JOURNAL OF MULTIVARIATE ANALYSIS 52, 308–324 (1995)

Estimation of a Normal Covariance Matrix with Incomplete Data under Stein's Loss*

YOSHIHIKO KONNO

Chiba University, Chiba 263, Japan

Suppose that we have $(n-a)$ independent observations from $N_p(0, \Sigma)$ and that, in addition, we have a independent observations available on the last $(p-c)$ coordinates. Assuming that both observations are independent, we consider the problem of estimating Σ under the Stein's loss function, and show that some estimators invariant under the permutation of the last $(p-c)$ coordinates as well as under those of the first c coordinates are better than the minimax estimators of Eaton. The estimators considered outperform the maximum likelihood estimator (MLE) under the Stein's loss function as well. The method involved here is computation of an unbiased estimate of the risk of an invariant estimator considered in this article. In addition we discuss its application to the problem of estimating a covariance matrix in a GMANOVA model since the estimation problem of the covariance matrix with extra data can be regarded as its canonical form.

1. INTRODUCTION

Suppose that we observe $S: p \times p$ following a Wishart distribution $W_p(\Sigma, n-a)$ with $n-a \geqslant p+1$. Partition Σ into $\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, where Σ_{11} is $c \times c$. It is also assumed that the extra data $X_1, X_2, \ldots, X_a$ are available such that each $X_i: (p-c) \times 1$ follows a $N_{p-c}(0, \Sigma_{22})$ distribution independently. The problem considered here is estimating Σ based on a sufficient statistic (S, W), where $W = \sum_{i=1}^{a} X_i X_i'$.

In this paper we employ the Stein's loss function,

$$\mathbf{L}(\hat{\Sigma}, \Sigma) = \operatorname{tr}(\hat{\Sigma}\Sigma^{-1}) - \log\det(\hat{\Sigma}\Sigma^{-1}) - p, \tag{1}$$

and evaluate the performance of an estimator by considering its risk function $\mathbf{R}(\hat{\Sigma}, \Sigma) = E[\mathbf{L}(\hat{\Sigma}, \Sigma)]$, where the expectation is taken with respect to the joint distribution of (S, W).

Received September 15, 1993; revised May 2, 1994.

AMS 1980 subject classification: 62H12, 62F20.

Key words and phrases: covariance matrix, incomplete data, Stein's loss, minimax estimator, GMANOVA model.

* Research supported by a Ministry of Education grant.

0047-259X/95 $6.00

Anderson (1957) derived the maximum likelihood estimator (see also Anderson and Olkin, 1985). Later Eaton (1970) obtained the best lower triangular invariant estimator under the loss function (1). The method involved there utilizes a representation of Bayes invariant procedures in terms of Haar measure from Zidek (1969). It is pointed out in Eaton (1970) that this estimator is minimax because of the result in Kiefer (1957), and that the maximum likelihood estimator is inadmissible under the loss (1) since it is also lower triangular invariant and different from the best one (see also Eaton, 1989, and Krishnamoorthy, 1991, for this argument). More recently, Krishnamoorthy (1991) considered a class of estimators which are invariant under the permutation of the first c coordinates and showed that some estimators are better than Eaton's minimax estimator under the loss function (1). Furthermore, Krishnamoorthy (1991) pointed out that an estimator should be invariant under the permutation of the last $(p-c)$ coordinates as well as under that of the first c coordinates and proposed an estimator having such a property. However, he did not show its minimaxness. In this paper we prove that his estimator is better than the minimax estimator of Eaton (1970) under the loss (1), which shows that his estimator is also minimax. The method involved here is a widely applied approach in eigenvalue estimation problems such as Stein (1973, 1977), Haff (1991), Dey and Srinivasan (1985), Loh (1988, 1991a, b), Muirhead and Verathaworn (1985), Bilodeau (1992), and Konno (1991) that is, computation of an unbiased estimate of the risk of certain invariant estimators from which the promising estimators are derived.

For notation, let $\Sigma_{11.2}=\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ and define $S_{11.2}$ and the partition of S like Σ. In Section 2, we reduce the problem considered here to three: first, the estimation problem of $\Sigma_{11.2}$ based on $S_{11.2}$ [distributed as $W_c(\Sigma_{11.2}, n-a-p+c)$] under the loss function $\operatorname{tr}(\hat{\Sigma}_{11.2}\Sigma_{11.2}^{-1})-\log\det(\hat{\Sigma}_{11.2}\Sigma_{11.2}^{-1})-c$; second, that of Σ_{22} based on (S_{22}, W) [$S_{22}+W$ being distributed as $W_{p-c}(\Sigma_{22}, n)$] under the quasi-loss $\operatorname{tr}(\hat{\Sigma}_{22}\Sigma_{22}^{-1})+c\operatorname{tr}(\hat{\Sigma}_{22}S_{22}^{-1})-\log\det(\hat{\Sigma}_{22}\Sigma_{22}^{-1})-(p-c)$; and third, that of $\Sigma_{12}\Sigma_{22}^{-1}$ based on $(S_{12}S_{22}^{-1}, S_{11.2})$ given S_{22} [the conditional distribution of $S_{12}S_{22}^{-1}$ given S_{22} being distributed as $N_{c\times(p-c)}(\Sigma_{12}\Sigma_{22}^{-1}, \Sigma_{11.2}\otimes S_{22}^{-1})$] under the loss matrix $(\hat{\xi}-\Sigma_{12}\Sigma_{22}^{-1})'\Sigma_{11.2}^{-1}(\hat{\xi}-\Sigma_{12}\Sigma_{22}^{-1})$, where $\hat{\xi}: c\times(p-c)$ is an estimator of $\Sigma_{12}\Sigma_{22}^{-1}$. Also, we obtain a representation of the minimax constant risk of Eaton's estimator which plays an important role in deriving an improved estimator.

Considering the first problem, Krishnamoorthy (1991) obtained improved estimators for Σ which are invariant under the permutation of the first c coordinates. By making use of the result in Bilodeau and Kariya (1989), Kubokawa *et al.* (1992) have taken up the third problem, in essence, in the context of estimating the regression coefficient in a GMANOVA model. As pointed out implicitly in Krishnamoorthy (1991), the main difficulty in deriving an invariant estimator under the permutation of the last $(p-c)$ coordinates is the second one. In Section 3, we tackle the second problem and obtain an unbiased estimate for $E[\operatorname{tr}(\hat{\Sigma}_{22}\Sigma_{22}^{-1}) + c\operatorname{tr}(\hat{\Sigma}_{22}S_{22}^{-1}) - \log\det(\hat{\Sigma}_{22})]$, where $\hat{\Sigma}_{22}$ is an orthogonally invariant estimator depending on (S_{22}, W). We apply the Wishart identity and calculus on eigenstructure in terms of $S_{22}+W$, fixing a certain random variable which is independent of $S_{22}+W$, and derive a partially differential inequality which may be regarded as an extension of the result concerned with a normal covariance matrix estimation in Stein (1977). In Section 4, obtaining solutions to this inequality, we show the minimaxness of Krishnamoorthy's estimator. In Section 5, we apply the result obtained in the previous sections to the problem of estimating a covariance matrix in GMANOVA, since it can be reduced to the estimation problem of the covariance matrix with incomplete data.

2. Reduction of the Problem and the Minimax Risk

Let $G_T(p)$ be a group of $p\times p$ lower triangular matrices, $G_T^+(p)$ its subgroup with positive diagonal elements, and $\mathcal{O}_p$ a group of $p\times p$ orthogonal matrices. Furthermore, let G be a group consisting of $p\times p$ nonsingular matrices of the form

$$g=\begin{pmatrix} g_{11} & g_{12} \\ 0 & g_{22}\end{pmatrix},$$

where g_{11} is $c\times c$, and let G_0 be a subgroup of G, where $g_{11}\in\mathcal{O}_c$, $g_{22}\in\mathcal{O}_{p-c}$, and $g_{12}=0$. Group actions on sample and parameter spaces are defined as

$$(S, W)\to(gSg', g_{22}Wg_{22}') \qquad \text{and} \qquad \Sigma\to g\Sigma g', \quad \forall g\in G.$$

Write Σ as

$$\Sigma=\begin{pmatrix} I_c & \Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I_{p-c}\end{pmatrix}\begin{pmatrix} \Sigma_{11.2} & 0 \\ 0 & \Sigma_{22}\end{pmatrix}\begin{pmatrix} I_c & 0 \\ \Sigma_{22}^{-1}\Sigma_{21} & I_{p-c}\end{pmatrix}, \tag{2}$$

which are called the partial Iwasawa coordinates, being an analogue of rectangular coordinates in the Poincaré upper half plane (see Terras, 1988 for this). It is also worth noting that the group actions defined above induce the actions given as

$$\begin{aligned}&(S_{11.2}, S_{22}, S_{12}S_{22}^{-1}, W)\\ &\quad\to(g_{11}S_{11.2}g_{11}', g_{22}S_{22}g_{22}', g_{11}S_{12}S_{22}^{-1}g_{22}^{-1}+g_{12}g_{22}^{-1}, g_{22}Wg_{22}').\end{aligned}$$

In the light of (2), we introduce estimators of the form

$$\tau(\hat{\Sigma}_{11.2}, \hat{\Sigma}_{22}, \hat{\xi}) = \begin{pmatrix} I_c & \hat{\xi} \\ 0 & I_{p-c} \end{pmatrix} \begin{pmatrix} \hat{\Sigma}_{11.2} & 0 \\ 0 & \hat{\Sigma}_{22} \end{pmatrix} \begin{pmatrix} I_c & 0 \\ \hat{\xi}' & I_{p-c} \end{pmatrix}, \tag{3}$$

where $\hat{\Sigma}_{11.2}: c \times c$, $\hat{\Sigma}_{22}: (p-c) \times (p-c)$, and $\hat{\xi}: c \times (p-c)$ depend only on $S_{11.2}$, (S_{22}, W), and $(S_{12}S_{22}^{-1}, S_{11.2})$, respectively. When $\hat{\xi} = S_{12}S_{22}^{-1}$, it turns out that $\hat{\Sigma}_{11.2}$ and $\hat{\Sigma}_{22}$ are invariant respectively under the groups of transformations of $G_T(c)$ and $G_T(p-c)$ if and only if τ is invariant under $G_T(p)$, and respectively under $\mathcal{O}_c$ and $\mathcal{O}_{p-c}$ if and only if τ is invariant under G_0. Furthermore, note that the estimator (3) is the maximum likelihood estimator if $\hat{\Sigma}_{11.2} = (n-a)^{-1} S_{11.2}$, $\hat{\Sigma}_{22} = n^{-1}(S_{22} + W)$, and $\hat{\xi} = S_{12}S_{22}^{-1}$, and that it is the constant risk minimax estimator of Eaton (1970) if

$$\begin{aligned} &\hat{\Sigma}_{11.2} = T_1 D_1^{-1} T_1', \\ &\hat{\Sigma}_{22} = [(T_2 D_2^{-1} T_2')^{-1} + cS_{22}^{-1}]^{-1}, \qquad \text{and} \qquad \hat{\xi} = S_{12}S_{22}^{-1}, \end{aligned} \tag{4}$$

where $S_{11.2} = T_1 T_1'$ and $S_{22} + W = T_2 T_2'$ for $T_1 \in G_T^+(c)$ and $T_2 \in G_T^+(p-c)$, and $D_1 = \mathrm{diag}(d_{11}, ..., d_{1c})$ and $D_2 = \mathrm{diag}(d_{21}, ..., d_{2,p-c})$ with

$$d_{1i} = n - a - p + 2c + 1 - 2i, \qquad d_{2i} = n + p - c + 1 - 2i. \tag{5}$$

Noting that

$$\Sigma^{-1} = \begin{pmatrix} I_c & 0 \\ -\Sigma_{22}^{-1}\Sigma_{21} & I_{p-c} \end{pmatrix} \begin{pmatrix} \Sigma_{11.2}^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix} \begin{pmatrix} I_c & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I_{p-c} \end{pmatrix},$$

we may see that

$$\begin{aligned} \mathrm{L}(\tau, \Sigma) = &\operatorname{tr}(\hat{\Sigma}_{11.2}\Sigma_{11.2}^{-1}) - \log\det(\hat{\Sigma}_{11.2}\Sigma_{11.2}^{-1}) - c \\ &+ \operatorname{tr}(\hat{\Sigma}_{22}\Sigma_{22}^{-1}) - \log\det(\hat{\Sigma}_{22}\Sigma_{22}^{-1}) - (p-c) \\ &+ \operatorname{tr}[(\hat{\xi} - \Sigma_{12}\Sigma_{22}^{-1})\hat{\Sigma}_{22}(\hat{\xi} - \Sigma_{12}\Sigma_{22}^{-1})' \Sigma_{11.2}^{-1}]. \end{aligned}$$

Furthermore, we have

$$\begin{aligned} &E[\operatorname{tr}(S_{12}S_{22}^{-1} - \Sigma_{12}\Sigma_{22}^{-1})\hat{\Sigma}_{22}(S_{12}S_{22}^{-1} - \Sigma_{12}\Sigma_{22}^{-1})' \Sigma_{11.2}^{-1} \mid S_{22}, W] \\ &\quad = cE[\operatorname{tr}(\hat{\Sigma}_{22}S_{22}^{-1}) \mid S_{22}, W], \end{aligned}$$

since the conditional distribution of $S_{12}S_{22}^{-1}$ given S_{22} is $N_{c \times (p-c)}(\Sigma_{12}\Sigma_{22}^{-1}, \Sigma_{11.2} \otimes S_{22}^{-1})$, $\hat{\Sigma}_{22}$ depends only on (S_{22}, W), and W is independent of S. On account of these, we may see that

$$\begin{aligned}\mathbf{R}(\tau, \Sigma) = {} & E[\operatorname{tr}(\hat{\Sigma}_{11.2}\Sigma_{22}^{-1}) - \log\det(\hat{\Sigma}_{11.2}\Sigma_{22}^{-1}) - c] \\ & + E[\operatorname{tr}(\hat{\Sigma}_{22}\Sigma_{22}^{-1}) + c\operatorname{tr}(\hat{\Sigma}_{22}S_{22}^{-1}) - \log\det(\hat{\Sigma}_{22}S_{22}^{-1}) - (p-c)] \\ & + E[\operatorname{tr}\{(\hat{\xi} - \Sigma_{12}\Sigma_{22}^{-1})\hat{\Sigma}_{22}(\hat{\xi} - \Sigma_{12}\Sigma_{22}^{-1})'\Sigma_{11.2}^{-1} \\ & - (S_{12}S_{22}^{-1} - \Sigma_{12}\Sigma_{22}^{-1})\hat{\Sigma}_{22}(S_{12}S_{22}^{-1} - \Sigma_{12}\Sigma_{22}^{-1})'\hat{\Sigma}_{11.2}\}], \qquad (6)\end{aligned}$$

if τ has the form (3). Denote the first, the second, and the last two expectations in the right hand side of (6) by $\mathbf{R}_1(\hat{\Sigma}_{11.2}, \Sigma_{11.2})$, $\mathbf{R}_2(\hat{\Sigma}_{22}, \Sigma_{22})$, and Δ_3, respectively. Hence we can split the original problem into three parts as shown below:

1. Estimation of $\Sigma_{11.2}$ is based on $S_{11.2} \sim W_c(\Sigma_{11.2}, n-a-p+c)$ under the Stein's loss function $\operatorname{tr}(\hat{\Sigma}_{11.2}\Sigma_{22}^{-1}) - \log\det(\hat{\Sigma}_{11.2}\Sigma_{11.2}^{-1}) - c$. So the results obtained by Stein (1973, 1977), Haff (1991), Dey and Srinivasan (1985), Perron (1992), Sheena and Takemura (1992), and a host of others are applicable.

2. Suppose that we observe S_{22} and W, where $S_{22}+W$ and S_{22} follow $W_{p-c}(\Sigma_{22}, n)$ and $W_{p-c}(\Sigma_{22}, n-a)$ distributions, respectively, and consider the estimation of Σ_{22} based on (S_{22}, W) under the quasi-loss

$$\mathbf{L}_2(\hat{\Sigma}_{22}, \Sigma_{22}) = \operatorname{tr}(\hat{\Sigma}_{22}\Sigma_{22}^{-1}) + c\operatorname{tr}(\hat{\Sigma}_{22}S_{22}^{-1}) - \log\det(\hat{\Sigma}_{22}\Sigma_{22}^{-1}) - (p-c). \qquad (7)$$

Note that (7) is not a loss function in the strict sense since it is not equal to zero even if $\hat{\Sigma}_{22}$ estimates its parameter correctly. However, we can regard it as a loss function without inconvenience.

3. Estimation of $\xi = \Sigma_{12}\Sigma_{22}^{-1}$ is based on $S_{12}S_{22}^{-1} \mid S_{22}$ and $S_{11.2}$ being distributed independently as $N_{c\times(p-c)}(\xi, \Sigma_{11.2} \otimes S_{22}^{-1})$ and $W_c(\Sigma_{11.2}, n-a-p+c)$, respectively. We define a loss matrix $(\hat{\xi}-\xi)'\Sigma_{22}^{-1}(\hat{\xi}-\xi)$ and evaluate an estimator $\hat{\xi}$ of ξ with its risk matrix $\mathbf{R}_3(\hat{\xi}, \xi) = E[(\hat{\xi}-\xi)'\Sigma_{22}^{-1}(\hat{\xi}-\xi) \mid S_{22}, W]$. Here an estimator $\hat{\xi}_1$ is better than another estimator $\hat{\xi}_2$ if $\mathbf{R}_3(\hat{\xi}_2, \xi) - \mathbf{R}_3(\hat{\xi}_1, \xi)$ is non-negative definite for all ξ and $\Sigma_{11.2}$. Note that $E[\operatorname{tr}(\hat{\xi}-\xi)\hat{\Sigma}_{11.2}(\hat{\xi}-\xi)'\Sigma_{22}^{-1}] = E[\operatorname{tr}\hat{\Sigma}_{22}\mathbf{R}_3(\hat{\xi}, \xi)]$ since $\hat{\Sigma}_{22}$ depends only on (S_{22}, W). Hence, if $\hat{\xi}$ is a better estimator than $S_{12}S_{22}^{-1}$ with respect to this criterion, then $\Delta_3 \leq 0$.

Here we derive a representation of the risk of Eaton's minimax estimator. We denote this estimator by $\hat{\Sigma}^M = \tau(\hat{\Sigma}_{11.2}^M, \hat{\Sigma}_{22}^M, S_{12}S_{22}^{-1})$, where $\hat{\Sigma}_{11.2}^M$ and $\hat{\Sigma}_{22}^M$ are given in (4).

THEOREM 2.1. *Let χ_k^2 be a chi-squared random variable with k degrees of freedom and let B be a $(p-c)\times(p-c)$ positive definite random matrix* *where* $B = T_2' S_{22} T_2$ *and* $S_{22} + W = T_2 T_2'$ *with* $T_2 \in G_T^+(p-c)$. *Then the minimax risk is obtained as*

$$\mathbf{R}(\hat{\Sigma}^M, \Sigma) = \mathbf{R}_1(\hat{\Sigma}^M_{11.2}, I_c) + \mathbf{R}_2(\hat{\Sigma}^M_{22}, I_{p-c})$$

with

$$\mathbf{R}_1(\hat{\Sigma}^M_{11.2}, I_c) = \sum_{i=1}^{c} \log(d_{1i}) - E(\log \chi_{n-a-p+c-i+1}),$$

$$\mathbf{R}_2(\hat{\Sigma}^M_{22}, I_{p-c}) = \sum_{i=1}^{p-c} \{\log(d_{2i}) - E(\log \chi_{n-i+1})\}$$
$$+ E \log \det[I_{p-c} + c \operatorname{diag}(d_{21}^{-1}, ..., d_{2,p-c}^{-1}) B],$$

where d_{1i} *and* d_{2i} *are given by* (5). *Furthermore, let* $b_1, b_2, ..., b_{p-c}$ *be eigenvalues of* B^{-1}. *Then* $W = \prod_{i=1}^{p-c}(1-b_i)$ *has the same distribution as* $\prod_{i=1}^{p-c} Z_{(n-a-i+1,a)}$ *if* $a \geqslant p-c$ *and as* $\prod_{i=1}^{a} Z_{(n-(p-c)-i+1,p-c)}$ *if* $p-c > a$, *where* $Z_{(k_1,k_2)}$ *is a beta random variable with parameters* $k_1/2$ *and* $k_2/2$.

Proof. To find the risk of $\hat{\Sigma}^M_{11.2}$, just note that this is the best lower triangular invariant estimator of $\Sigma_{11.2}$ based on $S_{11.2}$ and hence the result follows from James and Stein (1961). It is easily seen that $\mathbf{R}_2(\hat{\Sigma}^M_{22}, \Sigma_{22}) = E_I[\operatorname{tr} \hat{\Sigma}^M_{22} + c \operatorname{tr}(\hat{\Sigma}^M_{22} S_{22}^{-1}) - \log \det \hat{\Sigma}^M_{22} - (p-c)]$, where E_I denotes the expectation taken with respect to the joint distribution of $S_{22} + W \sim W_{p-c}(I_{p-c}, n)$ and $S_{22} \sim W_{p-c}(I_{p-c}, n-a)$, since $\hat{\Sigma}^M_{22}$ is invariant under group actions of lower triangular matrices. Using the standard distribution theory, we can see that B is independent of $S_{22} + W$. By virtue of an application of Bartellet's decomposition theorem to $S_{22} + W$, we may see that $E_I[T_2' D_2^{-1} T_2 \mid B] = I_{p-c}$, from which it follows that

$$E_I[\operatorname{tr}(\hat{\Sigma}_{22} + c\hat{\Sigma}_{22}\Sigma_{22}^{-1}) | B] = E_I[\operatorname{tr}\{T_2' D_2^{-1} T_2 (I_{p-c} + cD_2^{-1}B)^{-1}$$
$$+ cD_2^{-1}B(I_{p-c} + cD_2^{-1}B)^{-1}\} | B] = p - c.$$

In addition note that

$$\log \det[(T_2 D_2^{-1} T_2')^{-1} + cS_{22}^{-1}]$$
$$= \log \det[I_{p-c} + cD_2^{-1} T_2' S_{22}^{-1} T_2] - \log \det(S_{22} + W) + \log \det(D_2).$$

Finally, using Theorem 10.5.3 and Corollary 10.5.4 in Muirhead (1982), we can see that the eigenvalues of B^{-1} have the distribution claimed in this theorem, which completes the proof.

3. Unbiased Estimator of Risk

In this section we develop an unbiased estimate of the risk of an orthogonally invariant estimator proposed in Krishnamoorthy (1991). Let $S_{11.2} = R_1 L_1 R_1'$ and $S_{22} + W = R_2 L_2 R_2'$, where $R_1 \in \mathcal{O}_c$, $R_2 \in \mathcal{O}_{p-c}$, $L_1 = \mathrm{diag}(l_{11}, ..., l_{1c})$ with $l_{11} \geqslant \cdots \geqslant l_{1c}$, and $L_2 = \mathrm{diag}(l_{21}, ..., l_{2,p-c})$ with $l_{21} \geqslant \cdots \geqslant l_{2,p-c}$.

Furthermore, let

$$\Phi_1(L_1) = \mathrm{diag}(\phi_{11}(L_1), ..., \phi_{1c}(L_1)) \qquad \text{and}$$

$$\Phi_2(L_2) = \mathrm{diag}(\phi_{21}(L_2), ..., \phi_{2,p-c}(L_2)),$$

where ϕ_{1i} and ϕ_{2i} ($i = 1, ..., c$ or $p-c$) are non-negative functions from L_1 or L_2 to $[0, \infty)$, respectively. Consider an estimator of the form (3) where $\hat{\xi} = S_{12} S_{22}^{-1}$,

$$\hat{\Sigma}_{11.2} = R_1 \Phi_1^{-1}(L_1) R_1', \tag{8}$$

and

$$\hat{\Sigma}_{22} = [R_2 \Phi_2(L_2) R_2' + c S_{22}^{-1}]^{-1}. \tag{9}$$

In (8) and (9), we assume that ϕ_{1i} and ϕ_{2i} are differentiable on the regions $\{l_{11} > \cdots > l_{1c}\}$ and $\{l_{21} > \cdots > l_{2,p-c}\}$, respectively.

For the sake of simplicity, we remove the constant terms with respect to Σ from the risk function, and define $\mathbf{R}^* = \mathbf{R}_1^* + \mathbf{R}_2^*$ with $\mathbf{R}_1^* = E[\mathrm{tr}(\hat{\Sigma}_{11.2} \Sigma_{11.2}^{-1}) - \log\det(\hat{\Sigma}_{11.2})]$ and $\mathbf{R}_2^* = E[\mathrm{tr}(\hat{\Sigma}_{22} \Sigma_{22}^{-1}) + c\,\mathrm{tr}(\hat{\Sigma}_{22} S_{22}^{-1}) - \log\det(\hat{\Sigma}_{22})]$.

Theorem 3.1. *Consider an estimator τ of the form (3), where $\hat{\xi} = S_{12} S_{22}^{-1}$, and $\hat{\Sigma}_{11.2}$ and $\hat{\Sigma}_{22}$ are given by (8) and (9). Except for the terms constant with respect to Σ, an unbiased estimate of the risk $\mathbf{R}^*(\tau, \Sigma)$ is given as $\hat{\mathbf{R}}^* = \hat{\mathbf{R}}_1^* + \hat{\mathbf{R}}_2^*$ with*

$$\hat{\mathbf{R}}_1^* = \sum_{i=1}^{c} \left\{ \frac{n-a-p-1}{l_{1i}\phi_{1i}} - \frac{2}{\phi_{1i}^2} \frac{\partial \phi_{1i}}{\partial l_{1i}} + 2 \sum_{j>i} \frac{\phi_{1i}^{-1} - \phi_{1j}^{-1}}{l_{1i} - l_{1j}} + \log(\phi_{1i}) \right\}, \tag{10}$$

$$\begin{aligned}\hat{\mathbf{R}}_2^* = \sum_{i=1}^{p-c} &\left\{ \frac{(n-p+c+1)\,\omega_{ii}}{l_{2i}} - 2\frac{\omega_{ii}^2 \phi_{2i}}{l_{2i}} - 2 \sum_{j=1}^{p-c} \omega_{ij}^2 \frac{\partial \phi_{2i}}{\partial l_{2j}} + 2 \sum_{j>i} \frac{\omega_{ii} - \omega_{jj}}{l_{2i} - l_{2j}} \right\} \\ &+ c\,\mathrm{tr}(\hat{\Sigma}_{22} S_{22}^{-1}) - \log\det(\hat{\Sigma}_{22}),\end{aligned} \tag{11}$$

where $R_2' \hat{\Sigma}_{22} R_2 = (\omega_{ij})$.

Proof. From Dey and Srinivasan (1985), it is easily seen than an unbiased estimate of $\mathbf{R}_1^*$ is equal to (10). For $\mathbf{R}_2^*$, put $\tilde{B} = L_2^{1/2} R_2' S_{22}^{-1} R_2 L_2^{1/2}$. Then $\tilde{B}$ is independent of $V = S_{22} + W$, whose distribution is $W_{p-c}(\Sigma_{22}, n)$. Using the Wishart identity (see Haff, 1979) given $\tilde{B}$, we get

$$E[\operatorname{tr}(\hat{\Sigma}_{22}\Sigma_{11.2}^{-1}) | \tilde{B}] = E[2 \operatorname{tr}(\mathscr{D}\hat{\Sigma}_{22}) + (n-p+c-1) \operatorname{tr}(\hat{\Sigma}_{22} V^{-1}) | \tilde{B}],$$

from which it follows that

$$\begin{aligned} \mathbf{R}_2^*(\hat{\Sigma}_{22}, \Sigma_{22}) = E[&2 \operatorname{tr}(\mathscr{D}\hat{\Sigma}_{22}) + (n-p+c-1) \operatorname{tr}(\hat{\Sigma}_{22} V^{-1}) \\ &+ c \operatorname{tr}(\hat{\Sigma}_{22} S_{22}^{-1}) - \log\det(\hat{\Sigma}_{22})], \end{aligned} \tag{12}$$

where $\mathscr{D} = ((1+\delta_{ij})/2)\, \partial/\partial v_{ij}$ for $V = (v_{ij})$ and a Kronecker's delta δ_{ij}. Denote the (i, j)th elements of $\hat{\Sigma}_{22}$ and $\hat{\Sigma}_{22}^{-1}$ by u_{ij} and u^{ij}, respectively. Note that

$$\frac{\partial}{\partial u^{km}} u_{ji} = -\frac{u_{jk}u_{mi} + u_{jm}u_{ki}}{1+\delta_{km}} \tag{13}$$

for a symmetric matrix $\hat{\Sigma}_{22}^{-1}$. Using (13), the chain rule, and the symmetry of $\hat{\Sigma}_{22}$, we may observe that

$$\operatorname{tr}(\mathscr{D}\hat{\Sigma}_{22}) = -\sum_{i,j,k,m} u_{jk}u_{im} \frac{1+\delta_{ij}}{2} \frac{\partial}{\partial v_{ij}} u^{km}. \tag{14}$$

See the Appendix for verification of (14). From Haff (1991), we have the derivatives

$$\frac{1+\delta_{ij}}{2} \frac{\partial}{\partial v_{ij}} l_{2n} = r_{in} r_{jn}$$

and

$$\frac{1+\delta_{ij}}{2} \frac{\partial}{\partial v_{ij}} r_{km} = \frac{1}{2} \sum_{t \neq m} \frac{r_{kt}}{l_{2m} - l_{2t}} (r_{it} r_{jm} + r_{jt} r_{im}),$$

where $V = R_2 L_2 R_2'$, $L_2 = \operatorname{diag}(l_{21}, ..., l_{2,p-c})$, and $R_2 = (r_{ij})$. Recall that $u^{km} = \sum_{n_1} r_{kn_1} \phi_{2n_1} r_{mn_1} + cs^{km}$, where $S_{22}^{-1} = (s^{km})$. So, using the derivatives mentioned above and the symmetry of v_{ij}, we may get that

$$\begin{aligned} &\sum_{i,j,k,m} u_{jk}u_{im} \frac{1+\delta_{ij}}{2} \frac{\partial}{\partial v_{ij}} \sum_{n_1} r_{kn_1}\phi_{2n_1} r_{mn_1} \\ &\quad = \sum_i \left[\sum_{j=1} \omega_{ij}^2 \frac{\partial \phi_{2i}}{\partial l_{2j}} + \sum_{j>i} \frac{(\omega_{ij}^2 + \omega_{ii}\omega_{jj})(\phi_{2i} - \phi_{2j})}{l_{2i} - l_{2j}} \right]. \end{aligned} \tag{15}$$

Also, we recall that

$$
\begin{aligned}
& c \sum_{i,j,k,m} u_{jk} u_{im} \frac{1+\delta_{ij}}{2} \frac{\partial}{\partial v_{ij}} s^{km} \\
& = -\sum_i \left[\frac{\omega_{ii}}{l_{2i}} - \frac{\omega_{ii}^2 \phi_{2i}}{l_{2i}} + \sum_{j>i} \left\{ \frac{\omega_{ii}-\omega_{jj}}{l_{2i}-l_{2j}} + \frac{(\omega_{ij}^2 + \omega_{ii}\omega_{jj})(\phi_{2i}-\phi_{2j})}{l_{2i}-l_{2j}} \right\} \right]. \qquad (16)
\end{aligned}
$$

See the Appendix for verification of (15) and (16). Consequently, substituting (15) and (16) into (14) and simplifying this, we sum up that

$$
\operatorname{tr}(\mathscr{D}\hat{\Sigma}_{22}) = \sum_{i=1}^{p-c} \left[\frac{\omega_{ii}}{l_{2i}} - \frac{\omega_{ii}^2 \phi_{2i}}{l_{2i}} - \sum_{j=1}^{p-c} \omega_{ij}^2 \frac{\partial \phi_{2i}}{\partial l_{2j}} + \sum_{j>i} \frac{\omega_{ii}-\omega_{jj}}{l_{2i}-l_{2j}} \right].
$$

After some simplification, we can arrive at the desired result.

4. Alternative Estimations

In Krishnamoorthy (1991), a Monte Carlo study simulation indicated that a simple estimator invariant under the permutation of the first c-coordinates and of the last $(p-c)$-coordinates is minimax. However, it has not been established that this estimator is minimax analytically. Here we prove it by showing that its risk is smaller than the minimax risk.

Theorem 4.1. *Consider the estimator $\hat{\Sigma}^m$ of the form (3), where $\hat{\xi} = S_{12}S_{22}^{-1}$ and $\hat{\Sigma}_{11.2}^m$ and $\hat{\Sigma}_{22}^m$ are given by (8) and (9), respectively, and assume that Φ_1 and Φ_2 are given componentwise as*

$$
\phi_{1i} = \frac{d_{1i}}{l_{1i}} \quad (i=1, ..., c) \qquad \text{and} \qquad \phi_{2i} = \frac{d_{2i}}{l_{2i}} \quad (i=1, ..., p-c), \qquad (17)
$$

where d_{1i} and d_{2i} are given by (5). Then $\hat{\Sigma}^m$ is minimax under the loss (1).

Proof. From (6) and the result in Dey and Srinivasan (1985) and Krishnamoorthy (1991), it remains to prove that $\hat{\Sigma}_2^m$ is better than $\hat{\Sigma}_{22}^M$ with respect to the quasi-loss (7). By virtue of (12), it follows that the risk of $\hat{\Sigma}_{22}^m$, except for the constant terms, is given as

$$
\begin{aligned}
\mathbf{R}_2^*(\hat{\Sigma}_{22}^m, \Sigma_{22}) = E\Bigg[\sum_i \left\{ \frac{(n-p+c+1)\,\omega_{ii}}{l_{2i}} + 2\sum_{j>i} \frac{\omega_{ii}-\omega_{jj}}{l_{2i}-l_{2j}} \right\} \\
+ c \operatorname{tr}(\hat{\Sigma}_{22}^m S_{22}^{-1}) - \log\det(\hat{\Sigma}_{22}^m) \Bigg],
\end{aligned}
$$

since $\partial\phi_{2i}/\partial l_{2j}=0$ for $i\neq j$. Note that

$$\sum_{j>i}\frac{\omega_{ii}-\omega_{jj}}{l_{2i}-l_{2j}}=(p-c-i)\frac{\omega_{ii}}{l_{2i}}+\sum_{j>i}\frac{l_{2i}^{-1}\omega_{ii}-l_{2j}^{-1}\omega_{jj}}{l_{2j}^{-1}(l_{2i}-l_{2j})}.$$

Recall the relation $\tilde{B}=L_2^{1/2}R_2'S_{22}^{-1}R_2L_2^{1/2}$ and the fact that $D_2=\operatorname{diag}(d_{21},\ldots,d_{2,p-c})$, where $d_{21}>\cdots>d_{2,p-c}$. Since the distribution of $\tilde{B}$ is orthogonally invariant, we have that, for $j>i$,

$$E[l_{2i}^{-1}\omega_{ii}-l_{2j}^{-1}\omega_{jj}]=E[(D_2+c\tilde{B})^{ii}-(D_2+c\tilde{B})^{jj}]\leqslant 0,$$

where we denote by $(D_2+c\tilde{B})^{ii}$ the (i,i)th element of the matrix $(D_2+c\tilde{B})^{-1}$. Noting that $\tilde{B}$ is independent of L_2, we may see that

$$E\left[\sum_{j>i}\frac{l_{2i}^{-1}\omega_{ii}-l_{2j}^{-1}\omega_{jj}}{l_{2j}^{-1}(l_{2i}-l_{2j})}\right]\leqslant 0,$$

from which it follows that

$$\begin{aligned}\mathbf{R}_2^*(\hat{\Sigma}_{22}^m,\Sigma_{22})&\leqslant E\left[\sum_i\frac{(n+p-c+1-2i)\,\omega_{ii}}{l_{2i}}+c\operatorname{tr}(\hat{\Sigma}_{22}^mS_{22}^{-1})-\log\det(\hat{\Sigma}_{22}^m)\right]\\&=E[\operatorname{tr}(R_2'\hat{\Sigma}_{22}^mR_2\,D_2L_2^{-1})+c\operatorname{tr}(\hat{\Sigma}_{22}^mS_{22}^{-1})-\log\det(\hat{\Sigma}_{22}^m)]\\&=E[\operatorname{tr}(\hat{\Sigma}_{22}^m\{R_2\,D_2L_2^{-1}R_2'+cS_{22}^{-1}\})-\log\det(\hat{\Sigma}_{22}^m)]\\&=E[-\log\det(\hat{\Sigma}_{22}^m)]+(p-c).\end{aligned}$$

Furthermore, note that

$$\begin{aligned}&\log\det[R_2\,D_2L_2^{-1}R_2'+cS_{22}^{-1}]-\log\det(\Sigma_{22})\\&\quad=\log\det[I_{p-c}+c\,D_2^{-1}HBH']-\log\det((S_{22}+W)\,\Sigma_{22}^{-1})+\log\det(D_2),\end{aligned}$$

where H is a $(p-c)\times(p-c)$ orthogonal matrix. Since the distribution of H is independent of B, it follows that HBH' has the same distribution as B. Hence it follows that $\mathbf{R}_2(\hat{\Sigma}_{22}^m,\Sigma_{22})\leqslant\mathbf{R}_2(\hat{\Sigma}_{22}^M,\Sigma_{22})$, which completes the proof.

Remark. 4.1. The estimator given in Theorem 4.1 can be further improved upon just be replacing ϕ_{1i} in (7) with its superior due to Dey and Srinivasan (1985) when the dimension of the first c-coordinates is greater or equal to 3.

Following the arguments in Bilodeau and Kariya (1989) and Kubokawa *et al.* (1992), we can find the shrinkage-type estimators $\hat{\xi}_S$ for

$\Sigma_{12}\Sigma_{22}^{-1}$ based on $(S_{12}S_{22}^{-1}, S_{11.2})$ to improve upon $S_{12}S_{22}^{-1}$ with respect to the risk matrix $\mathbf{R}_3$, provided $c \geqslant 3$. For instance, we may choose $\hat{\xi}_S$ as

$$\hat{\xi}_S = \left(1 - \frac{h(F, S_{11.2})}{F}\right) S_{12}S_{22}^{-1}, \tag{18}$$

where h is a scalar function of $F = \operatorname{tr} S_{21}S_{11.2}^{-1}S_{12}S_{22}^{-1}$ and $S_{11.2}$, satisfying the following conditions:

(i) $0 \leqslant h(F, S_{11.2}) \leqslant \dfrac{2(c-2)}{n-a-p+3}$,

(ii) $\dfrac{\partial h(F, S_{11.2})}{\partial F} \geqslant 0$,

(iii) $\dfrac{\partial h(F, S_{11.2})}{\partial S_{11.2}} = \left(\dfrac{\partial h(F, S_{11.2})}{\partial (S_{11.2})_{ij}}\right)$ is non-negative definite.

Combining this shrinkage-estimator $\hat{\xi}_S$ with the estimator given in Theorem 4.1, we immediately obtain the following theorem.

THEOREM 4.2. *Consider the estimator* $\hat{\Sigma}^s = \tau(\hat{\Sigma}^m_{11.2}, \hat{\Sigma}^m_{22}, \hat{\xi}_S)$ *where* $(\hat{\Sigma}^m_{11.2}, \hat{\Sigma}^m_{22})$ *are described in Theorem* 4.1 *and* $\hat{\xi}_S$ *is given by* (18). *Then* $\hat{\Sigma}^s$ *is minimax under the loss* (1).

Remark 4.2. In Krishnamoorthy, a Monte Carlo simulation study was performed to compute the risk saving involved in the orthogonally invariant minimax estimator $\hat{\Sigma}_m$ given in Theorem 4.1. He showed that the percentage risk improvements of the estimator $\hat{\Sigma}_m$ over MLE and the minimax estimator given by (4) are about 15.6% and 10.1%, respectively, when the covariance matrix was chosen to be an identity matrix for $p = 5$, $c = 2$, $n = 25$, and $a = 5$. See Krishnamoorthy (1991) for the details.

5. COVARIANCE ESTIMATION IN GMANOVA MODEL

The model considered here can be summarized as follows;

$$Y: n \times p \sim N(ABC, I_n \otimes \tilde{\Sigma}), \tag{19}$$

where $A: n \times a$ $(n > a)$ is a known design matrix of rank a, $C: c \times p$ $(c < p)$ is a known matrix of regression matrix of rank c, and $B: a \times c$ is a matrix of unknown parameters. The maximum likelihood estimator for $\tilde{\Sigma}$ is given as

$$n\hat{\tilde{\Sigma}}_{\mathrm{mle}} = \tilde{S} + \tilde{S}C_o(C_o'\tilde{S}C_o)^{-1} W(C_o'\tilde{S}C_o)^{-1} C_o'\tilde{S}, \tag{20}$$

where $\tilde{S}: p\times p = Y'\{I_n - A(A'A)^{-1}A'\}Y$, $W: (p-c)\times(p-c) = C_o'Y'A(A'A)^{-1}A'YC_o$, and $C_o: p\times(p-c)$ is a matrix satisfying $CC_o = 0$ and $C_o'C_o = I_{p-c}$. Note that the estimate (20) does not depend on the choice of C_o. von Rosen (1991) showed that the estimator (20) is negatively biased and proposed an unbiased estimator whish is a function solely of $\hat{\tilde{\Sigma}}_{\text{mle}}$. On the other hand, Kariya (1989) referred to this problem with a view to deriving the best equivariant estimator in connection with the treatment of a model admitting an ancillary statistic. However, it has not been studied exhaustively from a decision-theoretic perspective. In this section, we consider the problem of estimating $\tilde{\Sigma}$ in the model (19) under the loss function

$$\mathrm{L}(\hat{\tilde{\Sigma}}, \tilde{\Sigma}) = \mathrm{tr}(\hat{\tilde{\Sigma}}\tilde{\Sigma}^{-1}) - \log\det(\hat{\tilde{\Sigma}}\tilde{\Sigma}^{-1}) - p, \tag{21}$$

and, by means of the result in the preceding sections, we propose an estimator which is better than both the maximum likelihood estimator and an unbiased estimator of von Rosen (1991) with respect to the loss function (21). To analyze the problem states, we reduce it to a canonical form. Define Γ as

$$\Gamma: p\times p = (C'(CC')^{-1}; C_o),$$

and write S and Σ as $S = \Gamma'\tilde{S}\Gamma$ and $\Sigma = \Gamma'\tilde{\Sigma}\Gamma$. Then it is easily seen that S is distributed as $W_p(\Sigma, n-a)$ and independent of W. Furthermore, it turns out that $S_{22} + W$ is distributed as $W_{p-c}(\Sigma_{22}, n)$ where the partitions of S and Σ are defined as in Section 2. Note that $\Gamma^{-1} = (C'; C_o)'$. If we restrict our attention to an estimator of the form

$$\hat{\tilde{\Sigma}} = (C'; C_o)\hat{\Sigma}(C'; C_o)', \tag{22}$$

where $\hat{\Sigma}: p\times p$ is an estimator of Σ based on (S, W), the loss function (21) is expressed as

$$\mathrm{L}(\hat{\tilde{\Sigma}}, \tilde{\Sigma}) = \mathrm{tr}(\hat{\Sigma}\Sigma^{-1}) - \log\det(\hat{\Sigma}\Sigma^{-1}) - p.$$

Hence comparison among estimators of the form (22) under the loss function (21) corresponds to comparison among estimators $\hat{\Sigma}$ based on (S, W) under the loss function (1), where (S, W) have the joint distribution stated above. By virtue of the relation $I_p - \tilde{S}^{-1}C'(C\tilde{S}^{-1}C')^{-1}C = C_o(C_o'\tilde{S}C_o)^{-1}C_o'\tilde{S}$ (see Siotani *et al.*, 1985, p. 311), we can see that $S_{11.2} = (C\tilde{S}^{-1}C')^{-1}$. From the identity $I_p - C'(CC')^{-1}C = C_oC_o'$, we may observe that $C'S_{12}S_{22}^{-1} + C_o = \tilde{S}C_o(C_o'\tilde{S}C_o)^{-1}$, from which it follows that

$$(C'; C_o)\begin{pmatrix} I_c & S_{12}S_{22}^{-1} \\ 0 & I_{p-c} \end{pmatrix} = (C'; \tilde{S}C_o(C_o'\tilde{S}C_o)^{-1}). \tag{23}$$

On account of these and writing (20) as

$$\tilde{S}C_o(C_o'\tilde{S}C_o)^{-1}C_o'Y'YC_o(C_o'\tilde{S}C_o)^{-1}C_o'\tilde{S}+C'(C\tilde{S}^{-1}C')^{-1}C,$$

we may summarize that the estimator (20) corresponds to the estimator

$$n\hat{\Sigma}_{\text{mle}}=\begin{pmatrix} I_c & S_{12}S_{22}^{-1} \\ 0 & I_{p-c} \end{pmatrix}\begin{pmatrix} S_{11.2} & 0 \\ 0 & S_{22}+W \end{pmatrix}\begin{pmatrix} I_c & 0 \\ S_{22}^{-1}S_{21} & I_{p-c} \end{pmatrix}.$$

Similarly, the unbiased estimator of von Rosen (1991),

$$\hat{\hat{\Sigma}}_U=\hat{\hat{\Sigma}}_{\text{mle}}+e_1C'(C\hat{\hat{\Sigma}}_{\text{mle}}^{-1}C')^{-1}C, \tag{24}$$

where

$$e_1=\frac{a(n-a-2(p-c)-1)}{(n-a-p+c-1)(n-a-p+c)},$$

corresponds to the estimator

$$n\hat{\Sigma}_U=\begin{pmatrix} I_c & S_{12}S_{22}^{-1} \\ 0 & I_{p-c} \end{pmatrix}\begin{pmatrix} (1+e_1)S_{11.2} & 0 \\ 0 & (S_{22}+W) \end{pmatrix}\begin{pmatrix} I_c & 0 \\ S_{22}^{-1}S_{21} & I_{p-c} \end{pmatrix}.$$

Since $\hat{\Sigma}_{\text{mle}}$ and $\hat{\Sigma}_U$ are invariant under the group actions of a lower triangular matrix, they are improved by the minimax estimator of Eaton (1970), from which it follows that they are also improved by the estimators given in Section 4 under the loss function (21). Our proposed estimator in the model (19) is written as

$$\hat{\hat{\Sigma}}=(C';C_o)\begin{pmatrix} I_c & S_{12}S_{22}^{-1} \\ 0 & I_{p-c} \end{pmatrix}\begin{pmatrix} \hat{\Sigma}_{11.2} & 0 \\ 0 & \hat{\Sigma}_{22} \end{pmatrix}\begin{pmatrix} I_c & 0 \\ S_{22}^{-1}S_{21} & I_{p-c} \end{pmatrix}\begin{pmatrix} C \\ C_o' \end{pmatrix}, \tag{25}$$

where $\hat{\Sigma}_{11.2}$ and $\hat{\Sigma}_{22}=[R_2\Phi_2(L_2)R_2'+c(C_o'\tilde{S}C_o)^{-1}]^{-1}$ are given in Theorem 4.1 with $(C\tilde{S}^{-1}C')^{-1}=R_1L_1R_1'$ and $C_o'\tilde{S}C_o+W=C_o'Y'YC_o=R_2L_2R_2'$. On account of (23), we may conclude that (25) is rewritten as

$$\hat{\hat{\Sigma}}=\tilde{S}C_o(C_o'\tilde{S}C_o)^{-1}\hat{\Sigma}_{22}(C_o'\tilde{S}C_o)^{-1}C_o'\tilde{S}+C'\hat{\Sigma}_{11.2}C. \tag{26}$$

Our estimator is also independent of the choice of C_o, since it is invariant under the permutation of the last $(p-c)$ coordinates.

Summing up these arguments we may obtain the following theorem.

THEOREM 5.1. *The estimator* (26) *has uniformly smaller risk than the estimator* (20) *and* (24) *with respect to the loss function* (21).

We assume that both matrices A and C are of full rank. But this is not essential since the technique adopted above will apply by replacing the inverse of the matrix and the parameters (a, c) by the generalized inverse of the matrix and the ranks of A and C.

Chinchilli and Elswick (1985) considered the mixed MANOVA–GMANOVA model expressed as

$$Y \sim N(A_1 B_1 C + A_2 B_2, I_n \otimes \tilde{\Sigma}), \tag{27}$$

where Y, C, and $\tilde{\Sigma}$ are as desribed in (19). However, A_1 and A_2 are $n \times a_1$ and $n \times a_2$ design matrices, the argument matrix $[A_1; A_2]$ being of full rank $a_1 + a_2$, and B_1 and B_2 are $a_1 \times c$ and $a_2 \times p$ unknown matrices. They obtain the maximum likelihood estimator for the model (27) as

$$\tilde{\Sigma} = \tilde{S}_1 + \tilde{S}_1 C_o (C_o' \tilde{S}_1 C_o)^{-1} W_1 (C_o' \tilde{S}_1 C_o)^{-1} C_o' \tilde{S}_1,$$

where

$$\tilde{S}_1 = X' \left[I_n - (A_1; A_2) \begin{pmatrix} A_1'A_1 & A_1'A_2 \\ A_2'A_1 & A_2'A_2 \end{pmatrix}^{-1} \begin{pmatrix} A_1' \\ A_2' \end{pmatrix} \right] X$$

and $W_1 = C_o' X' A (A'A)^{-1} A' X C_o$ with $X = (I_n - A_2(A_2'A_2)^{-1} A_2') Y$ and $A = A_1(I_n - A_2(A_2'A_2)^{-1} A_2')$. Since $\Gamma' \tilde{S}_1 \Gamma$ is distributed as $W_p(\Sigma, n - a_1 - a_2)$, being independent of W_1, where W_1 follows the $W_{p-c}(\Sigma_{22}, a_1)$ distribution, an improved estimator for the model (27) is easily obtained from Theorem 5.1 by replacing S and W with S_1 and W_1.

Appendix

The notation and terminology in Section 3 are as follows.

The calculation in line (14). We have

$$\begin{aligned}
\operatorname{tr}(\mathscr{D}\hat{\Sigma}_{22}) &= \sum_{i \geq j} \frac{\partial}{\partial v_{ij}} u_{ji} \\
&= \sum_{i \geq j} \sum_{k \geq m} \frac{\partial}{\partial u^{km}} u_{ji} \frac{\partial}{\partial v_{ij}} u^{km} \\
&= \sum_{i,j} \sum_{k,m} \frac{1+\delta_{km}}{2} \frac{\partial}{\partial u^{km}} u_{ji} \frac{1+\delta_{ij}}{2} \frac{\partial}{\partial v_{ij}} u^{km} \\
&= - \sum_{i,j,k,m} \frac{u_{jk} u_{mi} + u_{jm} u_{ki}}{2} \frac{1+\delta_{ij}}{2} \frac{\partial}{\partial v_{ij}} u^{km} \\
&= - \sum_{i,j,k,m} u_{jk} u_{im} \frac{1+\delta_{ij}}{2} \frac{\partial}{\partial v_{ij}} u^{km}.
\end{aligned}$$

The calculation in line (15). We have

$$
\begin{aligned}
&\sum_{i,j,k,m} u_{jk}u_{im}\frac{1+\delta_{ij}}{2}\frac{\partial}{\partial v_{ij}}\sum_{n_1} r_{kn_1}\phi_{2n_1}r_{mn_1}\\
&= \sum_{i,j,k,m} u_{jk}u_{im}\sum_{n_1}\Bigg[2\phi_{2n_1}r_{mn_1}\frac{1+\delta_{ij}}{2}\frac{\partial}{\partial v_{ij}}r_{kn_1}\\
&\quad+\sum_{n_2} r_{kn_1}r_{mn_1}\frac{\partial\phi_{2n_1}}{\partial l_{2n_2}}\frac{1+\delta_{ij}}{2}\frac{\partial}{\partial v_{ij}}l_{2n_2}\Bigg]\\
&= \sum_{i,j,k,m} u_{jk}u_{im}\sum_{n_1}\Bigg[\phi_{2n_1}r_{mn_1}\sum_{t\neq n_1}\frac{r_{kt}}{l_{2n_1}-l_{2t}}(r_{it}r_{jn_1}+r_{jt}r_{in_1})\\
&\quad+\sum_{n_2} r_{kn_1}r_{mn_1}r_{in_2}r_{jn_2}\frac{\partial\phi_{2n_1}}{\partial l_{2n_2}}\Bigg]\\
&=\sum_i\Bigg[\sum_{j=1}\omega_{ij}^2\frac{\partial\phi_{2i}}{\partial l_{2j}}+\sum_{j>i}\frac{(\omega_{ij}^2+\omega_{ii}\omega_{jj})(\phi_{2i}-\phi_{2j})}{l_{2i}-l_{2j}}\Bigg].
\end{aligned}
$$

The calculation in line (16). Write the (k, m)th element of $S_{22}{}^{1}$ as

$$
s^{km}=\sum_{n_1,n_2} r_{kn_1}l_{2n_1}{}^{1/2}b_{n_1n_2}l_{2n_2}^{-1/2}r_{mn_2}
$$

for $\tilde{B}=(b_{ij})$. Similarly, we get that

$$
\begin{aligned}
&c\sum_{i,j,k,m} u_{jk}u_{im}\frac{1+\delta_{ij}}{2}\frac{\partial}{\partial v_{ij}}s^{km}\\
&=2c\sum_{i,j,k,m,n_1,n_2} u_{jk}u_{im}b_{n_1n_2}\Bigg[r_{kn_1}r_{mn_2}l_{2n_2}^{-1/2}\frac{1+\delta_{ij}}{2}\frac{\partial}{\partial v_{ij}}l_{2n_1}{}^{1/2}\\
&\quad+(l_{2n_1}l_{2n_2})^{-1/2}r_{mn_2}\frac{1+\delta_{ij}}{2}\frac{\partial}{\partial v_{ij}}r_{kn_1}\Bigg]\\
&=-c\sum_{i,j,k,m,n_1,n_2}\Bigg[l_{2n_1}{}^{-1}r_{jn_1}u_{jk}r_{kn_1}l_{2n_1}{}^{1/2}b_{n_1n_2}l_{2n_2}{}^{-1/2}r_{mn_2}u_{im}r_{in_1}\\
&\quad-\sum_{t\neq n_1}\Bigg\{\frac{r_{kt}u_{jk}r_{jn_1}r_{it}u_{im}r_{mn_2}l_{2n_2}{}^{1/2}b_{n_1n_2}l_{2n_1}{}^{1/2}}{l_{2n_1}-l_{2t}}\\
&\quad-\frac{r_{kt}u_{jk}r_{jt}r_{in_1}u_{im}r_{mn_2}l_{2n_2}{}^{1/2}b_{n_1n_2}l_{2n_1}{}^{1/2}}{l_{2n_1}-l_{2t}}\Bigg\}\Bigg].
\end{aligned}
$$

Then, using the relations $\tilde{B}=L_2^{1/2}R_2'S_{22}{}^{1}R_2L_2^{1/2}$ and $cS_{22}{}^{1}=\hat{\Sigma}_{22}{}^{1}-R_2\Phi_2R_2'$, we can obtain the desired result.

ACKNOWLEDGMENTS

The author sincerely thanks the referees for their very careful reading and thoughtful comments which led to a much improved version of the paper.

REFERENCES

1. ANDERSON, T. W. (1957). Maximum likelihood estimators for a multivariate normal distribution when some observations are missing. *J. Amer. Statist. Assoc.* **52** 200–203.
2. ANDERSON, T. W., AND OLKIN, I. (1985). Maximum-likelihood estimation of the parameters of multivariate normal distribution. *Linear Algebra Appl.* **70** 147–171.
3. BILODEAU, M. (1992). Estimation of the eigenvalues of $\Sigma_1\Sigma_2^{-1}$. *J. Multivariate Anal.* **41** 1–13.
4. BILODEAU, M., AND KARIYA, T. (1989). Minimax estimators in the normal MANOVA model. *J. Multivariate Anal.* **28** 260–270.
5. CHINCHILLI, V. M., AND ELSWICK, K. (1985). A mixture of the MANOVA and GMANOVA models. *Comm. Statist. Theory Methods* **14** 3075–3089.
6. DEY, D. K., AND SRINIVASAN, C. (1985). Estimation of a covariance matrix under Stein's loss. *Ann. Statist.* **13** 1581–1591.
7. EATON, M. L. (1970). *Some Problems in Covariance Estimation*, Tech. Rep. 49. Department of Statistics, Stanford University.
8. EATON, M. L. (1989). Group invariance application in statistics. In *NSF–CBMS Regional Conference Series in Probability and Statistics.*
9. HAFF, L. R. (1979). An identity for the Wishart distribution with applications. *J. Multivariate Anal.* **9** 531–544.
10. HAFF, L. R. (1991). The variational form of certain Bayes estimators. *Ann. Statist.* **19** 1163–1190.
11. JAMES, W., AND STEIN, C. (1961). Estimation with quadratics loss. In *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. I, pp. 361–379. Univ. California Press, Berkeley.
12. KARIYA, T. (1989). Equivariant estimation in a model with an ancillary statistic. *Ann. Statist.* **17** 920–928.
13. KIEFER, J. (1957). Invariance, minimax sequential estimation, and continuous time processes. *Ann. Math. Statist.* **28** 573–601.
14. KONNO, Y. (1991). On estimation of a matrix of normal means with unknown covariance matrix. *J. Multivariate Anal.* **36** 44–55.
15. KRISHNAMOORTHY, K. (1991). Estimation of normal covariance and precision matrices with incomplete data. *Comm. Statist. Theory Methods* **20** 757–770.
16. KUBOKAWA, T., SALEH, A.K.MD.E., AND MORITA, K. (1992). Improving on MLE of coefficient matrix in a growth curve model. *J. Statist. Planning Inference* **31** 169–177.
17. LOH, W.-L. (1988). *Estimating Covariance Matrices*, Ph.D. thesis. Stanford University.
18. LOH, W.-L. (1991a). Estimating covariance matrices. *Ann. Statist.* **19** 283–296.
19. LOH, W.-L. (1991b). Estimating covariance matrices, II. *J. Multivariate Anal.* **36** 163–174.
20. MUIRHEAD, R. J. (1982). *Aspects of Multivariate Statistical Theory.* Wiley, New York.
21. MUIRHEAD, R. J., AND VERATHAWORN, T. (1985). On estimating the latent roots of $\Sigma_1\Sigma_2^{-1}$. In *Multivariate Analysis–VI* (P. R. Krishnaiah, Ed.), pp. 431–447. North-Holland, Amsterdam.
22. PERRON, F. (1992). Minimax estimators of a covariance matrix. *J. Multivariate Anal.* **43** 6–28.

23. Sheena, Y., and Takemura, A. (1992). Inadmissibility of non-order-preserving orthogonally invariant estimators of the covariance matrix in the case of Stein's loss. *J. Multivariate Anal.* **41** 117–131.
24. Siotani, M., Hayakawa, T., and Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis*, American Sciences Press, Columbus, OH.
25. Stein, C. (1973). Estimation of the mean of a multivariate normal distribution. In *Proceedings, Prague Symposium on Asymptotic Statistics*, pp. 345–381.
26. Stein, C. (1977). Lectures on multivariate estimation theory. In *Investigation on Statistical Estimation Theory I*, pp. 4–65. In *Zapiski Nauchych Seminarov LOMI*, Vol. 74. V. A. Steklova AN SSSR, Leningrad.
27. Terras, A. (1988). *Harmonic Analysis on Symmetric Spaces and Applications*, II. Springer-Verlag, New York/Berlin/Heidelberg.
28. von Rosen, D. (1991). Moments of maximum likelihood estimators in growth curve model. *Statistics* **22** 111–131.
29. Zidek, J. V. (1969). A representation of Bayes invariant procedures in terms of Haar measure. *Ann. Inst. Statist. Math.* **21** 291–308.