldsymbol{R}((\boldsymbol{\Theta}, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2), \widehat{\boldsymbol{\Theta}}^{EQ})$$
$$= \mathbb{E}\Bigg[q(r_2 - r_1) + \sum_{j=1}^{q}\Bigg\{2(r_1 - r_2)\phi_j + (n_1 - q - 1)\frac{(1-\phi_j)^2}{1-f_j}\{\boldsymbol{H}_1\}_{jj}$$
$$+4\{\boldsymbol{H}_1\}_{jj}(1-\phi_j)f_j\frac{\partial\phi_j}{\partial f_j} + 2\sum_{k\neq j}\{\boldsymbol{H}_1\}_{jj}(1-\phi_j)(\phi_j - \phi_k)\frac{f_k}{f_j - f_k}$$
$$+(n_2 - q - 1)\frac{\phi_j^2}{f_j}\{\boldsymbol{H}_2\}_{jj} + 4\{\boldsymbol{H}_2\}_{jj}\phi_j(1-f_j)\frac{\partial\phi_j}{\partial f_j}$$
$$+2\sum_{k\neq j}\{\boldsymbol{H}_2\}_{jj}\phi_j(\phi_j - \phi_k)\frac{1-f_k}{f_j - f_k}\Bigg\}\Bigg], \tag{16}$$

where $r_1 = m(n_1+p_1-q-1)/(n_1-1)$, $r_2 = \{(n_2+p_2-q-1)/(n_2-1)\}\,\mathrm{tr}\,\{(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})\}$, *and* $\{\boldsymbol{H}_1\}_{jj}$ *and* $\{\boldsymbol{H}_2\}_{jj}$ *are j-th diagonal elements of the matrices given by* (15*a*) *and* (15*b*), *respectively.*

2.5 Choice of $\boldsymbol{\Phi}$

From Theorem 2, we obtain the unbiased estimate of the risk of the subclass of equivariant estimators given by (11). We denote by $\widehat{\boldsymbol{R}}$ the unbiased estimate of the risk, i.e., the terms inside large bracket in the right-hand side of (16). Although we obtain the unbiased estimate of risk for the class of estimators given by (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 estimate of the risk from the unbiased estimate of the risk. The derivation below to give alternative estimators is no longer mathematically rigorous. However we believe that the following argument results in promising estimators which perform well in our simulation study.

First we replace $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ in (16) by their approximation. To this end, we observe that

$$\begin{aligned}
&\mathbb{E}[(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2)'(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2)] \\
&\qquad = m\tilde{r}_1\boldsymbol{\Sigma}_{11\cdot 2}^{(1)} + \tilde{r}_2 \operatorname{tr}\{(\boldsymbol{A}')^{-1}\boldsymbol{A}^{-1}\}\boldsymbol{\Sigma}_{11\cdot 2}^{(2)}, \\
&\mathbb{E}[(\boldsymbol{A}\widehat{\boldsymbol{\Theta}}_1 - \widehat{\boldsymbol{\Theta}}_2)'(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})(\boldsymbol{A}\widehat{\boldsymbol{\Theta}}_1 - \widehat{\boldsymbol{\Theta}}_2)] \\
&\qquad = \tilde{r}_1 \operatorname{tr}(\boldsymbol{C}'\boldsymbol{C})\boldsymbol{\Sigma}_{11\cdot 2}^{(1)} + \tilde{r}_2 \operatorname{tr}\{(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})\}\boldsymbol{\Sigma}_{11\cdot 2}^{(2)},
\end{aligned}$$

where $\tilde{r}_i = (n_i + p_i - q - 1)/(n_i - 1)$. Replacing $\boldsymbol{\Sigma}_{11\cdot 2}^{(i)}$, $i = 1, 2$, in right-hand side of the above equations with their maximum likelihood estimators $\boldsymbol{S}_i/n_i$, we approximate $\{\boldsymbol{H}_1\}_{jj}$ and $\{\boldsymbol{H}_2\}_{jj}$, $j = 1, 2, \ldots, q$, by

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

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

$$\frac{\partial \phi_j}{\partial f_j} = f_j \frac{\partial}{\partial f_j}\left(\frac{\phi_j}{f_j}\right) + \frac{\phi_j}{f_j} = (1 - f_j)\frac{\partial}{\partial(1 - f_j)}\left(\frac{1 - \phi_j}{1 - f_j}\right) + \frac{1 - \phi_j}{1 - f_j}, \tag{17}$$

we can see that the unbiased estimate for risk of $\widehat{\boldsymbol{\Theta}}^{EQ}$ given by (16) is approximated by

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

Ignoring the derivative terms, we get

$$
\begin{aligned}
\widehat{\boldsymbol{R}} \;\approx\;& q(r_2-r_1)+\sum_{j=1}^{q}\Big\{2(r_1-r_2)\phi_j \\
&+(n_1-q-1)\frac{(1-\phi_j)^2}{1-f_j}h_{1j}+4h_{1j}(1-\phi_j)^2\frac{f_j}{1-f_j} \\
&+2\sum_{k\neq j}h_{1j}(1-\phi_j)(\phi_j-\phi_k)\frac{f_k}{f_j-f_k} \\
&+(n_2-q-1)\frac{\phi_j^2}{f_j}h_{2j}+4h_{2j}\phi_j^2\frac{1-f_j}{f_j}+2\sum_{k\neq j}h_{2j}\phi_j(\phi_j-\phi_k)\frac{1-f_k}{f_j-f_k}\Big\}
\end{aligned}
$$

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

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

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

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

$$
\phi_j^{ST}=\frac{\hat{\beta}_j^{ST}}{\hat{\beta}_j^{ST}+\hat{\alpha}_j^{ST}}, \tag{18}
$$

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

Consequently we propose an estimator of the form

$$
\begin{aligned}
\mathrm{vec}\,(\widehat{\boldsymbol{\Theta}}^{STI}) &= \{\boldsymbol{I}_m\otimes(\boldsymbol{B}'\mathbf{diag}(\hat{\beta}_j^{ST})\boldsymbol{B})+(\boldsymbol{A}'\boldsymbol{A})\otimes(\boldsymbol{B}'\mathbf{diag}(\hat{\alpha}_j^{ST})\boldsymbol{B})\}^{-1}\\
&\quad\times\{\boldsymbol{I}_m\otimes(\boldsymbol{B}'\mathbf{diag}(\hat{\beta}_j^{ST})\boldsymbol{B})\,\mathrm{vec}\,(\widehat{\boldsymbol{\Theta}}_1)\\
&\quad+(\boldsymbol{A}'\boldsymbol{A})\otimes(\boldsymbol{B}'\mathbf{diag}(\hat{\alpha}_j^{ST})\boldsymbol{B})\mathrm{vec}(\boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2)\},
\end{aligned}
\tag{19}
$$

with (18). Because of complex nature of the estimation problem, we can not carry out analytic comparison between the Graybill-Deal type estimator (12) and our proposed estimator. However, we justify our proposed estimator via simulation study in Section 3.

Remark 2. From similar argument as in Remark 1, we can also see that the estimator (19) is unbiased estimator of $\boldsymbol{\Theta}$.

Remark 3. For the special case, the estimator (19) reduces a simple form. When $\boldsymbol{C}'\boldsymbol{C}=\boldsymbol{A}'\boldsymbol{A}$, $N_1=N_2$ and $p_1=p_2$, we have $r_1=r_2$. This case generalizes the results obtained by Loh (1991). When $\boldsymbol{C}'\boldsymbol{C}=\boldsymbol{I}_m$, we have $h_{1j}=h_{2j}$, $j=1,\dots,q$.

3 Numerical studies

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 (5a)–(6d) with special cases for (N_1, N_2, p_1, p_2, m, q). These results are given in Tables 1–3.

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

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

Note that the matrices A_{11} and A_{21} given above are obtained from some examples of monotone missing data patterns of one-sample growth curve model. Also note that we consider only the case $p_1 > p_2$ since this restriction can be naturally obtained from the monotone missing data set-up of one-sample growth curve model.

For $(\Sigma_{11\cdot 2}^{(1)}, \Sigma_{11\cdot 2}^{(2)})$, we assume that the eigenvalues of $\Sigma_{11\cdot 2}^{(2)}(\Sigma_{11\cdot 2}^{(1)})^{-1}$ are close together and that these eigenvalues are widely spread out. Furthermore, we put $\Theta = \mathbf{0}_{m\times q}$, $\gamma_i = \mathbf{0}_{(p_i-q)\times q}$, and $\Sigma_{22}^{(i)} = I_{p_i-q}$.

Recall that, when (Σ_1, Σ_2) are known, the maximum likelihood estimator of Θ in (5a) and (5b) is given by

$$\begin{aligned} \mathrm{vec}(\widetilde{\Theta}^{ML}) &= [I_m \otimes (\Sigma_{11\cdot 2}^{(1)})^{-1} + A'A \otimes (\Sigma_{11\cdot 2}^{(2)})^{-1}]^{-1} \\ &\quad \times [\{I_m \otimes (\Sigma_{11\cdot 2}^{(1)})^{-1}\}\mathrm{vec}(\widetilde{\Theta}_1) + \{A'A \otimes (\Sigma_{11\cdot 2}^{(2)})^{-1}\}\mathrm{vec}(A^{-1}\widetilde{\Theta}_2)], \end{aligned} \tag{20}$$

where $\widetilde{\Theta}_i = X_i - \gamma_i Z_i \, (i = 1, 2)$. Here the risk of $\mathrm{vec}(\widetilde{\Theta}^{ML})$ is evaluated as follows:

Lemma 1 *We have*

$$\begin{aligned} &R((\Theta, \Sigma_1, \Sigma_2), \widetilde{\Theta}^{ML}) \\ &\quad = \mathrm{tr}\,\{[I_m \otimes (\Sigma_{11\cdot 2}^{(1)})^{-1} + (C'C) \otimes (\Sigma_{11\cdot 2}^{(2)})^{-1}] \\ &\qquad \times [I_m \otimes (\Sigma_{11\cdot 2}^{(1)})^{-1} + (A'A) \otimes (\Sigma_{11\cdot 2}^{(2)})^{-1}]^{-1}\}. \end{aligned}$$

Furthermore, if $A'A = C'C$, *then* $R((\Theta, \Sigma_1, \Sigma_2), \widetilde{\Theta}^{ML}) = mq$.

When (Σ_1, Σ_2) are unknown, $\widetilde{\Theta}^{ML}$ is no longer estimator. However, its risk serves as a lower bound of risk of estimators.

In Tables 1–3, "ML" indicates the maximum likelihood estimator (20) and its risk value was calculated by Lemma 1. Moreover, "GD" and "ST" denote the Graybill-Deal type estimator (12) 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

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

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

$$\hat{\alpha}_j^{ST} = \left\{(n_2 - q - 1)h_{2j} + 4h_{2j}(1 - f_j) + 2h_{2j}\sum_{k\neq j}\frac{f_j(1 - f_k)}{f_j - f_k}\right\}\Big/ f_j,$$

$$\hat{\beta}_j^{ST} = \left\{(n_1 - q - 1)h_{1j} + 4h_{1j}f_j - 2h_{1j}\sum_{k\neq j}\frac{(1 - f_j)f_k}{f_j - f_k}\right\}\Big/ (1 - f_j).$$

Note that we modify $\hat{\alpha}_j$ and $\hat{\beta}_j$ in (18) 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 Tables 1–3 indicates the average of improvement in risk of ST against GD, i.e., AV $= 100(1 - \widehat{R}^{*ST}/\widehat{R}^{*GD})$ %, where $\widehat{R}^{*GD}$ 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 $\mathbf{\Sigma}_{11\cdot 2}^{(2)}(\mathbf{\Sigma}_{11\cdot 2}^{(1)})^{-1}$ are close together, the AVs are large. Specially, in cases when $\boldsymbol{A'A} = \mathbf{diag}\,(3, 1/3)$, $\boldsymbol{C'C} = \mathbf{diag}\,(1,1)$, $N_1 = N_2 = 20$, $p_1 = 7$, $p_2 = 6$, $m = 2$, $q = 5$, and these eigenvalues are equal to 1, the largest AV is 15.9% in Table 2.

2. On the contrary, when the eigenvalues of $\mathbf{\Sigma}_{11\cdot 2}^{(2)}(\mathbf{\Sigma}_{11\cdot 2}^{(1)})^{-1}$ are widely spread out, the AVs are small.

Remark 4. Under another assumptions for $\mathbf{\Sigma}_{11\cdot 2}^{(2)}(\mathbf{\Sigma}_{11\cdot 2}^{(1)})^{-1}$ as examined by Loh (1991), we simulated the risk values of GD and ST and obtained the results that ST performs better than GD.

Table 1. Estimated risks (Estimated standard errors are in parentheses)

$$(\boldsymbol{N_1, N_2, p_1, p_2, m, q}) = (\mathbf{12, 12, 7, 6, 2, 5})$$
$$\boldsymbol{\Theta} = \mathbf{0}_{m\times q},\quad \boldsymbol{\gamma_i} = \mathbf{0}_{(p_i-q)\times q},\quad \mathbf{\Sigma}_{22}^{(i)} = \boldsymbol{I}_{p_i-q}$$
$$\boldsymbol{A'A} = \mathbf{diag}\,(1,1),\quad \boldsymbol{C'C} = \mathbf{diag}\,(1,1)$$

$\mathbf{\Sigma}_{11\cdot 2}^{(2)}(\mathbf{\Sigma}_{11\cdot 2}^{(1)})^{-1}$	ML	GD	ST	AV
$\mathbf{diag}\,(1,1,1,1,1)$	10	17.767 (0.093)	15.175 (0.078)	14.6 %
$\mathbf{diag}\,(10,10,10,1,1)$	10	18.320 (0.104)	16.584 (0.090)	9.5 %
$\mathbf{diag}\,(100,1,1,1,1)$	10	17.840 (0.097)	16.344 (0.088)	8.4 %
$\mathbf{diag}\,(100,100,10,1,1)$	10	18.129 (0.108)	16.687 (0.094)	7.9 %
$\mathbf{diag}\,(10^4,10^4,10^2,1,1)$	10	17.853 (0.106)	16.520 (0.094)	7.5 %
$\mathbf{diag}\,(10^4,10^3,10^2,10,1)$	10	17.950 (0.119)	15.993 (0.094)	10.9 %
$\mathbf{diag}\,(10^8,1,1,1,1)$	10	17.797 (0.095)	16.288 (0.086)	8.5 %
$\mathbf{diag}\,(10^8,10^6,10^4,10^2,1)$	10	16.992 (0.120)	15.508 (0.096)	8.7 %

Table 2. Estimated risks (Estimated standard errors are in parentheses)
$(N_1,\ N_2,\ p_1,\ p_2,\ m,\ q) = (12,\ 12,\ 7,\ 6,\ 2,\ 5)$
$\Theta = 0_{m\times q},\quad \gamma_i = 0_{(p_i-q)\times q},\quad \Sigma_{22}^{(i)} = I_{p_i-q}$
$A'A = \mathbf{diag}\,(1/3, 3),\quad C'C = \mathbf{diag}\,(1, 1)$

$\Sigma_{11\cdot 2}^{(2)}(\Sigma_{11\cdot 2}^{(1)})^{-1}$	ML	GD	ST	AV
diag $(1, 1, 1, 1, 1)$	10.000	18.493 (0.111)	15.557 (0.091)	15.9 %
diag $(10, 10, 10, 1, 1)$	9.732	18.183 (0.118)	15.441 (0.089)	15.1 %
diag $(100, 1, 1, 1, 1)$	9.987	18.559 (0.117)	16.159 (0.097)	12.9 %
diag $(100, 100, 10, 1, 1)$	9.885	18.224 (0.121)	15.600 (0.092)	14.4 %
diag $(10^4, 10^4, 10^2, 1, 1)$	9.987	18.236 (0.121)	15.692 (0.088)	13.9 %
diag $(10^4, 10^3, 10^2, 10, 1)$	9.896	17.770 (0.125)	14.938 (0.086)	15.9 %
diag $(10^8, 1, 1, 1, 1)$	10.000	18.511 (0.114)	16.064 (0.091)	13.2 %
diag $(10^8, 10^6, 10^4, 10^2, 1)$	9.987	17.095 (0.127)	14.729 (0.089)	13.8 %

Table 3. Estimated risks (Estimated standard errors are in parentheses)
$(N_1,\ N_2,\ p_1,\ p_2,\ m,\ q) = (12,\ 12,\ 7,\ 6,\ 2,\ 5)$
$\Theta = 0_{m\times q},\quad \gamma_i = 0_{(p_i-q)\times q},\quad \Sigma_{22}^{(i)} = I_{p_i-q}$
$A'A = \mathbf{diag}\,(1/3, 3),\quad C'C = \mathbf{diag}\,(1/3, 3)$

$\Sigma_{11\cdot 2}^{(2)}(\Sigma_{11\cdot 2}^{(1)})^{-1}$	ML	GD	ST	AV
diag $(1, 1, 1, 1, 1)$	10	18.651 (0.105)	16.880 (0.095)	9.5 %
diag $(10, 10, 10, 1, 1)$	10	18.400 (0.109)	17.458 (0.102)	5.1 %
diag $(100, 1, 1, 1, 1)$	10	18.402 (0.105)	17.711 (0.102)	3.8 %
diag $(100, 100, 10, 1, 1)$	10	18.206 (0.112)	17.413 (0.106)	4.4 %
diag $(10^4, 10^4, 10^2, 1, 1)$	10	18.000 (0.109)	17.262 (0.103)	4.1 %
diag $(10^4, 10^3, 10^2, 10, 1)$	10	17.819 (0.117)	16.383 (0.100)	8.1 %
diag $(10^8, 1, 1, 1, 1)$	10	18.312 (0.103)	17.598 (0.099)	3.9 %
diag $(10^8, 10^6, 10^4, 10^2, 1)$	10	16.990 (0.117)	15.934 (0.103)	6.2 %

4 Proof of Theorem 2

In this section, we state lemmas which are useful in proving the main theorem. 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 2 *Let* $r_1 = m\tilde{r}_1$, $r_2 = \tilde{r}_2 \operatorname{tr}\{(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})\}$, $\tilde{r}_i = (n_i+p_i-q-1)/(n_i-1)$, $i = 1, 2$. *Then we have*

$$\mathbb{E}[\operatorname{tr}\{(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot 2}^{(1)})^{-1}(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})'\}] = qr_1, \tag{21a}$$

$$\mathbb{E}[\operatorname{tr}\{(\widehat{\boldsymbol{\Theta}}_2 - \boldsymbol{A}\boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot 2}^{(2)})^{-1}(\widehat{\boldsymbol{\Theta}}_2 - \boldsymbol{A}\boldsymbol{\Theta})'(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})\}] = qr_2, \tag{21b}$$

$$\begin{aligned}\mathbb{E}[\operatorname{tr}\{(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})&(\boldsymbol{\Sigma}_{11\cdot 2}^{(1)})^{-1}\boldsymbol{B}^{-1}(\boldsymbol{I}_q - \boldsymbol{\Phi})\boldsymbol{B}(\boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2 - \widehat{\boldsymbol{\Theta}}_1)'\}]\\ &= -\mathbb{E}\left[\left(q - \sum_{i=1}^{q}\phi_i\right)r_1\right],\end{aligned} \tag{21c}$$

$$\begin{aligned}\mathbb{E}[\operatorname{tr}\{(\widehat{\boldsymbol{\Theta}}_2 - \boldsymbol{A}\boldsymbol{\Theta})&(\boldsymbol{\Sigma}_{11\cdot 2}^{(2)})^{-1}\boldsymbol{B}^{-1}\boldsymbol{\Phi}\boldsymbol{B}(\boldsymbol{A}\widehat{\boldsymbol{\Theta}}_1 - \widehat{\boldsymbol{\Theta}}_2)'(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})\}]\\ &= -\mathbb{E}\left[\left(\sum_{i=1}^{q}\phi_i\right)r_2\right].\end{aligned} \tag{21d}$$

Proof. Note that

$$\begin{aligned}\widehat{\boldsymbol{\Theta}}_1 \mid \boldsymbol{Z}_1, \boldsymbol{W}_1 &\sim N_{m\times q}(\boldsymbol{\Theta}, (\boldsymbol{I}_m + \boldsymbol{Z}_1\boldsymbol{W}_1^{-1}\boldsymbol{Z}_1') \otimes \boldsymbol{\Sigma}_{11\cdot 2}^{(1)}),\\ \widehat{\boldsymbol{\Theta}}_2 \mid \boldsymbol{Z}_2, \boldsymbol{W}_2 &\sim N_{m\times q}(\boldsymbol{A}\boldsymbol{\Theta}, (\boldsymbol{I}_m + \boldsymbol{Z}_2\boldsymbol{W}_2^{-1}\boldsymbol{Z}_2') \otimes \boldsymbol{\Sigma}_{11\cdot 2}^{(2)}),\end{aligned}$$

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

$$\begin{aligned}
&\mathbb{E}[\operatorname{tr}\{(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot 2}^{(1)})^{-1}(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})'\}] = \mathbb{E}[q\operatorname{tr}(\boldsymbol{I}_m + \boldsymbol{Z}_1\boldsymbol{W}_1^{-1}\boldsymbol{Z}_1')],\\
&\mathbb{E}[\operatorname{tr}\{(\widehat{\boldsymbol{\Theta}}_2 - \boldsymbol{A}\boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot 2}^{(2)})^{-1}(\widehat{\boldsymbol{\Theta}}_2 - \boldsymbol{A}\boldsymbol{\Theta})'(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})\}]\\
&\quad = \mathbb{E}[q\operatorname{tr}\{(\boldsymbol{I}_m + \boldsymbol{Z}_2\boldsymbol{W}_2^{-1}\boldsymbol{Z}_2')(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})\}],\\
&\mathbb{E}[\operatorname{tr}\{(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot 2}^{(1)})^{-1}\boldsymbol{B}^{-1}(\boldsymbol{I}_q - \boldsymbol{\Phi})\boldsymbol{B}(\boldsymbol{A}^{-1}\widehat{\boldsymbol{\Theta}}_2 - \widehat{\boldsymbol{\Theta}}_1)'\}]\\
&\quad = -\mathbb{E}[\operatorname{tr}\{(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot 2}^{(1)})^{-1}\boldsymbol{B}^{-1}(\boldsymbol{I}_q - \boldsymbol{\Phi})\boldsymbol{B}(\widehat{\boldsymbol{\Theta}}_1 - \boldsymbol{\Theta})'\}]\\
&\quad = -\mathbb{E}[\operatorname{tr}\{\boldsymbol{B}^{-1}(\boldsymbol{I}_q - \boldsymbol{\Phi})\boldsymbol{B}\} \times \operatorname{tr}(\boldsymbol{I}_m + \boldsymbol{Z}_1\boldsymbol{W}_1^{-1}\boldsymbol{Z}_1')],\\
&\mathbb{E}[\operatorname{tr}\{(\widehat{\boldsymbol{\Theta}}_2 - \boldsymbol{A}\boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot 2}^{(2)})^{-1}\boldsymbol{B}^{-1}\boldsymbol{\Phi}\boldsymbol{B}(\boldsymbol{A}\widehat{\boldsymbol{\Theta}}_1 - \widehat{\boldsymbol{\Theta}}_2)'(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})\}]\\
&\quad = -\mathbb{E}[\operatorname{tr}\{(\widehat{\boldsymbol{\Theta}}_2 - \boldsymbol{A}\boldsymbol{\Theta})(\boldsymbol{\Sigma}_{11\cdot 2}^{(2)})^{-1}\boldsymbol{B}^{-1}\boldsymbol{\Phi}\boldsymbol{B}(\widehat{\boldsymbol{\Theta}}_2 - \boldsymbol{A}\boldsymbol{\Theta})'(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})\}]\\
&\quad = -\mathbb{E}[\operatorname{tr}\{\boldsymbol{B}^{-1}\boldsymbol{\Phi}\boldsymbol{B}\} \times \operatorname{tr}\{(\boldsymbol{I}_m + \boldsymbol{Z}_2\boldsymbol{W}_2^{-1}\boldsymbol{Z}_2')(\boldsymbol{C}\boldsymbol{A}^{-1})'(\boldsymbol{C}\boldsymbol{A}^{-1})\}].
\end{aligned}$$

Finally, from (6a) and (6d), we get (21a)–(21d). □

Lemma 3 (Stein-Haff identity) *Assume that a $q \times q$ positive definite matrix $\boldsymbol{S}$ follows the Wishart distribution $W_q(\boldsymbol{\Sigma}, a)$. Also let*

$$\mathcal{D} = \left(\frac{1}{2}(1+\delta_{ij}) \frac{\partial}{\partial s_{ij}} \right), \tag{22}$$

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

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

Lemma 4 (Loh, 1988 and 1991) *For $i = 1, 2$, let $\mathcal{D}_i$ be $q \times q$ differential operators which are define by* (22) *with replacing $\boldsymbol{S}$ by $\boldsymbol{S}_i$. Also let $\boldsymbol{x}$ be a $q \times 1$ vector which is independent of $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$. Then*

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

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

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

Lemma 5

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

where $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are given by (15a) and (15b), respectively.

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

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

Hence, from this equation and Lemma 4, we get the first expression. The second expression can be obtained from the similar argument. □

Proof of Theorem 2. First apply Lemma 3 to the third and sixth terms in right-hand side of (14) and then use Lemma 2 to the other terms in right-hand side of (14) to get that the risk $\boldsymbol{R}((\boldsymbol{\Theta}, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2), \widehat{\boldsymbol{\Theta}}^{EQ})$ is rewritten as

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

Finally apply Lemma 5 to the third and fourth terms inside the expectation of the above equation to complete the theorem. □

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

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

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

[9] Johnson, M. (1987). *Multivariate Statistical Simulation.* John Wiley and Sons, New York.

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

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

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

[13] Khatri, C.G. and Shah, K.R. (1974). Estimation of location parameters from two linear models under normality. *Comm. Statist. Theory Methods*, **3**, 647–663.

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

[15] Kubokawa, T. and Srivastava, M.S. (2002). Prediction in multivariate mixed linear models. Discussion paper CIRJE-F-180, Faculty of Economics, University of Tokyo.

[16] 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.

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

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

[19] Shinozaki, N. (1978). A note on estimating the common mean of k normal distributions and Stein problem. *Comm. Statist. Theory Methods* **7**, 1421-1432.

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

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

[22] 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.).