Ann. Inst. Statist. Math.
Vol. 47, No. 4, 665–674 (1995)

ON CONSTRUCTION OF IMPROVED ESTIMATORS IN MULTIPLE-DESIGN MULTIVARIATE LINEAR MODELS UNDER GENERAL RESTRICTION

T. SHIRAISHI[1] AND Y. KONNO[2]

[1] *Department of Mathematical Sciences, Yokohama City University, 22-2 Seto, Kanazawa-ku, Yokohama 236, Japan*
[2] *Department of Mathematics and Informatics, Chiba University, 1-33 Yayoi-cho, Chiba 263, Japan*

(Received December 20, 1993; revised January 4, 1995)

Abstract. Consider a set of p equations $\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\xi}_i + \boldsymbol{\epsilon}_i$, $i = 1, \dots, p$, where the rows of the random error matrix $(\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_p) : n \times p$ are mutually independent and identically distributed with a p-variate distribution function $F(\boldsymbol{x})$ having null mean and finite positive definite variance-covariance matrix $\boldsymbol{\Sigma}$. We are mainly interested in an improvement upon a feasible generalized least squares estimator (FGLSE) for $\boldsymbol{\xi} = (\boldsymbol{\xi}_1', \dots, \boldsymbol{\xi}_p')'$ when it is a priori suspected that $\boldsymbol{C\xi} = \boldsymbol{c}_0$ may hold. For this problem, Saleh and Shiraishi (1992, *Nonparametric Statistics and Related Topics* (ed. A. K. Md. E. Saleh), 269–279, North-Holland, Amsterdam) investigated the property of estimators such as the shrinkage estimator (SE), the positive-rule shrinkage estimator (PSE) in the light of their asymptotic distributional risks associated with the Mahalanobis loss function. We consider a general form of estimators and give a sufficient condition for proposed estimators to improve on FGLSE with respect to their asymptotic distributional quadratic risks (ADQR). The relative merits of these estimators are studied in the light of the ADQR under local alternatives. It is shown that the SE, the PSE and the Kubokawa-type shrinkage estimator (KSE) outperform the FGLSE and that the PSE is the most effective among the four estimators considered under $\boldsymbol{C\xi} = \boldsymbol{c}_0$. It is also observed that the PSE and the KSE fairly improve over the FGLSE. Lastly, the construction of estimators improved on a generalized least squares estimator is studied, assuming normality when $\boldsymbol{\Sigma}$ is known.

Key words and phrases: Shrinkage estimators, generalized least squares estimators, asymptotic distribution, seemingly unrelated regression model.

1. Introduction

Consider p different regression models with cross correlation,

$$(1.1) \qquad \boldsymbol{Y}_i : n \times 1 = \boldsymbol{X}_i\boldsymbol{\xi}_i + \boldsymbol{\epsilon}_i \quad \text{and} \quad E[\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_j'] = \sigma_{ij}\boldsymbol{I}_n, \qquad (i, j = 1, \dots, p)$$

where $\boldsymbol{X}_i$ $(n \times q_i)$ is a design matrix of full rank, $\boldsymbol{\xi}_i$ $(q_i \times 1)$ is a column vector of unknown parameter, and $\boldsymbol{\epsilon}_i$ $(n \times 1)$ is an error term vector continuously distributed with $E[\boldsymbol{\epsilon}_i] = \mathbf{0}$. We also assume that $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=1,\dots,p}$ is positive definite. The model (1.1) is referred to as the seemingly unrelated regression model of Zellner (1962). We can rewrite the model (1.1) as follows:

$$\text{(1.2)} \qquad \boldsymbol{Y} = \boldsymbol{X\xi} + \boldsymbol{\epsilon} \quad \text{and} \quad \mathrm{Cov}(\boldsymbol{\epsilon}) = \boldsymbol{\Sigma} \otimes \boldsymbol{I}_n,$$

where $\boldsymbol{Y} : np \times 1 = (\boldsymbol{Y}_1', \dots, \boldsymbol{Y}_p')'$, $\boldsymbol{\xi} : q \times 1 = (\boldsymbol{\xi}_1', \dots, \boldsymbol{\xi}_p')'$, $\boldsymbol{\epsilon} : np \times 1 = (\boldsymbol{\epsilon}_1', \dots, \boldsymbol{\epsilon}_p')'$,

$$\boldsymbol{X} : np \times 1 = \begin{pmatrix} \boldsymbol{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{X}_p \end{pmatrix}, \quad \text{and} \quad q = \sum_{i=1}^{p} q_i.$$

In this paper, we are primarily interested in the estimation of $\boldsymbol{\xi}$ when it is suspected that a general restriction

$$\text{(1.3)} \qquad H_0 : \boldsymbol{C\xi} = \boldsymbol{c}_0$$

holds, where $\boldsymbol{C}$ is an $r \times q$ matrix of rank r and $\boldsymbol{c}_0$ is an r-dimensional column vector. Taking account of the form of the covariance matrix in (1.2) and utilizing any consistent estimators $\hat{\boldsymbol{\Sigma}}$ of $\boldsymbol{\Sigma}$, we have a feasible generalized least squares estimator (FGLSE)

$$\text{(1.4)} \qquad \tilde{\boldsymbol{\xi}}_n = \{\boldsymbol{X}'(\hat{\boldsymbol{\Sigma}}_n^{-1} \otimes \boldsymbol{I}_n)\boldsymbol{X}\}^{-1}\boldsymbol{X}'(\hat{\boldsymbol{\Sigma}}_n^{-1} \otimes \boldsymbol{I}_n)\boldsymbol{Y}.$$

For example, we may take

$$\hat{\boldsymbol{\Sigma}}_n = n^{-1}[\operatorname*{mat}_{n\times p}(\boldsymbol{Y} - \boldsymbol{X}\tilde{\boldsymbol{\xi}}_n)]'[\operatorname*{mat}_{n\times p}(\boldsymbol{Y} - \boldsymbol{X}\tilde{\boldsymbol{\xi}}_n)],$$

as a consistent estimator. Here, we define $\mathrm{mat}_{n\times p}(\cdot)$ by $\mathrm{mat}_{n\times p}(\mathbb{Z}) = (\mathbb{Z}_1, \dots, \mathbb{Z}_p)$ for $\mathbb{Z} : np \times 1 = (\mathbb{Z}_1', \dots, \mathbb{Z}_p')'$ and p-dimensional column vectors $\mathbb{Z}_i$, $i = 1, \dots, p$. See Srivastava and Giles (1987) for a review of the estimator (1.4). Furthermore, taking the restriction (1.3) into consideration, we have a restricted estimator (RE),

$$\text{(1.5)} \qquad \hat{\boldsymbol{\xi}}_n = \tilde{\boldsymbol{\xi}}_n - (\hat{\boldsymbol{\Sigma}}_n^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{C}'\{\boldsymbol{C}(\hat{\boldsymbol{\Sigma}}_n^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{C}'\}^{-1}(\boldsymbol{C}\tilde{\boldsymbol{\xi}}_n - \boldsymbol{c}_0),$$

where $\boldsymbol{D}_n : q \times q = (\boldsymbol{D}_{nij})_{i,j=1,\dots,p}$ with $\boldsymbol{D}_{nij} : q_i \times q_j = \boldsymbol{X}_i\boldsymbol{X}'_j/n$, and the matrix operator $\#$ is defined by $\boldsymbol{A}\#\boldsymbol{B} = (a_{ij}\boldsymbol{B}_{ij})_{i,j=1,\dots,p}$ for $\boldsymbol{A} = (a_{ij})_{i,j=1,\dots,p}$ and $q \times q$ matrix $\boldsymbol{B} = (\boldsymbol{B}_{ij})_{i,j=1,\dots,p}$ so that $\boldsymbol{B}_{ij}$ is $q_i \times q_j$.

$\hat{\boldsymbol{\xi}}_n$ may be biased and even inconsistent unless the restriction (1.3) holds, while it performs better than $\tilde{\boldsymbol{\xi}}_n$ when (1.3) holds. So we propose a weighted combination of $\tilde{\boldsymbol{\xi}}_n$ and $\hat{\boldsymbol{\xi}}_n$, being of the form

$$\text{(1.6)} \qquad \hat{\boldsymbol{\xi}}_n^g = \hat{\boldsymbol{\xi}}_n + \{\boldsymbol{I}_q - g(\mathcal{L}_n)(\hat{\boldsymbol{\Sigma}}_n^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{C}'\hat{\boldsymbol{\Gamma}}_n^{-1}\boldsymbol{C}\}(\tilde{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_n),$$

where

$$\mathcal{L}_n = n(\tilde{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_n)'\boldsymbol{C}'\hat{\boldsymbol{\Gamma}}_n^{-1}\boldsymbol{C}(\tilde{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_n) \quad \text{and}$$
$$\hat{\boldsymbol{\Gamma}}_n = \boldsymbol{C}(\hat{\boldsymbol{\Sigma}}_n^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{Q}_n(\hat{\boldsymbol{\Sigma}}_n^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{C}'.$$

Here, $\mathcal{L}_n$ is a test statistic for testing $H_0 : \boldsymbol{C\xi} = \boldsymbol{c}_0$ v.s. $H_A : \boldsymbol{C\xi} \neq \boldsymbol{c}_0$, and $\boldsymbol{Q}_n$ is a consistent estimator of $\boldsymbol{Q}$ which is a weight matrix of full rank associated with a quadratic loss function

$$(1.7) \qquad (\hat{\boldsymbol{\xi}}^+ - \boldsymbol{\xi})'\boldsymbol{Q}(\hat{\boldsymbol{\xi}}^+ - \boldsymbol{\xi})$$

for an estimator $\hat{\boldsymbol{\xi}}^+$ of $\boldsymbol{\xi}$. The estimator (1.6) reduces to one considered in Saleh and Shiraishi (1992) when $\boldsymbol{Q}_n = \hat{\boldsymbol{\Sigma}}_n^{-1}\#\boldsymbol{D}_n$. We obtain a sufficient condition that the proposed estimator outperforms the FGLSE (1.1) with respect to their asymptotic distributional risks (i.e., the risk by reference to the asymptotic distribution of an estimator) associated with the loss function (1.7) under the local alternatives. We construct the shrinkage estimator (SE), the positive-rule shrinkage estimators (PSE), and the Kubokawa-type estimator (KSE), which are better than the FGLSE.

2. Asymptotic distributional quadratic risks (ADQR)

First, we introduce the assumption to compute the asymptotic distributional risks of estimators. Let

$$(2.1) \qquad \tilde{\boldsymbol{\xi}}_n^* = \{\boldsymbol{X}'(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_n)\boldsymbol{X}\}^{-1}\boldsymbol{X}'(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_n)\boldsymbol{Y},$$

which is referred to as the generalized least squares estimator. The variance-covariance matrix of $\tilde{\boldsymbol{\xi}}_n^*$ is given by

$$\mathrm{Var}(\tilde{\boldsymbol{\xi}}_n^*) = \{\boldsymbol{X}'(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_n)\boldsymbol{X}\}^{-1} = (\boldsymbol{\Sigma}^{-1}\#\boldsymbol{D}_n)^{-1}.$$

Now, we set

ASSUMPTION 1. $\lim_{n\to\infty} n^{-1}\boldsymbol{D}_n = \boldsymbol{\Delta}$.

ASSUMPTION 2.

$$\sqrt{n}(\tilde{\boldsymbol{\xi}}_n^* - \boldsymbol{\xi}) \xrightarrow{L} N_q(\boldsymbol{0}, (\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1}),$$

where $\xrightarrow{L}$ denotes convergence in distribution.

Assumption 2 holds if Lindeberg's condition is satisfied. We consider the following contiguous sequence of alternatives

$$A_n : \boldsymbol{C\xi} = \boldsymbol{c}_0 + \boldsymbol{\theta}/\sqrt{n}; \boldsymbol{\theta} = (\theta_1, \ldots, \theta_r)'.$$

Then, we have the following Lemma from Saleh and Shiraishi (1992).

LEMMA 2.1. *Under Assumptions* 1, 2 *and under* A_n, *the asymptotic distributions of* $\sqrt{n}(\hat{\boldsymbol{\xi}}_n - \boldsymbol{\xi})$, $\sqrt{n}\boldsymbol{C}(\tilde{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_n)$, $\mathcal{L}_n$ *and* $\sqrt{n}(\tilde{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_n)$ *are given by*

$$\sqrt{n}(\hat{\boldsymbol{\xi}}_n - \boldsymbol{\xi}) \xrightarrow{L} \boldsymbol{Y} \sim N_q(-\boldsymbol{\mu}_0, (\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1} - \boldsymbol{B}_1), \tag{2.2}$$

$$\sqrt{n}\boldsymbol{C}(\tilde{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_n) \xrightarrow{L} \boldsymbol{Z} \sim N_r(\boldsymbol{\theta}, \boldsymbol{B}_2), \tag{2.3}$$

$$\mathcal{L}_n \xrightarrow{L} \boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z},$$

and

$$\sqrt{n}(\tilde{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_n) \xrightarrow{L} (\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1}\{\boldsymbol{C}'\boldsymbol{B}_2^{-1}\boldsymbol{C}\}^{-1}\boldsymbol{Z} \sim N_q(\boldsymbol{\mu}_0, \boldsymbol{B}_1), \tag{2.4}$$

where $\boldsymbol{B}_1 = (\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1}\boldsymbol{C}'\boldsymbol{B}_2\boldsymbol{C}(\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1}$, $\boldsymbol{\mu}_0 = (\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1}\{\boldsymbol{C}'\boldsymbol{B}_2^{-1}\boldsymbol{C}\}^{-1}\boldsymbol{\theta}$, $\boldsymbol{B}_2 = \boldsymbol{C}(\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1}\boldsymbol{C}'$ *and* $\boldsymbol{\Gamma} = \boldsymbol{C}(\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1}\boldsymbol{Q}(\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1}\boldsymbol{C}'$. *Furthermore* $\sqrt{n}(\hat{\boldsymbol{\xi}}_n - \boldsymbol{\xi})$, *and* $\sqrt{n}(\tilde{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_n)$ *are asymptotically independent under* A_n.

Assume that the asymptotic c.d.f. is obtained as

$$G_r^+(\boldsymbol{x}) = \lim_{n\to\infty} P\{\sqrt{n}(\hat{\boldsymbol{\xi}}_n^+ - \boldsymbol{\xi}) \le \boldsymbol{x}\}$$

for an estimator $\hat{\boldsymbol{\xi}}_n^+$ of $\boldsymbol{\xi}$. Then we define the asymptotic distributional quadratic risk (ADQR) of $\hat{\boldsymbol{\xi}}_n^+$ by the expression

$$\begin{aligned} R(\hat{\boldsymbol{\xi}}_n^+ : \boldsymbol{Q}) &= \lim_{b\to\infty}\lim_{n\to\infty} E[\min\{n(\hat{\boldsymbol{\xi}}_n^+ - \boldsymbol{\xi})'\boldsymbol{Q}(\hat{\boldsymbol{\xi}}_n^+ - \boldsymbol{\xi}), b\}] \\ &= \int_{R^r} \boldsymbol{x}'\boldsymbol{Q}\boldsymbol{x}\,dG^+(\boldsymbol{x}) = \mathrm{tr}(\boldsymbol{Q}\boldsymbol{\Sigma}^+), \end{aligned} \tag{2.5}$$

where $\boldsymbol{\Sigma}^+ = \int_{R^r} \boldsymbol{x}\boldsymbol{x}'dG^+(\boldsymbol{x})$ and $\boldsymbol{Q}$ is the weight matrix in the loss function (1.7). In order to obtain a sufficient condition for the proposed estimators to outperform the FGLSE, we investigate the ADQR for $\hat{\boldsymbol{\xi}}_n^g$.

LEMMA 2.2. *Under Assumptions* 1, 2 *and under* A_n, *the ADQR* $R(\hat{\boldsymbol{\xi}}_n^g \boldsymbol{Q})$ *for the estimator* $\hat{\boldsymbol{\xi}}_n^g$ *defined by* (1.6) *has the following expression,*

$$\begin{aligned} &\mathrm{tr}((\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1}\boldsymbol{Q}) - E[2rg(\boldsymbol{Y}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Y}) + 4g'(\boldsymbol{Y}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Y})\boldsymbol{Y}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Y} \\ &\qquad + g^2(\boldsymbol{Y}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Y})\boldsymbol{Y}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Y}]. \end{aligned} \tag{2.6}$$

PROOF. Using (2.3) and (2.4), we get

$$\begin{aligned} \sqrt{n}(\hat{\boldsymbol{\xi}}_n^g - \boldsymbol{\xi}) \xrightarrow{L} \boldsymbol{Y} &+ [\{(\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1}\boldsymbol{C}'\boldsymbol{B}_2{}^{-1}\boldsymbol{C}\}^{-1} \\ &- g(\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Y})(\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1}\boldsymbol{C}'\boldsymbol{\Gamma}^{-1}]\boldsymbol{Y}, \end{aligned} \tag{2.7}$$

where $\boldsymbol{Y}$ and $\boldsymbol{Z}$ are independent and are respectively defined in (2.2) and (2.3). From (2.7), we get that the ADQR $R(\hat{\boldsymbol{\xi}}_n^g : \boldsymbol{Q})$ is expressed as

$$E\{(\boldsymbol{Y}+\boldsymbol{\mu}_0)'\boldsymbol{Q}(\boldsymbol{Y}+\boldsymbol{\mu}_0)\} + E[(\boldsymbol{Z} - g(\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z})\boldsymbol{B}_2\boldsymbol{\Gamma}^{-1}\boldsymbol{Z} - \boldsymbol{\theta})'\boldsymbol{B}_2{}^{-1}\boldsymbol{\Gamma}\boldsymbol{B}_2{}^{-1} \\ \times(\boldsymbol{Z} - g(\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z})\boldsymbol{B}_2\boldsymbol{\Gamma}^{-1}\boldsymbol{Z} - \boldsymbol{\theta})].$$

After a simple computation, it follows that the ADQR $R(\hat{\boldsymbol{\xi}}_n^g : \boldsymbol{Q})$ is equal to

$$\operatorname{tr}(\boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})^{-1}\boldsymbol{Q} - 2E[\operatorname{tr}(\boldsymbol{Z}-\boldsymbol{\theta})'\boldsymbol{B}_2{}^{-1}(g(\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z})\boldsymbol{Z}) \\ - \operatorname{tr} g^2(\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z})\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z}]. \tag{2.8}$$

Under the regularity conditions stated in Stein (1981) or Bilodeau and Kariya (1989), we have

$$E[g(\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z})\boldsymbol{Z}(\boldsymbol{Z}-\boldsymbol{\theta})'\boldsymbol{B}_2{}^{-1}] = E[(\partial/\partial\boldsymbol{Z})(g(\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z})\boldsymbol{Z})]$$

where $(\partial/\partial\boldsymbol{Z})(g(\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z})\boldsymbol{Z})$ is an $r \times r$ matrix whose (i,j) element is $(\partial/\partial Z_j)(g(\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z})Z_i)$ for $\boldsymbol{Z} = (Z_1,\dots,Z_r)'$ and $\partial/\partial\boldsymbol{Z} = (\partial/\partial Z_1,\dots,\partial/\partial Z_r)'$. From chain-rule and straightforward calculation, it follows that

$$E[(\partial/\partial\boldsymbol{Z})(g(\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z})\boldsymbol{Z})] = E[g(\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z})\boldsymbol{I}_r + 2g'(\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}\boldsymbol{Z})\boldsymbol{Z}\boldsymbol{Z}'\boldsymbol{\Gamma}^{-1}].$$

Finally, putting this equation into (2.8), we get (2.6). □

3. Asymptotic dominance

Using Lemma 2.2, we give a sufficient condition for the estimators (1.6) to outperform the FGLSE.

THEOREM 3.1. *Suppose that $r \geq 3$. Under Assumptions* 1, 2 *and under A_n, the estimator $\hat{\boldsymbol{\xi}}_n^g$ asymptotically outperforms $\tilde{\boldsymbol{\xi}}_n$ if* (i) $0 \leq g(u) \leq 2(r-2)$ *for any u and* (ii) $g(u)$ *is nonincreasing.*

PROOF. This is a direct consequence from Lemma 2.2. □

Set $g(u) = I(u \leq \mathcal{L}_{n,\alpha})$, $(r-2)\cdot u^{-1}$, $1 - \{1-(r-2)\cdot u^{-1}\}\cdot I(u \geq r-2)$, and $\phi(u)\cdot u^{-1}$ in (1.6), where I_A is an indicator function, $\phi(u)$ is increasing, $\lim_{t\to\infty}\phi(u) = r-2$,

$$\phi(u) \geq \phi_0(u) = r - 2 - 2f_r(u)\Big/\int_0^u t^{-1}f_r(t)dt, \tag{3.1}$$

and $f_r(t)$ is a density of the chi-square distribution with r degrees of freedom. Then, we define a preliminary test estimator (PTE), a shrinkage estimator (SE), a positive-rule shrinkage estimator (PSE), and a Kubokawa-type shrinkage estimator (KSE) $\hat{\boldsymbol{\xi}}_n^{KS}$ as

$$\begin{aligned}
\hat{\boldsymbol{\xi}}_n^{PT} &= \hat{\boldsymbol{\xi}}_n + \{\boldsymbol{I}_q - I(\mathcal{L}_n \le \mathcal{L}_{n,\alpha})(\hat{\boldsymbol{\Sigma}}_n^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{C}'\hat{\boldsymbol{\Gamma}}_n^{-1}\boldsymbol{C}\}(\tilde{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_n),\\
\hat{\boldsymbol{\xi}}_n^{S} &= \hat{\boldsymbol{\xi}}_n + \{\boldsymbol{I}_q - (r-2)\mathcal{L}_n^{-1}(\hat{\boldsymbol{\Sigma}}_n^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{C}'\hat{\boldsymbol{\Gamma}}_n^{-1}\boldsymbol{C}\}(\tilde{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_n),\\
\hat{\boldsymbol{\xi}}_n^{PS} &= \hat{\boldsymbol{\xi}}_n + \{\boldsymbol{I}_q - [1 - \{1-(r-2)\mathcal{L}_n^{-1}\}\cdot I(\mathcal{L}_n^{-1} \ge r-2)]\\
&\qquad\qquad\qquad \cdot(\hat{\boldsymbol{\Sigma}}_n^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{C}'\hat{\boldsymbol{\Gamma}}_n^{-1}\boldsymbol{C}\}(\tilde{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_n),
\end{aligned}$$

and

$$\hat{\boldsymbol{\xi}}_n^{KS} = \hat{\boldsymbol{\xi}}_n + \{\boldsymbol{I}_q - \phi(\mathcal{L}_n)\cdot\mathcal{L}_n^{-1}(\hat{\boldsymbol{\Sigma}}_n^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{C}'\hat{\boldsymbol{\Gamma}}_n^{-1}\boldsymbol{C}\}(\tilde{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_n),$$

respectively. Now, we have the following corollary:

COROLLARY 3.1. *Suppose that $r \ge 3$. Under Assumptions* 1, 2 *and under A_n, the estimators $\hat{\boldsymbol{\xi}}_n^{S}$, $\hat{\boldsymbol{\xi}}_n^{PS}$ and $\hat{\boldsymbol{\xi}}_n^{KS}$ asymptotically outperform the FGLSE $\tilde{\boldsymbol{\xi}}_n$ with respect to the ADQR defined by* (2.5).

THEOREM 3.2. *Put $\boldsymbol{Q}_n = (\hat{\boldsymbol{\Sigma}}_n^{-1}\#\boldsymbol{D}_n)^{-1}$ in $\hat{\boldsymbol{\xi}}_n^{g}$ of* (1.6), *and suppose that $r \ge 3$. Under Assumptions* 1, 2 *and under H_0, the ADQR of $\hat{\boldsymbol{\xi}}_n^{PS}$ is less than those of $\hat{\boldsymbol{\xi}}_n^{S}$, $\hat{\boldsymbol{\xi}}_n^{KS}$, $\hat{\boldsymbol{\xi}}_n^{PT}$ and $\tilde{\boldsymbol{\xi}}_n$, where $\phi_0(u)$ defined in* (3.1) *is taken as $\phi(u)$ in $\hat{\boldsymbol{\xi}}_n^{KS}$ and $k \le r-2$ in $\hat{\boldsymbol{\xi}}_n^{PT}$. Furthermore we have, for any positive definite matrix $\boldsymbol{S}$,*

$$R(\hat{\boldsymbol{\xi}}_n^{PS} : \boldsymbol{S}) \le R(\hat{\boldsymbol{\xi}}_n^{S} : \boldsymbol{S}) = R(\hat{\boldsymbol{\xi}}_n^{KS} : \boldsymbol{S}), R(\hat{\boldsymbol{\xi}}_n^{PT} : \boldsymbol{S}) \le R(\tilde{\boldsymbol{\xi}}_n : \boldsymbol{S}). \tag{3.2}$$

PROOF. From Theorem 3.1 in Saleh and Shiraishi (1992), the ADQR $R(\hat{\boldsymbol{\xi}}_n^{g} : \boldsymbol{S})$ is an increasing function of $E\{1 - g(\chi_{r+2}^2(0))\}^2$ under H_0. Hence, we get

$$E\{1 - g(\chi_{r+2}^2(0))\}^2 = \begin{cases}
\displaystyle\int_{r-2}^{\infty}\{1-(r-2)/x\}^2 f_{r+2}(x)dx & \text{if } \hat{\boldsymbol{\xi}}_n^{g} = \hat{\boldsymbol{\xi}}_n^{PS},\\
\displaystyle\int_{0}^{\infty}\{1-(r-2)/x\}^2 f_{r+2}(x)dx & \text{if } \hat{\boldsymbol{\xi}}_n^{g} = \hat{\boldsymbol{\xi}}_n^{S},\\
\displaystyle\int_{0}^{\infty}\{1-\phi_0(x)/x\}^2 f_{r+2}(x)dx & \text{if } \hat{\boldsymbol{\xi}}_n^{g} = \hat{\boldsymbol{\xi}}_n^{KS},\\
\displaystyle\int_{k}^{\infty} f_{r+2}(x)dx & \text{if } \hat{\boldsymbol{\xi}}_n^{g} = \hat{\boldsymbol{\xi}}_n^{PT},\\
\displaystyle\int_{0}^{\infty} f_{r+2}(x)dx & \text{if } \hat{\boldsymbol{\xi}}_n^{g} = \tilde{\boldsymbol{\xi}}_n,
\end{cases} \tag{3.3}$$

from which it follows that

$$R(\hat{\boldsymbol{\xi}}_n^{PS} : \boldsymbol{S}) \le R(\hat{\boldsymbol{\xi}}_n^{S} : \boldsymbol{S}), R(\hat{\boldsymbol{\xi}}_n^{PT} : \boldsymbol{S}) \le R(\tilde{\boldsymbol{\xi}}_n : \boldsymbol{S}).$$

Furthermore, using equation (4.3) in Kubokawa (1994), we get

$$\int_0^\infty \{1-(r-2)/x\}^2 f_{r+2}(x)dx - \int_0^\infty \{1-\phi_0(x)/x\}^2 f_{r+2}(x)dx = 0,$$

which implies

$$R(\hat{\boldsymbol{\xi}}_n^{KS} : \boldsymbol{S}) = R(\hat{\boldsymbol{\xi}}_n^{S} : \boldsymbol{S}). \qquad \square$$

If we employ the Mahalanobis loss function, i.e., $\boldsymbol{S} = \boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta}$ in (3.2), we can see stronger ordering than that given in Corollary 3.1.

THEOREM 3.3. *Set* $\boldsymbol{Q}_n = (\hat{\boldsymbol{\Sigma}}_n^{-1}\#\boldsymbol{D}_n)^{-1}$, *and suppose that* $r \geq 3$. *Under Assumptions* 1, 2 *and under* A_n, *the PSE and the KSE asymptotically outperform the SE with respect to the ADQR defined by* (2.4) *with* $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta}$.

PROOF. Setting $\boldsymbol{\Gamma} = \boldsymbol{B}_1$ and $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta}$ in (2.6), we get

$$\begin{aligned} R(\hat{\boldsymbol{\xi}}_n^{g} : \boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta}) = q &- E\{2rg(\boldsymbol{Z}_0'\boldsymbol{Z}_0) + 4g'(\boldsymbol{Z}_0'\boldsymbol{Z}_0)\boldsymbol{Z}_0'\boldsymbol{Z}_0\} \\ &+ E\{g^2(\boldsymbol{Z}_0'\boldsymbol{Z}_0)\boldsymbol{Z}_0'\boldsymbol{Z}_0\}, \end{aligned}$$

where $\boldsymbol{Z}_0 = \boldsymbol{B}_1{}^{-1/2}\boldsymbol{Z} \sim N_r(\boldsymbol{\nu_0}, \boldsymbol{I}_r)$. Since the relation $R(\hat{\boldsymbol{\xi}}_n^{PS} : \boldsymbol{Q}) \leq R(\hat{\boldsymbol{\xi}}_n^{S} : \boldsymbol{Q})$ is shown in Saleh and Shiraishi (1992), it suffices to prove $R(\hat{\boldsymbol{\xi}}_n^{KS} : \boldsymbol{Q}) \leq R(\hat{\boldsymbol{\xi}}_n^{S} : \boldsymbol{Q})$. This follows from the argument of the proof of Theorem 4.1 in Kubokawa (1994). $\square$

Under the Mahalanobis loss function, Theorem 3.3 implies that the PSE and the KSE outperform the FGLSE and the SE. Hence, we investigate how much the PSE and the KSE improve upon the FGSLE by utilizing an efficiency. We define an asymptotic relative risk efficiency (ARRE) of $\hat{\boldsymbol{\xi}}_n^{+}$ with respect to $\tilde{\boldsymbol{\xi}}_n$ (FGLSE) by $R(\tilde{\boldsymbol{\xi}}_n : \boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})/R(\hat{\boldsymbol{\xi}}_n^{+} : \boldsymbol{\Sigma}^{-1}\#\boldsymbol{\Delta})$, which is denoted by $ARRE(\hat{\boldsymbol{\xi}}_n^{+}, \tilde{\boldsymbol{\xi}}_n)$. If $ARRE(\hat{\boldsymbol{\xi}}_n^{+}, \tilde{\boldsymbol{\xi}}_n) > 1$ (< 1), $\hat{\boldsymbol{\xi}}_n^{+}$ is better (worse) than $\tilde{\boldsymbol{\xi}}_n$. The ARRE's of the estimators depend on underlying distribution only through the noncentrality parameter δ^2. For $\delta^2 = 0.0(0.5)20$, the values of the risks for $\hat{\boldsymbol{\xi}}_n^{PS}$ and $\hat{\boldsymbol{\xi}}_n^{KS}$ were estimated from a Monte Carlo simulation with 10,000 repetitions. In the first setting, we restricted our attention to the case that $q = 6$ and $r = 4$. In Fig. 1, we drew the graphical picture of the ARRE's of the estimators $\hat{\boldsymbol{\xi}}_n^{PS}$ and $\hat{\boldsymbol{\xi}}_n^{KS}$ with respect to the estimator $\tilde{\boldsymbol{\xi}}_n$. From Fig. 1, we can see that (i) the improvements of the PSE and the KSE over the FGLSE decrease in δ^2, and (ii) the PSE is better than the KSE for all δ^2 satisfying $0 \leq \delta^2 \leq 2.6$ while the latter is better than the former for $2.8 \leq \delta^2$. Further, we find that, for $q = 2 + r$, the ARRE's of the PSE and the KSE relative to the FGLSE increase in r.

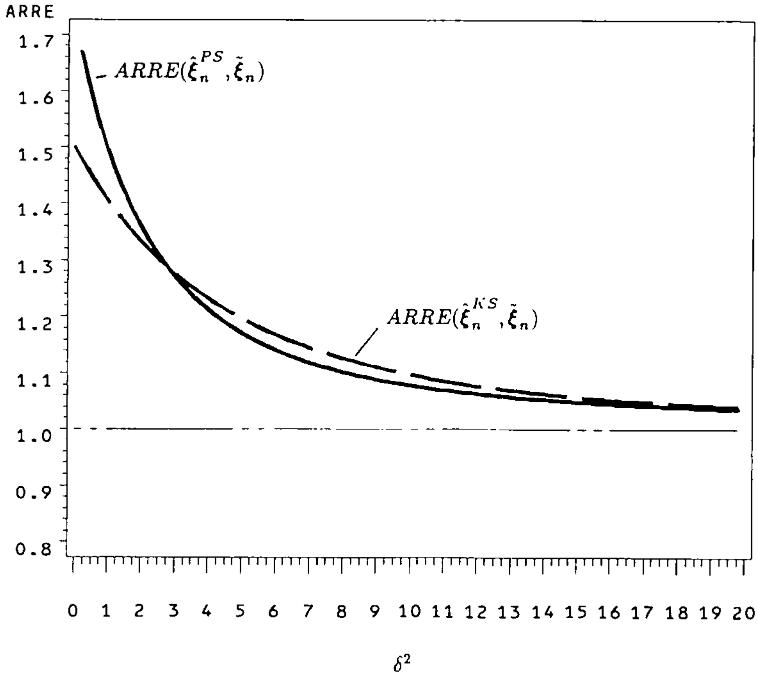

Fig. 1. The asymptotic relative risk efficiency of $\hat{\boldsymbol{\xi}}_n^{PS}$ and $\hat{\boldsymbol{\xi}}_n^{KS}$ with respect to $\hat{\boldsymbol{\xi}}_n$.

4. Normal theory

In this section, we assume that the underlying distribution $F(\boldsymbol{x})$ is a p-variate normal distribution with null mean and known positive-definite variance-covariance matrix $\boldsymbol{\Sigma}$. We state the normal theory shrinkage estimators, which is a general theory of a classical result. Based on $\tilde{\boldsymbol{\xi}}_n^*$ defined by (2.1), estimators are constructed. The weighted combination of the generalized least squares estimator and restricted estimator corresponding to (1.6) is proposed, being of the form

$$\hat{\boldsymbol{\xi}}_n^{*g} = \hat{\boldsymbol{\xi}}_n^* + \{\boldsymbol{I}_q - g(\mathcal{L}_n^*)(\boldsymbol{\Sigma}^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{C}'\boldsymbol{\Gamma}^{*-1}\boldsymbol{C}\}(\tilde{\boldsymbol{\xi}}_n^* - \hat{\boldsymbol{\xi}}_n^*), \tag{4.1}$$

where

$$\begin{aligned}
\hat{\boldsymbol{\xi}}_n^* &= \tilde{\boldsymbol{\xi}}_n^* - (\boldsymbol{\Sigma}^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{C}'\{\boldsymbol{C}(\boldsymbol{\Sigma}^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{C}'\}^{-1}(\boldsymbol{C}\tilde{\boldsymbol{\xi}}_n^* - \boldsymbol{c}_0),\\
\mathcal{L}_n^* &= n(\tilde{\boldsymbol{\xi}}_n^* - \hat{\boldsymbol{\xi}}_n^*)'\boldsymbol{C}'\boldsymbol{\Gamma}^{*-1}\boldsymbol{C}(\tilde{\boldsymbol{\xi}}_n^* - \hat{\boldsymbol{\xi}}_n^*) \quad \text{and}\\
\boldsymbol{\Gamma}^* &= \boldsymbol{C}(\boldsymbol{\Sigma}^{-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{Q}(\boldsymbol{\Sigma}^{*-1}\#\boldsymbol{D}_n)^{-1}\boldsymbol{C}'.
\end{aligned}$$

By setting $g(u) = I(u \le \mathcal{L}_{n,\alpha})$, $(r-2)\cdot u^{-1}$, $1 - \{1-(r-2)\cdot u^{-1}\}\cdot I(u \ge r-2)$, and $\phi(u)\cdot u^{-1}$ in (4.1), we define the preliminary test estimator (PTE) $\hat{\boldsymbol{\xi}}_n^{*PT}$, shrinkage estimator (SE) $\hat{\boldsymbol{\xi}}_n^{*S}$, positive-rule shrinkage estimator (PSE) $\hat{\boldsymbol{\xi}}_n^{*PS}$, and Kubokawa-type shrinkage estimator (KSE) $\hat{\boldsymbol{\xi}}_n^{*KS}$, respectively.

The quadratic risk (QR) of $\hat{\boldsymbol{\xi}}_n^+$ is given by

$$\mathcal{R}(\hat{\boldsymbol{\xi}}_n^+ : \boldsymbol{Q}) = E\{(\hat{\boldsymbol{\xi}}_n^+ - \boldsymbol{\xi})'\boldsymbol{Q}(\hat{\boldsymbol{\xi}}_n^+ - \boldsymbol{\xi})\}.$$

Corresponding to Corollary 3.1 and Theorem 3.3, we get Theorems 4.1 and 4.2 respectively.

THEOREM 4.1. *Suppose that $r \geq 3$. The $\hat{\boldsymbol{\xi}}_n^{*S}$, $\hat{\boldsymbol{\xi}}_n^{*PS}$ and $\hat{\boldsymbol{\xi}}_n^{*KS}$ dominate $\tilde{\boldsymbol{\xi}}_n^*$, namely,*

$$\mathcal{R}(\hat{\boldsymbol{\xi}}_n^{*S} : \boldsymbol{Q}), \quad \mathcal{R}(\hat{\boldsymbol{\xi}}_n^{*PS} : \boldsymbol{Q}), \quad \mathcal{R}(\hat{\boldsymbol{\xi}}_n^{*KS} : \boldsymbol{Q}) \leq \mathcal{R}(\tilde{\boldsymbol{\xi}}_n^* : \boldsymbol{Q}).$$

THEOREM 4.2. *Suppose that $r \geq 3$. For $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1} \# \boldsymbol{D}_n$, namely, when Mahalanobis loss is taken, $\hat{\boldsymbol{\xi}}_n^{*PS}$ and $\hat{\boldsymbol{\xi}}_n^{*KS}$ dominate $\hat{\boldsymbol{\xi}}_n^{*S}$, namely,*

$$\mathcal{R}(\hat{\boldsymbol{\xi}}_n^{*PS} : \boldsymbol{Q}), \quad \mathcal{R}(\hat{\boldsymbol{\xi}}_n^{*KS} : \boldsymbol{Q}) \leq \mathcal{R}(\hat{\boldsymbol{\xi}}_n^{*S} : \boldsymbol{Q}) \leq \mathcal{R}(\tilde{\boldsymbol{\xi}}_n^* : \boldsymbol{Q}).$$

The MDM linear model is an extension of the standard multivariate linear model. Our results stated in this section include those of many papers discussed shrinkage problems. As special cases, we state the following two examples.

Example 1. When $n = 1$, $\boldsymbol{X}_1 = \cdots = \boldsymbol{X}_p = 1$, and $\boldsymbol{\Sigma} = \boldsymbol{C} = \boldsymbol{I}_p$, James and Stein (1961), Sclove *et al.* (1972), and Kubokawa (1994) discussed the results of Theorem 4.2.

Example 2. When $n = 1$, $\boldsymbol{X}_1 = \cdots = \boldsymbol{X}_p = 1$, and $\boldsymbol{C} = \boldsymbol{I}_p$, Berger (1975) and Hudson (1974) discussed the results of Theorem 4.1.

Acknowledgements

The authors would like to thank the referees for their constructive comments on an earlier draft of this article.

REFERENCES

Berger, J. O. (1975). Minimax estimation of location vectors for a wide class of densities, *Ann. Statist.*, **3**, 1318–1328.

Bilodeau, M. and Kariya, T. (1989). Minimax estimators in the normal MANOVA model, *J. Multivariate Anal.*, **28**, 260–270.

Hudson, J. O. (1974). Empirical Bayes estimation, Tech. Report, No. 58, NSF contract GP 30711X-2.

James, W. and Stein, C. (1961). Estimation with quadratic loss, *Proc. 4th Berkley Symp. on Math. Statist. Probab.*, Vol. 1, 361–379, University of California Press, Berkeley.

Kubokawa, T. (1994). A unified approach to improving equivariant estimators, *Ann. Statist.*, **22**, 290–298.

Saleh, A. K. Md. E. and Shiraishi, T. (1992). On improved least squares estimators in multiple-design multivariate linear models under general restriction, *Nonparametric Statistics and Related Topics* (ed. A. K. Md. E. Saleh), 269–279, North-Holland, Amsterdam.

Sclove, S. L., Morris, C. and Radhakrishnan, R. (1972). Non-optimality of preliminary test estimators for the mean of a multivariate normal distribution, *Ann. Math. Statist.*, **43**, 1481–1490.

Srivastava, V. K. and Giles, D. E. A. (1987). *Seemingly Unrelated Regression Equations Models*, Dekker, New York.

Stein, C. (1981). Estimation of the mean of a multivariate normal distribution, *Ann. Statist.*, **9**, 1135–1151.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *J. Amer. Statist. Assoc.*, **58**, 348–368.