A NOTE ON ESTIMATION UNDER THE QUADRATIC LOSS IN MULTIVARIATE CALIBRATION

Hisayuki Tsukuma*

The problem of estimation in multivariate linear calibration with multivariate response and explanatory variables is considered. In this calibration problem two estimators are well-known; one is the classical estimator and the other is the inverse estimator. In this paper we show that the inverse estimator is a proper Bayes estimator under the quadratic loss with respect to a prior distribution which is considered by Kiefer and Schwartz (1965, *Ann. Math. Statist.*, **36**, 747–770) for proving admissibility of the likelihood ratio test about equality of covariance matrices under the normality assumption. We also show that the Bayes risk of the inverse estimator is finite and hence the inverse estimator is admissible under the quadratic loss. Further we consider an improvement on the classical estimator under the quadratic loss. First, the expressions for the first and the second moments of the classical estimator are given with expectation of a function of a noncentral Wishart matrix. From these expressions, we propose an alternative estimator which can be regarded as an extension of an improved estimator derived by Srivastava (1995, *Commun. Statist.-Theory Meth.*, **24**, 2753–2767) and we show, through numerical study, that the alternative estimator performs well as compared with the classical estimator.

Key words and phrases: Admissibility, inverse regression, multivariate linear model, noncentral Wishart distribution, quadratic loss.

1. Introduction

Let x be a $q \times 1$ vector of explanatory variables and y a $p \times 1$ vector of response variables. We assume that

$$(1.1) \quad y = \alpha + \Theta'x + e,$$

where α and Θ are $p \times 1$ and $q \times p$ unknown parameters, respectively, and e is an error vector with a p -variate normal distribution with mean zero vector and unknown covariance matrix Σ , denoted by $N_p(\mathbf{0}, \Sigma)$. Suppose that a training (calibration) sample (y_i, x_i) , $i = 1, \dots, n$, with relation to (1.1) has been given and, furthermore, that we obtain new observations y_{0j} , $j = 1, \dots, m$, corresponding to unknown explanatory vector x_0 from (1.1).

The calibration model for the training sample can be written as

$$(1.2) \quad Y = \mathbf{1}_n \alpha' + X\Theta + \epsilon,$$

where $\mathbf{1}_n$ denotes the $n \times 1$ vector consisting of ones, $Y = (y_1, \dots, y_n)'$ is an $n \times p$ random matrix of response variables, $X = (x_1, \dots, x_n)'$ is an $n \times q$ matrix of

Received November 22, 2001. Revised April 4, 2002. Accepted June 27, 2002.

*Graduate School of Science and Technology, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan.

explanatory variables with full rank, and ϵ is an $n \times p$ error matrix whose rows are independently, identically distributed as $N_p(\mathbf{0}, \Sigma)$. The prediction model corresponding to the observation y_0 can be expressed as

$$(1.3) \quad Y_0 = \mathbf{1}_m \alpha' + \mathbf{1}_m x_0' \Theta + \epsilon_0,$$

where $Y_0 = (y_{01}, \dots, y_{0m})'$ is an $m \times p$ random matrix of response variables and the rows of ϵ_0 are distributed as $N_p(\mathbf{0}, \Sigma)$ and independent of ϵ . We here assume that $p \geq q$, $n - q - 1 \geq p$, $m - q - 1 \geq p$. Our problem is to estimate x_0 in model (1.3) based on the training sample X and Y from model (1.2) and Y_0 from model (1.3).

From the model (1.2), the least squares estimators of α and Θ are given by

$$(1.4) \quad \hat{\alpha} = \bar{y} - \hat{\Theta}' \bar{x}, \quad \hat{\Theta} = [X'(I_n - \mathbf{1}_n \mathbf{1}_n' / n)X]^{-1} X'(I_n - \mathbf{1}_n \mathbf{1}_n' / n)Y,$$

where $\bar{y} = Y' \mathbf{1}_n / n$ and $\bar{x} = X' \mathbf{1}_n / n$. Let

$$(1.5) \quad \bar{y}_0 = Y_0' \mathbf{1}_m / m, \quad S_0 = (Y_0 - \mathbf{1}_m \bar{y}_0')(Y_0 - \mathbf{1}_m \bar{y}_0'), \\ S = (Y - \mathbf{1}_n \hat{\alpha}' - X \hat{\Theta}')(Y - \mathbf{1}_n \hat{\alpha}' - X \hat{\Theta}), \quad \text{and} \quad V = S + S_0.$$

We here note that V is distributed as the Wishart distribution with degrees of freedom $n + m - q - 2$ and scale matrix Σ and that V is independent of $\hat{\alpha}$ and $\hat{\Theta}$. Hence an unbiased estimator of Σ is $\hat{\Sigma} = V / (n + m - q - 2)$.

From (1.3), if the parameter (α, Θ, Σ) is known, then the maximum likelihood estimator of x_0 is $\hat{x}_0 = (\Theta \Sigma^{-1} \Theta')^{-1} \Theta \Sigma^{-1} (\bar{y}_0 - \alpha)$. Here, replacing (α, Θ, Σ) by $(\hat{\alpha}, \hat{\Theta}, V / (n + m - q - 2))$, we get the classical estimator

$$(1.6) \quad \hat{x}_0 = \bar{x} + (\hat{\Theta} V^{-1} \hat{\Theta}')^{-1} \hat{\Theta} V^{-1} (\bar{y}_0 - \bar{y}).$$

Hence in a sense the classical estimator is the maximum likelihood estimator.

On the other hand the inverse estimator is a regression predictor when we regress x on y (see Brown (1982)) and it is given by

$$\check{x}_0 = \bar{x} + \hat{D}' (\bar{y}_0 - \bar{y}),$$

where $\hat{D} = [Y'(I_n - \mathbf{1}_n \mathbf{1}_n' / n)Y]^{-1} Y'(I_n - \mathbf{1}_n \mathbf{1}_n' / n)X$. Here, using the fact that

$$Y'(I_n - \mathbf{1}_n \mathbf{1}_n' / n)Y = S + \hat{\Theta}' X'(I_n - \mathbf{1}_n \mathbf{1}_n' / n)X \hat{\Theta}, \\ \hat{\Theta} = [X'(I_n - \mathbf{1}_n \mathbf{1}_n' / n)X]^{-1} X'(I_n - \mathbf{1}_n \mathbf{1}_n' / n)Y,$$

and applying Lemma 1 (see Appendix A.1), we obtain the inverse estimator

$$(1.7) \quad \check{x}_0 = \bar{x} + X'(I_n - \mathbf{1}_n \mathbf{1}_n' / n)X \hat{\Theta} \{S + \hat{\Theta}' X'(I_n - \mathbf{1}_n \mathbf{1}_n' / n)X \hat{\Theta}\}^{-1} (\bar{y}_0 - \bar{y}) \\ = \bar{x} + \{[X'(I_n - \mathbf{1}_n \mathbf{1}_n' / n)X]^{-1} + \hat{\Theta} S^{-1} \hat{\Theta}'\}^{-1} \hat{\Theta} S^{-1} (\bar{y}_0 - \bar{y}).$$

For $n \rightarrow \infty$ and $m \rightarrow \infty$, the classical estimator (1.6) is consistent when $\Theta \neq \mathbf{0}$ but the inverse estimator (1.7) is not consistent. For details of comparison

between the classical and the inverse estimators see, for example, Brown (1993), Osborne (1991) and Sundberg (1999).

The main interest of this paper is an examination of distinctive features on the classical and the inverse estimators from a decision-theoretic point of view. When $q = 1$, $\Sigma = \sigma^2 I_p$ and σ^2 is unknown in models (1.2) and (1.3), Kubokawa and Robert (1994) showed, under the squared loss, that the classical estimator is inadmissible and that the inverse estimator is admissible. Srivastava (1995) showed the inadmissibility of the classical estimator and the admissibility of the inverse estimator when $q = 1$ and Σ is fully unknown. However there has not been any literature on either admissibility or inadmissibility results of these estimators in models (1.2) and (1.3) when $q > 1$.

This paper is organized in the following manner: First in Section 2, a canonical form of the calibration problem above is constructed. In Section 3 we show the admissibility of the inverse estimator under the quadratic loss. Next in Section 4, we give expressions for the first and the second moments of the classical estimator with the expectation of a function of a noncentral Wishart matrix and we propose an alternative estimator which can be regarded as extension of an improved estimator derived by Srivastava (1995). Through a Monte Carlo simulation, we show that the alternative estimator performs well compared with the classical estimator. Finally in the Appendix we state some technical lemmas and give proofs of Theorems in Sections 3 and 4.

2. Canonical form

In this section, we give a canonical form of the calibration problem. Without loss of generality we assume that $m = 1$ and then, in (1.5), $V = S$.

We first define the following notation. The Kronecker product of matrices A and C is denoted by ' $A \otimes C$ '. For any $q \times p$ matrix $Z = (z_1, \dots, z_q)'$, we write $\text{vec}(Z') = (z_1', \dots, z_q')'$. ' $Z \sim N_{q \times p}(M, A \otimes C)$ ' indicates that $\text{vec}(Z')$ follows multivariate normal distribution with mean $\text{vec}(M')$ and covariance matrix $A \otimes C$. Furthermore, $W_p(\Sigma, k)$ stands for the Wishart distribution with degrees of freedom k and scale matrix Σ .

The classical and the inverse estimators for unknown x_0 can be rewritten as

$$(2.1) \quad \hat{x}_0 = \bar{x} + (\hat{\Theta}S^{-1}\hat{\Theta}')^{-1}\hat{\Theta}S^{-1}(y_0 - \bar{y})$$

and

$$(2.2) \quad \tilde{x}_0 = \bar{x} + [\{X'(I_n - \mathbf{1}_n \mathbf{1}_n'/n)X\}^{-1} + \hat{\Theta}S^{-1}\hat{\Theta}]^{-1}\hat{\Theta}S^{-1}(y_0 - \bar{y}),$$

where $\bar{x}$, $\bar{y}$, $\hat{\Theta}$, and S are given in (1.4) and (1.5). We here note that $\bar{y}$, $\hat{\Theta}$, $\hat{\Sigma}$, and y_0 are mutually and independently distributed as

$$\begin{aligned} \bar{y} &\sim N_p(\alpha + \Theta'\bar{x}, (1/n)\Sigma), & \hat{\Theta} &\sim N_{q \times p}(\Theta, [X'(I_n - \mathbf{1}_n \mathbf{1}_n'/n)X]^{-1} \otimes \Sigma), \\ S &\sim W_p(\Sigma, n - q - 1), & \text{and} & \quad y_0 \sim N_p(\alpha + \Theta'x_0, \Sigma), \end{aligned}$$

for $n - q - 1 \geq p$. The estimators (2.1) and (2.2) can be interpreted as extensions of the *classical* estimator (Eisenhart (1939)) and the *inverse* regression estimator (Krutchkoff (1967)) in univariate linear model, respectively.

Let $B = [X'(I_n - \mathbf{1}_n \mathbf{1}_n'/n)X]^{1/2} \hat{\Theta}$, $z = c_n^{-1/2}(y_0 - \bar{y})$ and $c_n = 1 + 1/n$. Here we denote by $A^{1/2}$ a symmetric matrix such that $A = A^{1/2} A^{1/2}$. Then the distributions of B , S , and z are mutually and independently distributed as

$$(2.3) \quad B \sim N_{q \times p}(\beta, I_q \otimes \Sigma), \quad S \sim W_p(\Sigma, n - q - 1), \quad \text{and} \quad z \sim N_p(\beta' \xi, \Sigma),$$

where $\beta = [X'(I_n - \mathbf{1}_n \mathbf{1}_n'/n)X]^{1/2} \Theta$ and $\xi = c_n^{-1/2} [X'(I_n - \mathbf{1}_n \mathbf{1}_n'/n)X]^{-1/2} (x_0 - \bar{x})$.

To express the estimators (2.1) and (2.2) in terms of B , S and z , we put

$$\hat{\xi} = c_n^{-1/2} [X'(I_n - \mathbf{1}_n \mathbf{1}_n'/n)X]^{-1/2} (\hat{x}_0 - \bar{x})$$

and

$$\check{\xi} = c_n^{-1/2} [X'(I_n - \mathbf{1}_n \mathbf{1}_n'/n)X]^{-1/2} (\check{x}_0 - \bar{x}).$$

Then we have

$$(2.4) \quad \hat{\xi} = (BS^{-1}B')^{-1}BS^{-1}z$$

and

$$(2.5) \quad \check{\xi} = (I_q + BS^{-1}B')^{-1}BS^{-1}z.$$

In this paper we treat the calibration problem on the model (2.3) and discuss the properties of the estimators (2.4) and (2.5).

3. Admissibility of the inverse estimator

In this section we show the admissibility of the inverse estimator (2.5) under the quadratic loss function

$$(3.1) \quad L(\tilde{\xi}, \xi) = (\tilde{\xi} - \xi)'(\tilde{\xi} - \xi),$$

where $\tilde{\xi}$ is an estimator of ξ . The corresponding quadratic risk is given by

$$(3.2) \quad R(\theta, \tilde{\xi}) = E_\theta[L(\tilde{\xi}, \xi)],$$

where $\theta = (\beta, \Sigma, \xi)$ and the expectation is taken with respect to (2.3).

We first show that the inverse estimator is a Bayes estimator for a proper prior distribution. The prior distribution of (β, Σ) is similar to that of Kiefer and Schwartz (1965) for their proving admissibility of the likelihood ratio test about equality of covariance matrices and the prior distribution of ξ is a vector-valued t distribution.

Let

$$(3.3) \quad (\beta, \Sigma) = [(I_q + \xi\xi')^{-1/2} \Delta \Gamma' (I_p + \Gamma\Gamma')^{-1}, (I_p + \Gamma\Gamma')^{-1}],$$

where Δ and Γ are $q \times r$ and $p \times r$ random matrices, respectively. The conditional distribution of Δ given Γ follows $N_{q \times r}(\mathbf{0}, I_q \otimes [I_r - \Gamma'(I_p + \Gamma\Gamma')^{-1}\Gamma]^{-1})$, i.e., the conditional probability density function (abbreviated by 'p.d.f.') is given by

$$(3.4) \quad p_1(\Delta | \Gamma) \propto |I_r - \Gamma'(I_p + \Gamma\Gamma')^{-1}\Gamma|^{q/2} \exp[-2^{-1} \text{tr} \Delta \{I_r - \Gamma'(I_p + \Gamma\Gamma')^{-1}\Gamma\} \Delta'].$$

Further the marginal distribution of Γ is the matrix-variate t distribution whose density is given by

$$(3.5) \quad p_2(\Gamma) \propto |I_p + \Gamma\Gamma'|^{-(n-q)/2}.$$

Note that the p.d.f. (3.5) is integrable provided that $n \geq p+q+r$. The distribution of ξ is the q -variate t distribution with degrees of freedom $r - q$ whose density is given by

$$(3.6) \quad p_3(\xi) \propto (1 + \xi'\xi)^{-r/2}.$$

Now we state main theorems in this section and the proofs put into Appendix A.1. The following theorem is an extension of Section 4 in Srivastava (1995).

THEOREM 1. *Under the quadratic loss (3.1), the inverse estimator $\check{\xi}$ given in (2.5) is a proper Bayes estimator for the priors (3.4)–(3.6).*

Next, to show the admissibility of $\check{\xi}$, we make sure that the Bayes risk is finite.

THEOREM 2. *If $n \geq p+2q+3$ the Bayes risk is finite, and thus the inverse estimator $\check{\xi}$ is admissible.*

4. Improvement on the classical estimator

In this section, we consider an improvement of risk of the classical estimator (2.4) under the quadratic loss (3.1). First, in the next theorem we give the expressions of expectation and risk for the classical estimator (2.4). The proof of the next theorem is postponed to Appendix A.2.

THEOREM 3. *Let W be a random matrix distributed as the noncentral Wishart distribution with degrees of freedom p , scale matrix I_q , and noncentrality parameter matrix $\beta\Sigma^{-1}\beta'$. If $n - p - 2 > 0$, then the expressions of expectation and risk for the classical estimator are expressed as*

$$(4.1) \quad E_{\theta}[\hat{\xi}] = \xi - (p - q - 1)E[W^{-1}]\xi \quad \text{for } p - q - 1 > 0$$

and, for $p - q - 3 > 0$,

$$(4.2) \quad E_{\theta}[(\hat{\xi} - \xi)'(\hat{\xi} - \xi)] = \frac{n - q - 2}{n - p - 2}E[\text{tr } W^{-1}] + \xi' \left\{ E[C_1] + \frac{p - q}{n - p - 2}E[C_2] \right\} \xi,$$

where

$$\begin{aligned} C_1 &= \{(p - q - 1)(p - q - 2) + 2\}W^{-2} \\ &\quad - 2(p - q - 1)(\text{tr } W^{-1})W^{-1} + (\text{tr } W^{-1})I_q, \\ C_2 &= 2W^{-2} - (p - q - 1)(\text{tr } W^{-1})W^{-1} + (\text{tr } W^{-1})I_q. \end{aligned}$$

From the expression (4.2) for risk, it seems that the risk of the classical estimator is small if the noncentrality parameter matrix $\beta\Sigma^{-1}\beta'$ is large (with matrix argument) and that the risk of the classical estimator is large otherwise.

Hence, when $\beta\Sigma^{-1}\beta'$ is large, we should use another estimator instead of the classical estimator.

From the expression (4.1) for expectation of $\hat{\xi}$ in Theorem 3, we see that the bias of the classical estimator is $-E[W^{-1}]\xi/(p-q-1)$. Hence, replacing $E[W^{-1}]$ by $(BS^{-1}B)^{-1}/(n-p-1)$ and ξ by $\hat{\xi}$, we may propose a bias-corrected estimator

$$\hat{\xi}^{BC} = \left(I_q + \frac{p-q-1}{n-p-1} (BS^{-1}B')^{-1} \right) \hat{\xi}.$$

When $q = 1$ and Σ is unknown in model (2.3), Srivastava (1995) showed that the classical estimator is inadmissible under the squared loss and derived an improved estimator of the form

$$(4.3) \quad \hat{\xi}^{SR} = \min \left(\frac{1}{BS^{-1}B'}, \frac{c}{1 + BS^{-1}B'} \right) BS^{-1}z,$$

where c is a suitable constant. Using Theorem 2.3 of Kubokawa and Robert (1994), Srivastava (1995) proved that $\hat{\xi}^{SR}$ dominates the classical estimator.

When $q > 1$, on the analogy of (4.3), we propose an alternative estimator of the form

$$(4.4) \quad \hat{\xi}^{AC} = \begin{cases} \alpha(I_q + BS^{-1}B')^{-1}BS^{-1}z, & \text{if } 1/l_{\max} > \alpha/(1 + l_{\max}), \\ (BS^{-1}B')^{-1}BS^{-1}z, & \text{otherwise,} \end{cases}$$

where $l_{\max}$ is the maximum eigenvalue of $BS^{-1}B'$ and α is a constant. Since it is expected that the matrix $\beta\Sigma^{-1}\beta'$ is small if $l_{\max}$ is sufficiently small, then we should take $\hat{\xi}^{AC} = \alpha(I_q + BS^{-1}B')^{-1}BS^{-1}z$ instead of the classical estimator. However it is difficult to evaluate the risk of the alternative estimator (4.4) analytically.

We also consider an estimator of the form

$$(4.5) \quad \hat{\xi}^{GE} = (kI_q + BS^{-1}B')^{-1}BS^{-1}z,$$

where k is a nonnegative constant. This estimator is extended to the generalized inverse regression estimator proposed by Miwa (1985) and Takeuchi (1997). However, from numerical studies when $p = q = 1$ in Miwa (1985) and when $p > 1$ and $q = 1$ in Takeuchi (1997), the estimator (4.5) is not expected to dominate the classical estimator (2.4) for the whole parameter space.

Remark 1. The expectation and the risk of the classical estimator are finite if $p-q-1 \geq 0$ and $p-q-2 \geq 0$, respectively (see Nishii and Krishnaiah (1988)). In Theorem 3, $p-q-1 > 0$ in (4.1) and $p-q-3 > 0$ in (4.2) are the conditions that we express the expectation and the risk by using functions of a noncentral Wishart matrix.

Remark 2. To establish the inadmissibility of the classical estimator, the author also tried to compare risk of the classical estimator with that of the

generalized regression estimator. However the inadmissibility of the classical estimator under the quadratic loss could not be established. It is difficult to evaluate the expectations of the noncentral distribution with matrix argument when the risks of the estimators are compared.

Numerical studies.

We have carried out Monte Carlo simulations in order to investigate the risk performances of the classical estimator $\hat{\xi}$, the alternative estimator $\hat{\xi}^{AC}$, and the generalized regression estimator $\hat{\xi}^{GE}$. Our simulations are based on 10,000 independent replications. For the simulations, we take $n = 15$, $p = 5$ and $q = 2$ and we also put $\alpha = (n - 1)/(n - p - 1)$ for $\hat{\xi}^{AC}$ and $k = 0.5$ for $\hat{\xi}^{GE}$. The estimated risks when $\xi = (1, 1)'$ and when $\xi = (-\sqrt{2}, 0)'$ are given in Tables 1 and 2, respectively. In Tables 1 and 2, 'CL', 'AC', 'IN', and 'GE' denote the classical, the alternative, the inverse, and the generalized regression estimators, respectively, and their estimated standard deviations are in parentheses. We suppose that the parameter $\beta\Sigma^{-1}\beta'$ is the diagonal matrix with typical elements.

Our simulations suggest that the alternative estimator given in (4.4) is as good as the classical estimator, and that the alternative estimator substantially reduces the risk when diagonal elements of $\beta\Sigma^{-1}\beta'$ are small. Therefore, in spite of a simple extension of the estimator given in (4.3), the results in Tables 1 and 2 indicate that our estimator performs better than the classical estimator under the quadratic loss (3.1).

Further we observe that the generalized regression estimator does not uniformly improve to the classical estimator. However the generalized regression estimator has a smaller risk than either of the classical and the inverse estimators. Hence, as a result of estimating ξ with the generalized regression estimator,

Table 1. Estimated risks when $\xi = (1, 1)'$

$\beta\Sigma^{-1}\beta'$	CL	AC	IN	GE
diag(1, 1)	2.9016 (0.0445)	2.0979 (0.0232)	1.8637 (0.0065)	1.8774 (0.0090)
diag(1, 10)	2.0505 (0.0468)	1.6831 (0.0320)	1.3496 (0.0064)	1.3023 (0.0085)
diag(1, 100)	1.7725 (0.0394)	1.7725 (0.0394)	0.9932 (0.0058)	1.0188 (0.0082)
diag(1, 0.1)	3.1116 (0.0454)	2.2066 (0.0171)	1.9655 (0.0064)	2.0099 (0.0089)
diag(10, 10)	0.8201 (0.0184)	0.7677 (0.0165)	0.8114 (0.0049)	0.6796 (0.0056)
diag(10, 0.1)	2.4649 (0.0500)	2.0601 (0.0434)	1.4578 (0.0066)	1.4436 (0.0089)
diag(100, 100)	0.0823 (0.0009)	0.0823 (0.0009)	0.0857 (0.0008)	0.0792 (0.0008)
diag(100, 0.01)	2.3534 (0.0528)	2.3534 (0.0528)	1.1059 (0.0060)	1.1795 (0.0087)

Table 2. Estimated risks when $\xi = (-\sqrt{2}, 0)'$

$\beta\Sigma^{-1}\beta'$	CL	AC	IN	GE
diag(1, 1)	2.9045 (0.0515)	2.0955 (0.0347)	1.8528 (0.0067)	1.8677 (0.0094)
diag(1, 10)	2.2784 (0.0340)	2.0783 (0.0282)	1.8345 (0.0065)	1.8135 (0.0090)
diag(1, 100)	2.0645 (0.0486)	2.0645 (0.0486)	1.7840 (0.0062)	1.7301 (0.0087)
diag(1, 0.1)	3.0958 (0.0802)	2.0794 (0.0202)	1.8545 (0.0067)	1.8756 (0.0094)
diag(10, 10)	0.8353 (0.0186)	0.7634 (0.0162)	0.7964 (0.0049)	0.6671 (0.0054)
diag(10, 0.1)	2.2922 (0.0608)	1.6445 (0.0495)	0.8467 (0.0054)	0.7961 (0.0070)
diag(100, 100)	0.0833 (0.0010)	0.0833 (0.0010)	0.0848 (0.0008)	0.0791 (0.0008)
diag(100, 0.01)	2.0328 (0.0604)	2.0328 (0.0604)	0.1825 (0.0026)	0.2988 (0.0047)

the risks resulting from our use of the classical and the inverse estimators seem to be reduced.

For another values of ξ in the consideration of directions, for example $\xi = (-1, 1)'$ and $\xi = (0, -\sqrt{2})'$, we have carried out Monte Carlo simulations with the same values of $\beta\Sigma^{-1}\beta'$ in Tables 1 and 2, and we obtained results which are similar to those in Tables 1 and 2.

5. Concluding remarks

In this paper, we showed the admissibility of the inverse estimator and we proposed an alternative estimator over the classical estimator. However, the following problems remain to be solved: (i) Is the inverse estimator the proper Bayes for a prior distribution on (β, Σ, ξ) in the canonical form (2.3) such that the prior of ξ and that of (β, Σ) are mutually independent as ordinary Bayesian situations on the calibration problem? (see Brown (1982, 1993)); (ii) How do we give analytical proof for the inadmissibility of the classical estimator?

Recently Branco *et al.* (2000) considered a calibration problem with an elliptical error and showed that the inverse estimator is a Bayes estimator. However, since the prior distributions of parameters on their model are improper, it is not known whether the inverse estimator with elliptical error is admissible or not. Therefore it will be important to continue study of these problems in the future.

Appendix

A.1 Proof of Theorems 1 and 2

For proofs of Theorems 1 and 2, we list some lemmas.

LEMMA 1. Let A be a $p \times p$ nonsingular matrix and let B and C be $q \times p$ matrices. If $A + B'C$ and $I_q + CA^{-1}B'$ are nonsingular, then

$$(A + B'C)^{-1} = A^{-1} - A^{-1}B'(I_q + CA^{-1}B')^{-1}CA^{-1}$$

and

$$(A + B'C)^{-1}B' = A^{-1}B'(I_q + CA^{-1}B')^{-1}.$$

LEMMA 2. Let A be a $p \times p$ nonsingular matrix and y a $p \times 1$ vector. Then $|A + yy'| = |A|(1 + y'A^{-1}y)$.

LEMMA 3 (Khatrı (1966)). Let B be a $q \times p$ matrix with rank q , where $p > q$. Also, let B_1 a $(p - q) \times p$ matrix with rank $p - q$ such that $B_1B' = 0$. If S is a symmetric positive definite matrix, then $S^{-1} - S^{-1}B'(BS^{-1}B')^{-1}BS^{-1} = B_1'(B_1SB_1')^{-1}B_1$.

LEMMA 4 (Anderson and Takemura (1982)). Let B_1 and B_2 be $p \times q$ matrices. If both A_1 and A_2 are $p \times p$ positive definite matrices, then

$$(B_1 + B_2)'(A_1 + A_2)^{-1}(B_1 + B_2) \leq B_1'A_1^{-1}B_1 + B_2'A_2^{-1}B_2.$$

PROOF OF THEOREM 1. From (2.3), the joint p.d.f. of B , S , and z is

$$(A.1) \quad L(B, S, z \mid \beta, \Sigma, \xi) \\ \propto |\Sigma^{-1}|^{n/2} \exp[-\text{tr} \Sigma^{-1}\{W - \beta'(B + \xi z') \\ -(B + \xi z')'\beta + \beta'(I_q + \xi \xi')\beta\}/2],$$

where $W = S + B'B + zz'$.

First, from (3.3)–(3.6) and (A.1), we can write the posterior density of (ξ, β, Σ) given the data $D = (B, S, z)$ as

$$p(\xi, \beta, \Sigma \mid D) \\ \propto p_3(\xi) |I_p + \Gamma \Gamma'|^{-(n-q)/2} |I_r - \Gamma'(I_p + \Gamma \Gamma')^{-1} \Gamma|^{q/2} \\ \times \exp[-\text{tr} \Delta \{I_r - \Gamma'(I_p + \Gamma \Gamma')^{-1} \Gamma\} \Delta' / 2] \times |I_p + \Gamma \Gamma'|^{n/2} \\ \times \exp[-\text{tr} (I_p + \Gamma \Gamma') \{W - \tilde{\beta}'(B + \xi z') - (B + \xi z')' \tilde{\beta} + \tilde{\beta}'(I_q + \xi \xi') \tilde{\beta}\} / 2],$$

where $W = S + B'B + zz'$ and $\tilde{\beta} = (I_q + \xi \xi')^{-1/2} \Delta \Gamma' (I_p + \Gamma \Gamma')^{-1}$. Then it can be seen from simple calculation that

$$(A.2) \quad p(\xi, \beta, \Sigma \mid D) \\ \propto p_3(\xi) |I_p + \Gamma \Gamma'|^{q/2} |I_r - \Gamma'(I_p + \Gamma \Gamma')^{-1} \Gamma|^{q/2} \\ \times \exp[-\text{tr} W / 2] \exp[-\text{tr} \Delta \Delta' / 2] \\ \times \exp[-\text{tr} \{\Gamma' W \Gamma - \Gamma'(B + \xi z')'(I_q + \xi \xi')^{-1/2} \Delta \\ - ((B + \xi z')'(I_q + \xi \xi')^{-1/2} \Delta)' \Gamma\} / 2].$$

We here note that W depends on the data (B, S, z) only. Hence we can omit the factor $\exp[-2^{-1}\text{tr} W]$ in (A.2). Define $\tilde{\Gamma} = W^{-1}(B + \xi z')(I_q + \xi \xi')^{-1/2} \Delta$ to have

$$\begin{aligned} \Gamma' W \Gamma - \Gamma'(B + \xi z')(I_q + \xi \xi')^{-1/2} \Delta - ((B + \xi z')(I_q + \xi \xi')^{-1/2} \Delta)' \Gamma \\ = (\Gamma - \tilde{\Gamma})' W (\Gamma - \tilde{\Gamma}) \\ - \Delta'(I_q + \xi \xi')^{-1/2} (B + \xi z') W^{-1} (B + \xi z')(I_q + \xi \xi')^{-1/2} \Delta. \end{aligned}$$

From this equation and the relation $|I_r - \Gamma'(I_p + \Gamma \Gamma')^{-1} \Gamma| = |I_p + \Gamma \Gamma'|^{-1}$, we get

$$\begin{aligned} p(\xi, \beta, \Sigma | D) \\ \propto p_3(\xi) \exp[-\text{tr} \Delta \Delta' / 2] \exp[-\text{tr} (\Gamma - \tilde{\Gamma})' W (\Gamma - \tilde{\Gamma}) / 2] \\ \times \exp[\text{tr} \Delta'(I_q + \xi \xi')^{-1/2} (B + \xi z') W^{-1} (B + \xi z')(I_q + \xi \xi')^{-1/2} \Delta / 2]. \end{aligned}$$

Next, integrating out Γ , we can express the posterior density of (ξ, Δ) as

$$(A.3) \quad p(\xi, \Delta | D) \propto p_3(\xi) \exp[-\text{tr} \Delta \Delta' \{I_q - (I_q + \xi \xi')^{-1/2} (B + \xi z') W^{-1} (B + \xi z')(I_q + \xi \xi')^{-1/2}\} / 2].$$

We put $W = S + B' B + z z' \equiv V + z z'$ and use Lemma 1 to obtain

$$\begin{aligned} z' W^{-1} B' &= (1 + z' V^{-1} z)^{-1} z' V^{-1} B', \\ z' W^{-1} z &= z' V^{-1} z (1 + z' V^{-1} z)^{-1}, \\ z' V^{-1} B' &= z' S^{-1} B' (I_q + B S^{-1} B')^{-1} = \check{\xi}'. \end{aligned}$$

Then the matrix in the braces in the right-hand side (r.h.s.) of (A.3) can be replaced by

$$(I_q + \xi \xi')^{-1/2} [D_1 + (1 + z' V^{-1} z)^{-1} (\xi - \check{\xi})(\xi - \check{\xi})'] (I_q + \xi \xi')^{-1/2},$$

where $D_1 = I_q - B W^{-1} B' - \check{\xi} \check{\xi}' (1 + z' V^{-1} z)^{-1}$.

Hence, integrating out the r.h.s. in (A.3) with respect to Δ and applying Lemma 2, we have

$$\begin{aligned} p(\xi | D) \\ \propto (1 + \xi' \xi)^{-r/2} | (I_q + \xi \xi')^{-1/2} [D_1 + (1 + z' V^{-1} z)^{-1} (\xi - \check{\xi})(\xi - \check{\xi})'] \\ \cdot (I_q + \xi \xi')^{-1/2} |^{-r/2} \\ \propto \{1 + (\xi - \check{\xi})' D_2^{-1} (\xi - \check{\xi})\}^{-r/2}, \end{aligned}$$

where $D_2 = (1 + z' V^{-1} z) D_1$. Since the posterior distribution of ξ is q -variate t distribution with mean $\check{\xi}$ with respect to the proper priors (3.4)–(3.6), we can see that $\check{\xi}$ is a proper Bayes estimator.

PROOF OF THEOREM 2. Taking expectation of the loss function with respect to z , we can write the quadratic risk (3.2) as

$$\begin{aligned} (A.4) \quad R(\theta, \check{\xi}) &= \text{tr} E_\theta [S^{-1} B' (I_q + B S^{-1} B')^{-2} B S^{-1}] \Sigma \\ &\quad + \xi' E_\theta \{ \{ (I_q + B S^{-1} B')^{-1} B S^{-1} \beta' - I_q \}' \\ &\quad \cdot \{ (I_q + B S^{-1} B')^{-1} B S^{-1} \beta' - I_q \} \} \xi \\ &\equiv \text{tr} E_1 + \xi' E_2 \xi, \end{aligned}$$

where $E_1 = E_\theta[\Sigma^{1/2}S^{-1}B'(I_q + BS^{-1}B')^{-2}BS^{-1}\Sigma^{1/2}]$ and $E_2 = E_\theta[\{I_q + BS^{-1}(B - \beta)'\}'(I_q + BS^{-1}B')^{-2}\{I_q + BS^{-1}(B - \beta)'\}']$. Applying the inequality $(I_q + BS^{-1}B')^{-2} \leq (BS^{-1}B')^{-1}$ (' $A \leq B$ ' indicates that ' $B - A$ ' is positive semidefinite) to $\text{tr } E_1$, we obtain

$$\text{tr } E_1 \leq \text{tr } E_\theta[\Sigma^{1/2}S^{-1}B'(BS^{-1}B')^{-1}BS^{-1}\Sigma^{1/2}].$$

Here it follows from Lemma 3 that

$$(A.5) \quad S^{-1} - S^{-1}B'(BS^{-1}B')^{-1}BS^{-1} \geq 0.$$

Thus, since $S \sim W_p(\Sigma, n - q - 1)$, we get

$$\text{tr } E_1 \leq \text{tr } E_\theta[\Sigma^{1/2}S^{-1}\Sigma^{1/2}] = p/(n - p - q - 2)$$

for $n \geq p + q + 3$.

We next evaluate $\xi'E_2\xi$ in the r.h.s. of (A.4). Using the fact that $(I_q + BS^{-1}B')^{-2} \leq (I_q + BS^{-1}B')^{-1}$ and using Lemma 4, we have

$$(A.6) \quad \begin{aligned} \xi'E_2\xi &\leq \xi'E_\theta[\{I_q + BS^{-1}(B - \beta)'\}'(I_q + BS^{-1}B')^{-1} \\ &\quad \cdot \{I_q + BS^{-1}(B - \beta)'\}']\xi \\ &\leq \xi'E_\theta[I_q + (B - \beta)S^{-1}B'(BS^{-1}B')^{-1}BS^{-1}(B - \beta)']\xi. \end{aligned}$$

Applying (A.5) to the matrix of the second term in brackets of the r.h.s. of (A.6) and noting that $B \sim N_{q \times p}(\beta, I_q \otimes \Sigma)$ being independent of S , we can see that

$$\begin{aligned} \xi'E_2\xi &\leq \xi'E_\theta[I_q + (B - \beta)S^{-1}(B - \beta)']\xi \\ &= \{(n - q - 2)/(n - p - q - 2)\}\xi'\xi. \end{aligned}$$

Hence the quadratic risk (A.4) can be evaluated as

$$(A.7) \quad R(\theta, \tilde{\xi}) \leq \{p + (n - q - 2)\xi'\xi\}/(n - p - q - 2)$$

for $n \geq p + q + 3$.

We finally make sure of the finiteness of the Bayes risk. Using the inequality (A.7), we find that the Bayes risk is finite if

$$\int \xi'\xi p_3(\xi) d\xi < \infty,$$

where $p_3(\xi)$ is the prior density of ξ . Since the prior distribution of ξ is the q -variate t distribution with degrees of freedom $r - q$, the l.h.s. of the above inequality is finite if $r - q > 2$. Hence, combining $r - q > 2$ and the condition for the integrability of the prior distribution (3.5), i.e., $n \geq p + q + r$, we get $n \geq p + 2q + 3$.

A.2 Proof of Theorem 3

In this section we give the proof of Theorem 3. First we define some notation to prove Theorem 3 and next we list some lemmas.

Let A be a $q \times p$ random matrix distributed as $N_{q \times p}(M, I_q \otimes I_p)$. Denote $W = AA'$. Furthermore let P be a $p \times q$ matrix whose elements are functions of A . Also let g and G be, respectively, a scalar function of W and a $q \times q$ matrix-valued function of W . Denote differential operators in terms of $A = (A_{ij})$ and $W = (W_{ij})$ by, respectively,

$$\nabla_A = (\partial/\partial A_{ij}) \quad \text{and} \quad D_W = \left(\frac{1 + \delta_{ij}}{2} \frac{\partial}{\partial W_{ij}} \right).$$

The actions of ∇_A on $P = (P_{ij})$ and of D_W on g and $G = (G_{ij})$ are defined as

$$\nabla_A P = \left(\sum_{k=1}^p \frac{\partial P_{kj}}{\partial A_{ik}} \right), \quad D_W g = \left(\frac{1 + \delta_{ij}}{2} \frac{\partial g}{\partial W_{ij}} \right),$$

and

$$D_W G = \left(\sum_{k=1}^q \frac{1 + \delta_{ik}}{2} \frac{\partial G_{kj}}{\partial W_{ik}} \right),$$

where δ_{ij} is Kronecker's delta.

Let $\tilde{D}_W$ be a $q \times q$ matrix whose elements are linear combinations of $\partial/\partial W_{ij}$ ($i = 1, \dots, q, j = 1, \dots, q$). Also, let G and H be $q \times q$ matrices whose elements are functions of W . Then we have the following lemma which is due to Haff (1981).

LEMMA 6.

$$\tilde{D}_W GH = \{\tilde{D}_W G\}H + (G'\tilde{D}_W)'H.$$

Let $\tilde{\nabla}_A$ be a $q \times p$ matrix whose elements are linear combinations of $\partial/\partial A_{ij}$ ($i = 1, \dots, q, j = 1, \dots, p$). Furthermore, let P and Q be, respectively, $p \times q$ and $q \times q$ matrices whose elements are functions of A . Then we have

$$(A.8) \quad \tilde{\nabla}_A PQ = \{\tilde{\nabla}_A P\}Q + (P'\tilde{\nabla}_A)'Q,$$

as similar to Lemma 6.

We next list some equalities with respect to the operators ∇_A and D_W .

LEMMA 7. Let M be a $q \times p$ constant matrix. Also let P, Q, G , and H be the same as defined above. Then we have

- (i) $\nabla_A A' G Q = p G Q + \{(A \nabla_A')' G\} Q + (G' A \nabla_A')' Q$,
- (ii) $(P' \nabla_A')' A (A - M)' = P' (A - M)' + (\text{tr} AP) I_q$,
- (iii) $(P' \nabla_A')' G = 2(P' A' D_W)' G$.

PROOF. (i): The proof follows from (A.8) and the component-wise calculation.

(ii): The proof follows from the component-wise calculation.

(iii): Let $P = (P_{ij})$ and $G = (G_{ij})$. Then, by chain rule, we can write

$$\begin{aligned}
[(P'\nabla'_A)'G]_{ij} &= \sum_{k,l} P_{kl} \frac{\partial}{\partial A_{ik}} G_{lj} \\
&= \sum_{k,l} P_{kl} \sum_{m,n} \frac{1 + \delta_{mn}}{2} \frac{\partial G_{lj}}{\partial W_{mn}} \frac{\partial W_{mn}}{\partial A_{ik}}.
\end{aligned}$$

Hence, from $W_{mn} = \sum_{a=1}^p A_{ma}A_{na}$, the last l.h.s. above can be expressed as

$$2 \sum_{k,l,m} P_{kl} A_{mk} \frac{1 + \delta_{im}}{2} \frac{\partial G_{lj}}{\partial W_{im}} = 2[(P'A'D_W)'G]_{ij}.$$

LEMMA 8. Let G be a $q \times q$ matrix. Then we have

- (i) $(GD_W)'W^{-1} = -((\text{tr} G'W^{-1})W^{-1} + W^{-1}G'W^{-1})/2$,
- (ii) $(GD_W)'W^{-2} = -((\text{tr} G'W^{-1})W^{-2} + (\text{tr} G'W^{-2})W^{-1} + W^{-2}G'W^{-1} + W^{-1}G'W^{-2})/2$,
- (iii) $(GD_W)'(\text{tr} W^{-1}) = -W^{-2}G'$,
- (iv) $(GD_W)'W^{-1}(\text{tr} W^{-1}) = -(\text{tr} W^{-1})(\text{tr} G'W^{-1})W^{-1} + W^{-1}G'W^{-1})/2 - W^{-2}G'W^{-1}$.

PROOF. (i): Denote $W^{-1} = (W^{ij})$. Using the fact that

$$\frac{1}{2}(1 + \delta_{ia}) \frac{\partial}{\partial W_{ia}} W^{bc} = -(W^{ba}W^{ic} + W^{ib}W^{ac})/2,$$

we have

$$\begin{aligned}
[(GD_W)'W^{-1}]_{ij} &= \sum_{a,b} G_{ba} \frac{1 + \delta_{ia}}{2} \frac{\partial}{\partial W_{ia}} W^{bj} \\
&= - \sum_{a,b} G_{ba} (W^{ba}W^{ij} + W^{ib}W^{aj})/2 \\
&= -(\text{tr} G'W^{-1})[W^{-1}]_{ij}/2 - [W^{-1}G'W^{-1}]_{ij}/2.
\end{aligned}$$

(ii): From Lemma 6, we have

$$(GD_W)'W^{-2} = \{(GD_W)'W^{-1}\}W^{-1} + (W^{-1}GD_W)'W^{-1}.$$

Hence, from the component-wise calculation, we get (ii).

(iii): The proof follows from $\text{tr} W^{-1} = \sum_a W^{aa}$ and the component-wise calculation.

(iv): Using Lemma 6 and applying (i) and (iii), we have (iv).

LEMMA 9 (Bilodeau and Kariya (1989)). Let $A \sim N_{q \times p}(M, I_q \otimes I_p)$. Also let P be a $p \times q$ random matrix whose elements are functions of A . If the conditions of Bilodeau and Kariya (1989) hold, then

$$E[(A - M)P] = E[\nabla_A P].$$

LEMMA 10. Let $A \sim N_{q \times p}(M, I_q \otimes I_p)$ and let $W = AA'$. Then

- (i) $E[(A - M)A'(AA')^{-1}] = E[(p - q - 1)W^{-1}]$,
- (ii) $E[(A - M)A'(AA')^{-2}] = E[(p - q - 2)W^{-2} - (\text{tr} W^{-1})W^{-1}]$,
- (iii) $E[(A - M)A'(AA')^{-1}(\text{tr}(AA')^{-1})] = E[(p - q - 1)(\text{tr} W^{-1})W^{-1} - 2W^{-2}]$.

PROOF. Using Lemma 9, Lemma 7 (i) and (iii), we can write the l.h.s. of (i)–(iii) above as, respectively,

$$\begin{aligned}
 E[\nabla_A A'(AA')^{-1}] &= E[p(AA')^{-1} + (A\nabla'_A)'(AA')^{-1}] \\
 &= E[pW^{-1} + 2(WD_W)'W^{-1}], \\
 E[\nabla_A A'(AA')^{-2}] &= E[p(AA')^{-2} + (A\nabla'_A)'(AA')^{-2}] \\
 &= E[pW^{-2} + 2(WD_W)'W^{-2}], \\
 E[\nabla_A A'(AA')^{-1}(\text{tr}(AA')^{-1})] \\
 &= E[p(\text{tr}(AA')^{-1})(AA')^{-1} + (A\nabla'_A)'(AA')^{-1}(\text{tr}(AA')^{-1})] \\
 &= E[p(\text{tr} W^{-1})W^{-1} + 2(WD_W)'W^{-1}(\text{tr} W^{-1})].
 \end{aligned}$$

Thus, applying Lemma 8, we can get the desired results.

LEMMA 11. Let $A \sim N_{q \times p}(M, I_q \otimes I_p)$ and let $W = AA'$. Then we have

- (i) $E[(\text{tr}(AA')^{-1})(A - M)(A - M)']$
 $= E[p(\text{tr} W^{-1})I_q - 2(p - q - 2)W^{-2} + 2(\text{tr} W^{-1})W^{-1}],$
- (ii) $E[(\text{tr}(AA')^{-1})(A - M)A'(AA')^{-1}A(A - M)']$
 $= E[((p - q)(p - q - 1) + 2)(\text{tr} W^{-1})W^{-1}$
 $- 4(p - q - 1)W^{-2} + q(\text{tr} W^{-1})I_q],$
- (iii) $E[(A - M)A'(AA')^{-2}A(A - M)']$
 $= E[((p - q - 1)(p - q - 2) + 2)(\text{tr} W^{-1})W^{-1}$
 $- 2(p - q - 1)W^{-2} + (\text{tr} W^{-1})I_q].$

PROOF. (i): From Lemma 9 and Lemma 7 (i), it follows that

$$\begin{aligned}
 \text{(A.9)} \quad E[(\text{tr}(AA')^{-1})(A - M)(A - M)'] \\
 &= E[\nabla_A(A - M)'(\text{tr}(AA')^{-1})] \\
 &= E[p(\text{tr}(AA')^{-1})I_q + ((A - M)\nabla'_A)' \text{tr}(AA')^{-1}].
 \end{aligned}$$

Here, using Lemma 7 (iii) and Lemma 8 (iii), we have

$$\begin{aligned}
 \text{(A.10)} \quad E[((A - M)\nabla'_A)' \text{tr}(AA')^{-1}] &= 2E[((A - M)A'D'_W)'(\text{tr} W^{-1})] \\
 &= -2E[(AA')^{-2}A(A - M)'].
 \end{aligned}$$

Hence, combining (A.9) and (A.10) and applying Lemma 10 (ii), we get

$$\begin{aligned}
 E[(\text{tr}(AA')^{-1})(A - M)(A - M)'] \\
 &= E[p(\text{tr}(AA')^{-1})I_q - 2(AA')^{-2}A(A - M)'] \\
 &= E[p(\text{tr} W^{-1})I_q - 2(p - q - 2)W^{-2} + 2(\text{tr} W^{-1})W^{-1}].
 \end{aligned}$$

(ii): Similarly, from Lemma 9 and Lemma 7 (i), we can write the l.h.s. of (ii) as

$$\begin{aligned}
 (A.11) \quad & E[(\text{tr}(AA')^{-1})(A - M)A'(AA')^{-1}A(A - M)'] \\
 &= E[\nabla_A A'(AA')^{-1}A(A - M)'(\text{tr}(AA')^{-1})] \\
 &= E[p(AA')^{-1}A(A - M)'(\text{tr}(AA')^{-1}) \\
 &\quad + \{(A\nabla'_A)'(AA')^{-1}(\text{tr}(AA')^{-1})\}A(A - M)' \\
 &\quad + (\text{tr}(AA')^{-1})((AA')^{-1}A\nabla_A)'A(A - M)'].
 \end{aligned}$$

From Lemma 10 (iii), the first term of the last r.h.s. in (A.11) can be expressed as

$$\begin{aligned}
 (A.12) \quad & E[p(AA')^{-1}A(A - M)'(\text{tr}(AA')^{-1})] \\
 &= E[p(p - q - 1)(\text{tr}W^{-1})W^{-1} - 2pW^{-2}].
 \end{aligned}$$

Next for the second term of the r.h.s. in (A.11), we use Lemma 7 (iii) and Lemma 8 (iii) to give

$$\begin{aligned}
 & E[\{(A\nabla'_A)'(AA')^{-1}(\text{tr}(AA')^{-1})\}A(A - M)'] \\
 &= E[2\{(WD_W)'W^{-1}(\text{tr}W^{-1})\}A(A - M)'] \\
 &= E[-(q + 1)(\text{tr}(AA')^{-1})(AA')^{-1}A(A - M)' - 2(AA')^{-2}A(A - M)'].
 \end{aligned}$$

Thus, applying Lemma 10 (ii) and (iii) to the last r.h.s. above, we obtain

$$\begin{aligned}
 (A.13) \quad & E[\{(A\nabla'_A)'(AA')^{-1}(\text{tr}(AA')^{-1})\}A(A - M)'] \\
 &= E[-((q + 1)(p - q - 1) - 2)(\text{tr}W^{-1})W^{-1} - 2(p - 2q - 3)W^{-2}].
 \end{aligned}$$

Next, from Lemma 7 (ii), the last term of the r.h.s. in (A.11) can be rewritten as

$$\begin{aligned}
 & E[(\text{tr}(AA')^{-1})((AA')^{-1}A\nabla_A)'A(A - M)'] \\
 &= E[(\text{tr}(AA')^{-1})(AA')^{-1}A(A - M)' + q(\text{tr}(AA')^{-1})I_q].
 \end{aligned}$$

Using Lemma 10 (iii), we have

$$\begin{aligned}
 (A.14) \quad & E[(\text{tr}(AA')^{-1})((AA')^{-1}A\nabla_A)'A(A - M)'] \\
 &= E[(p - q - 1)(\text{tr}W^{-1})W^{-1} - 2W^{-2} + q(\text{tr}W^{-1})I_q].
 \end{aligned}$$

Finally, combining (A.12)–(A.14), we obtain the expression (ii).

(iii): The proof is similar to that of (ii) and is omitted.

For moments of the classical estimator, taking expectation with respect to S and z , we have the following lemma, which is due to Fujikoshi and Nishii (1986).

LEMMA 12. *Let $A \sim N_{q \times p}(M, I_q \otimes I_p)$ and $M = \beta\Sigma^{-1/2}$. Then the expectation and the risk for the classical estimator can be expressed as, respectively,*

$$E_\theta[\hat{\xi}] = E[(AA')^{-1}AM']\xi$$

and

$$\begin{aligned} E_{\theta}[(\hat{\xi} - \xi)'(\hat{\xi} - \xi)] \\ = \frac{n - q - 2}{n - p - 2} E[\text{tr}(AA')^{-1}] + \xi' E[(A - M)A'(AA')^{-2}A(A - M)'] \xi \\ + \xi' E[(\text{tr}(AA')^{-1})(A - M)(I_p - A'(AA')^{-1}A)(A - M)'] \xi / (n - p - 2). \end{aligned}$$

PROOF OF THEOREM 3. Applying Lemma 10 (i) to

$$E_{\theta}[\hat{\xi}] = E_{\theta}[I_q - (AA')^{-1}A(A - M)'] \xi,$$

we get the expression (4.1) for expectation of $\hat{\xi}$.

For the expression (4.2) for the risk of $\hat{\xi}$, we have to apply Lemma 11 to Lemma 12.

REFERENCES

- Anderson, T. W. and Takemura, A. (1982). A new proof of admissibility of tests in the multivariate analysis of variance, *J. Multivariate Anal.*, **12**, 457-468.
- Bilodeau, M. and Kariya, T. (1989). Minimax estimators in the MANOVA models, *J. Multivariate Anal.*, **28**, 260-270.
- Branco, M., Bolfarine, H., Iglesias, P., and Arellano-Valle, R. B. (2000). Bayesian analysis of the calibration problem under elliptical distributions, *J. Statist. Plan. Infer.*, **90**, 69-85.
- Brown, P. J. (1982). Multivariate calibration (with discussion), *J. Roy. Statist. Soc.*, **B 44**, 287-321.
- Brown, P. J. (1993). *Measurement, Regression, and Calibration*, Oxford University Press, Oxford.
- Eisenhart, C. (1939). The interpretation of certain regression methods and their use in biological and industrial research, *Ann. Math. Statist.*, **10**, 162-186.
- Fujikoshi, Y. and Nishii, R. (1986). Selection of variables in a multivariate inverse regression problem, *Hiroshima Math. J.*, **16**, 269-277.
- Haff, L. R. (1981). Further identities for the Wishart distribution with applications in regression, *Canad. J. Statist.*, **9**, 215-224.
- Haff, L. R. (1982). Identities for the inverse Wishart distribution with computational results in linear and quadratic discrimination, *Sankhyā, ser. B* **44**, 245-258.
- Khatri, C. G. (1966). A note on a MANOVA model applied to problems in growth curve, *Ann. Inst. Statist. Math.*, **18**, 75-86.
- Kiefer, J. and Schwartz, R. (1965). Admissible Bayes character of T^2 -, R^2 -, and other fully invariant tests for classical multivariate normal problems, *Ann. Math. Statist.*, **36**, 747-770.
- Konno, Y. (1991). On estimation of a matrix of normal means with unknown covariance matrix, *J. Multivariate Anal.*, **36**, 44-55.
- Krutchkoff, R. G. (1967). Classical and inverse regression methods of calibration, *Technometrics*, **9**, 425-439.
- Kubokawa, T. and Robert, C. P. (1994). New perspectives on linear calibration, *J. Multivariate Anal.*, **51**, 178-200.
- Lieftinck-Koeijers, C. A. J. (1988). Multivariate calibration: a generalization of the classical estimator, *J. Multivariate Anal.*, **25**, 31-44.
- Miwa, T. (1985). Comparison among point estimators in linear calibration in terms of mean squared error (in Japanese), *Japan J. appl. Statist.*, **14**, 83-93.
- Nishii, R. and Krishnaiah, P. R. (1988). On the moments of classical estimates of explanatory variables under a multivariate calibration model, *Sankhyā, Ser. A* **50**, 137-148.

- Oman, S. D. and Srivastava, M. S. (1996). Exact mean squared error comparisons of the inverse and classical estimators in multi-univariate linear calibration, *Scand. J. Statist.*, **23**, 473–488.
- Osborne, C. (1991). Statistical calibration: a review, *Internat. Statist. Rev.*, **59**, 309–336.
- Srivastava, M. S. (1995). Comparison of the inverse and classical estimators in multi-univariate linear calibration, *Commun. Statist.-Theory Meth.*, **24**, 2753–2767.
- Sundberg, R. (1999). Multivariate calibration—direct and indirect regression methodology (with discussion), *Scand. J. Statist.*, **26**, 161–207.
- Takeuchi, H. (1997). A generalized inverse regression estimator in multi-univariate linear calibration, *Commun. Statist.-Theory Meth.*, **26**, 2645–2669.