IMPROVED ESTIMATION OF MATRIX OF NORMAL MEAN AND EIGENVALUES IN THE MULTIVARIATE F-DISTRIBUTION

A DISSERTATION
SUBMITTED TO THE INSTITUTE OF MATHEMATICS
OF UNIVERSITY OF TSUKUBA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF SCIENCE

By
KONNO Yoshihiko
(Ishinomaki Senshu University)

February, 1992

ABSTRACT

This thesis focuss on the following estimation problems from the decision-theoretic point of view:

Matrix of normal mean

Let X be an $m \times p$ matrix normally distributed with matrix of mean B and a covariance matrix $I_m \otimes \Sigma$, where Σ is a $p \times p$ unknown positive definite matrix. We wish to estimate B under the loss function

$$\operatorname{tr} \Sigma^{-1}(\hat{B} - B)'(\hat{B} - B).$$

Here we denote the transpose of the matrix A by A'. Formulae are obtained for an unbiased estimate of the risk (the expected loss function) of certain forms of estimate of the mean matrix B. From these, improved estimators which beat the commonly used estimator X are proposed.

Eigenvalues in the multivariate F-distribution

Let $U_{p \times p}$ have the multivariate F-distribution with a positive definite $p \times p$ matrix Δ of scale parameters and degrees of freedom k_1 and k_2. We wish to estimate the eigenvalues of the scale matrix Δ under the loss functions

$$\begin{aligned} \mathbf{L}_1(\Delta, \hat{\Delta}) &= \operatorname{tr}(\Delta^{-1}\hat{\Delta}) - \log\det(\Delta^{-1}\hat{\Delta}) - p, \\ \mathbf{L}_2(\Delta, \hat{\Delta}) &= \operatorname{tr}(\Delta^{-1}\hat{\Delta} - I_p)^2. \end{aligned}$$

By recursive use of the F-identity (integration by parts formula for this distribution), improved estimators which beat the best scalar multiple of U are proposed with respect to

the loss function $\mathbf{L}_1$ or $\mathbf{L}_2$. For the case where the scale parameter Δ is 2×2, orthogonally invariant minimax estimators are given with respect to the loss function $\mathbf{L}_1$.

ACKNOWLEDGMENTS

I am indebted to my adviser, Professor Nariaki Sugiura for his support, encouragement, and guidance which made this thesis possible.

I wish to take this opportunity to express my deep gratitude to Professor Masafumi Akahira, Drs. Taka-aki Shiraishi and Tatsuya Kubokawa for encouragement and suggestions.

I would like to thank Professors M. Bilodeau, L. Haff, F.Perron, B.K. Sinha and Mr. M. Fujimoto for sending me their unpublished papers or giving me valuable comments.

Finally, special thanks go to E. Costello, the Dire Straites, D. Edmunds, J. Hiatt, N.Lowe, T. Petty, the Rolling Stones, and the World Party for musical support.

CONTENTS

CHAPTER 1

INTRODUCTION

Since the discovery of so-called Stein effect, the multiparameter estimation which occurs rather naturally in statistical decision-theoretic problems has been studied in many literatures. More recently the estimation problem in which procedures revolve mainly around the eigenstructures of random and parameter matrices has been received a lot of attention. There is a growing literature relating to the eigenvalues estimation: A useful review paper by Muirhead[41] gives many references. This thesis is concerned with possible decision theoretic approach to the following problems:

1. Let X be an $m \times p$ matrix normally distributed with matrix of mean B and a covariance matrix $I_m \otimes \Sigma$ and let S be a $p \times p$ Wishart matrix with n degrees of freedom and mean $n\Sigma$. We wish to estimate the mean matrix B.

2. Let U be a $p \times p$ random matrix, having the multivariate F-distribution with a scale matrix Δ and degrees of freedom k_1 and k_2. We wish to estimate the eigenvalues of Δ.

The usual or classical estimator for these problems, i.e., the maximum likelihood estimator or the unbiased estimator, is fully equivariant. It means that, under any groups of transformations, it transforms in the same way as its estimand. Equivariance is desirable but not compelling property as shown in Stein[48] that, in estimating a p-variate normal mean vector with identity covariance matrix, the sample mean is inadmissible in terms of sum of the squared-error loss if $p \geq 3$. Moreover, such a classical estimate ignores information about the ordered eigenvalues of the random and parameter matrices in the multivariate situation. All these arguments suggest that a systematic treatment of above problems can probably give superior alternatives to the commonly used estimator.

1.1. DEFINITION

The notation $A : p \times q$ is a matrix of p rows and q columns. The transpose of A is denoted by A'. When $p = q$, then $\det A$ or $|A|$ is the determinant of the square matrix A. The trace of A is denoted by $\operatorname{tr} A$, that is, the sum of diagonal elements of square matrix A. $D = \operatorname{diag}(d_1, d_2, \ldots, d_p)$ is a diagonal matrix with diagonal elements $d_1, d_2, \cdots, d_P$. The expected value of a random variable X is denoted by $\mathbf{E}\, X$.

Next let us recall some of terminology and definitions involved in decision-theoretic estimation.

Let X denote a random variable whose distribution depends on an unknown parameter θ. Here X can be a vector or matrix, as can θ. Let $\hat{\theta}(X)$ denote an estimate of θ. We shall denote its estimate of parameter by "hat" over parameter. A ***loss function*** $\mathbf{L}(\theta, \hat{\theta}(X))$ is a non-negative function of θ and $\hat{\theta}(X)$. The ***risk function*** $\mathbf{R}(\theta, \hat{\theta}(X))$ is the expectation of this loss function with respect to the distribution of X when θ represents the true value of the parameter. An estimate $\theta_1(X)$ is said to be ***better than*** or ***beat*** estimate $\hat{\theta}_2(X)$ if

$$\mathbf{R}(\theta, \hat{\theta}_1(X)) \leq \mathbf{R}(\theta, \hat{\theta}_2(X)) \quad \text{for} \quad \forall \theta \in \mathbf{\Omega},$$

where $\mathbf{\Omega}$ is the parameter space of the distribution of X, and

$$\mathbf{R}(\theta, \hat{\theta}_1(X)) < \mathbf{R}(\theta, \hat{\theta}_2(X)) \quad \text{for at least one} \quad \theta.$$

An estimate is said to be ***admissible*** if there exists no estimate which beats it. If there is an estimate which beats it , it is called ***inadmissible***. An estimate $\theta_0(X)$ is said to be ***minimax*** if there is no other estimate whose risk function has smaller supremum, i.e.,

$$\sup_{\theta \in \mathbf{\Omega}} \mathbf{R}(\theta, \hat{\theta}_0(X)) = \inf_{\hat{\theta}(X) \in \mathbf{C}} \sup_{\theta \in \mathbf{\Omega}} \mathbf{R}(\theta, \hat{\theta}(X)),$$

where $\mathbf{C}$ denotes the class of estimators of θ.

To introduce invariance into decision problem, let G be a topological group which acts on $\mathbf{X}$, $\mathbf{\Omega}$, and $\mathbf{A}$ where $\mathbf{X}$ is a sample space and $\mathbf{A}$ is an action space. Let $\mathcal{P} = \{P_\theta \,|\, \theta \in \mathbf{\Omega}\}$ be a family of distribution of X. The family $\mathcal{P}$ is said to be ***invariant*** under G if

$$gP_\theta = P_{g\theta} \quad \text{for} \quad g \in G,\, \theta \in \mathbf{\Omega}.$$

The loss function is said to be ***invariant*** under G if

$$\mathbf{L}(g\theta, g\hat{\theta}) = \mathbf{L}(\theta, \hat{\theta}) \quad \text{for} \quad g \in G,\, \theta \in \mathbf{\Omega},\, \hat{\theta} \in \mathbf{A}.$$

An estimation problem is said to be ***invariant*** under G if the family of distributions and the loss function are invariant.

An estimator $\hat{\theta}_0(X)$ is said to be ***equivariant*** under G if

$$\hat{\theta}_0(gX) = g\hat{\theta}_0(X) \quad \text{for} \quad g \in G, X \in \mathbf{X}.$$

Eaton[13] provides a good introduction to invariant decision problem and some techniques for finding "good" invariant decision rules.

1.2. THE GENERAL METHOD OF IMPROVING UPON USUAL ESTIMATOR

As seen in Stein[48], a fully equivariant estimator such as the maximum likelihood estimator or the unbiased estimator can be improved by simply requiring less invariance. However, once we look outside the class of fully equivariant estimators, the expression of its risk function becomes intractable in the multivariate situation. To overcome this, we shall employ the following method, which was first introduced by Stein[48]. The following description of this method comes from Loh[36].

1. Narrow the class of the estimators, for example, using equivariance in the problem and work out the form of the estimators.

2. Compute the unbiased estimate of the risk of the estimators under consideration using the integration by parts formula.

3. Determine promising alternative estimators from the unbiased estimate of risk.

By reducing the size of contenders, the problem reduces the difficulty of the search for a superior alternative to the commonly used estimator. Furthermore, equivariance is imposed on pooling across rows and columns, which implies that the resulting unbiased risk estimate depends only on the maximum invariant statistics if the estimation problem considered is invariant. The final step is rather difficult to deal with, since it has not been established to obtain the widely applicable way of deriving promising estimators (which has frequentist risk uniformly smaller than the classical estimator) from unbiased risk estimate. Since the introduction of this approach, numerous researchers have applied this technique to the problems in statistical decision theory. The literature includes Berger[2], Dey and Srinivasan[11], Efron and Morris[15], Haff[19, 22, 23], Loh[35, 36, 37], Muirhead and Verathworn[42], and Perron[46]. See Berger[3] for an extensive reference.

1.3. SUMMARY OF RESULTS

In Chapter 2, the problem of estimating matrix of normal mean is considered relative to the invariant loss function where the covariance matrix is unknown. Certain form of invariant estimators is introduced. Then the unbiased estimate of the risk is obtained, which depends on the eigenvalues of the usual F-matrix $X'XS^{-1}$. It facilitates an extensive search for superior alternatives to the commonly used estimator X. New classes of invariant minimax estimators are proposed, which are multivariate extensions of the estimators of James and Stein[24], Baranchik[1], and Lin and Tsai[35]. These results also extend the estimators of mean matrix of Stein[48] and Zheng[54] for the case in which the covariance matrix is known. Furthermore, following an approach by Haff[23], an alternative estimator as the solution of the Euler-Lagrange system of partial differential equations is derived.

Chapter 3 is devoted to the estimation of the eigenvalues in the multivariate F-distribution. First, using the F-identity we derive the second order moments of this distribution, which is useful to the statistical inference. For the Stein's loss function, we obtain the improved estimators (Haff-type and Perron-type) which beat the unbiased estimator. Furthermore, we derive the orthogonally invariant minimax estimators when the scale matrix is 2×2. For the squared loss function, we give the Haff-type improved estimators which beat the best scalar multiple estimator.

CHAPTER 2

ESTIMATING MATRIX OF NORMAL MEAN

Assume that

$$\begin{aligned} &X:\ m \times p \sim N(B, I_m \otimes \Sigma), \\ &S:\ p \times p \sim W_p(\Sigma, n), \\ &X \text{ and } S \text{ are independent,} \\ &B \text{ and } \Sigma \text{ are unknown.} \end{aligned} \tag{2.1}$$

Based on (X, S) we consider the problem of estimating B with respect to the loss function

$$\mathbf{L}((B, \Sigma), \hat{B}) = \operatorname{tr} \Sigma^{-1}(\hat{B} - B)'(\hat{B} - B). \tag{2.2}$$

The risk function corresponding to this loss function is

$$\mathbf{R}((B, \Sigma), \hat{B}) = \mathbf{E}_{\boldsymbol{B},\boldsymbol{\Sigma}}[L((B, \Sigma), \hat{B})].$$

Several authors have considered the minimaxity under this risk function. Baranchik[1] obtained a class of minimax estimators when $m \geq 3$ and $p = 1$ and Straderman[50] extended Baranchik's class of minimax estimators, while Lin and Tsai[35] treated the case where $m = 1$ and $p \geq 3$, and obtained a class of minimax estimators similar to that of Baranchik. They assume that variance or covariance matrix is unknown, but it is noted that their method to prove minimaxity is a direct evaluation of the risk function. Our interest is how these theoretical results concerning the parameter of univariate normal law may be extended to the multivariate one.

The case $\Sigma = I_p$ is considered by Stein[48] where our basic approach described in Section 1.2 was introduced. Using the result of Stein, Zheng[54] extended the results of Baranchik to the multivariate one. Later, the case where Σ is unknown and $p < m$ is considered by Zidek[55] where the underlying method, a multivariate version of that of James and Stein[24] , uses zonal polynomials expansions for the distributions of certain noncentral statistics, while Efron and Morris[15] proposed minimax estimator, called Efron-Morris type estimator, from an empirical Bayes argument. More recently, Bilodeau and Kariya[5] treated the case where $m \geq p$ and proposed several types of minimax estimators by using the unbiased risk estimate.

The objective of this chapter is to demonstrate a systematic search for superior alternatives to the commonly used estimator X by following the basic approach described in Section 1.2. In Section 2.1, we provide the basic identities, called Normal-identity and Wishart-identity, and record some calculus on eigenstructures to help in computation of the unbiased risk estimate. In Section 2.2, a certain form of equivariant estimators under the group of natural transformations is introduced so that the representation of the unbiased risk estimate could be obtained in terms of the eigenvalues of the usual F-matrix $X'XS^{-1}$. In Section 2.3, such calculation is undertaken. Following an approach by Haff[23] , an alternative estimator as the solution of the Euler-Lagrange system of partial differential equations is obtained. Furthermore, the new classes of minimax estimators are proposed. In Section 2.5, other forms of estimators are considered for the case where $m > p + 1$, which beat the commonly used estimator X.

2.1. PRELIMINARIES

In this section we shall state basic identities and some useful calculus lemmas on eigenstructures. For this end, we introduce additional notation.

Let ∇_X be $m \times p$ differential operator whose (i,j) element is given by $(\partial/\partial X_{ij})$ for $X = (X_{ij})$ and let D_S be a $p \times p$ differential operator whose (i,j) element is given by $(1/2)(1+\delta_{ij})(\partial/\partial S_{ij})$ for a Kronecker's delta δ_{ij} and $S = (S_{ij})$. We define $\nabla_X T(X) = (\sum_{k=1}^{p} \partial t_{kj}/\partial X_{ik})$ as a formal product followed by differentiation at the component level for a matrix $T(X) = (t_{ij})$, where t_{ij} is a differentiable function from $R^{m\times p}$ to R, and $\nabla_X t(X) = (\partial t(X)/\partial X_{ij})$ for a scalar function $t(X)$. The operation of D_S on a matrix or a scalar valued function is defined in the same way of that of ∇_X.

LEMMA 2.1.1 (Normal-identity). *Let $y:\ p\times 1 \sim N_p(\xi,\Sigma)$ and $f:\ R^p \to R^p$ be differentiable with $\mathrm{E}\,|\partial f_i/\partial y_j| < \infty$ $(i,j = 1,\ldots,p)$, where $y = (y_i)$ and $f = (f_i)$. Then*

$$\mathbf{E}\,[f(y)(y-\xi)'] = \mathbf{E}\,[\partial f(y)/\partial y]\Sigma,$$

where $\partial f(y)/\partial y = (\partial f_i(y)/\partial y_j)$.

This lemma is taken from Bilodeau and Kariya[5] and hence the proof is omitted. Essentially the same lemma can be seen in Loh[36].

Let $Q = Q(S)$ be a $p \times p$ matrix-valued function whose (i,j) elements q_{ij} are differentiable with $\mathbf{E}\,|q_{ij}| < \infty$ and $\mathrm{E}\,|\partial q_{ij}/\partial S_{ij}| < \infty$. The following lemma and its proof can be seen in Haff[17] and Loh[36].

LEMMA 2.1.2 (Wishart-identity). *Let S be a $p\times p$ Wishart matrix with n degrees of freedom and mean $n\Sigma$. Then*

$$\mathbf{E}\,[\,\mathrm{tr}\,Q\Sigma^{-1}] = 2\mathbf{E}\,[\,\mathrm{tr}\,D_S Q] + (n-p-1)\mathbf{E}\,[\,\mathrm{tr}\,S^{-1}Q].$$

Combining these identities gives the next lemma, which is taken from Bilodeau and Kariya[5].

LEMMA 2.1.3. *Assume that $G(X,S)$ is an $m \times p$ matrix whose elements are absolutely continuous functions of X and S such that*

$$\mathbf{E}\, G_{ij}^2 < \infty, \mathbf{E}\, |\partial G_{ij}/\partial X_{kl}| < \infty, \text{ and } \mathbf{E}\, (\partial G_{ij}/\partial S_{kl})^2 < \infty$$

and that the conditions in Theorem 2.1 (Haff[18]) *are satisfied. Then we get*

$$\begin{aligned}\mathbf{R}((B,\Sigma), X + G(X,S)) = & pm + 2\mathbf{E}\,[\,\operatorname{tr} \nabla_x' G(X,S)] + 2\mathbf{E}\,[\,\operatorname{tr} D_s G'(X,S) G(X,S)] \\ & + (n-p-1)\mathbf{E}\,[\,\operatorname{tr} G'(X,S) G(X,S) S^{-1}].\end{aligned}$$

Next we record two lemmas which state the action of the differential operator. See Haff[18, 21] for the detail of the proofs.

LEMMA 2.1.4. *For any symmetric matrix $S = (S_{kl})$, the derivatives of S^{-1} are given by*

$$\partial S^{-1}/\partial S_{kl} = -S^{-1}(e_k e_l' + e_l e_k') S^{-1}/(1 + \delta_{kl}),$$

where e_k denotes the kth unit column vector and δ_{kl} is a Kronecker's delta.

LEMMA 2.1.5. *Let Q and T be matrix functions of S. Assuming all relevant products and derivatives exist, we have*

$$D_s QT = (D_s Q)T + (Q' D_s')' T.$$

REMARK 2.1.1. The familiar law for the transposing products is not available in above lemma. The product $Q' D_s'$ is computed, then the transpose is to be taken.

LEMMA 2.1.6. *Assume that $m > p$ and that Q is a $p \times p$ matrix-valued function of $W = (W_{ij}) = X'X$ and S. Furthermore, let $D_w = (d_w^{ij})$ where $d_w^{ij} = (1/2)(1+\delta_{ij})(\partial/\partial W_{ij})$ for a Kronecker's delta δ_{ij}. Assuming that the derivatives exist, we have*

(i) $$\operatorname{tr} \nabla_x' X Q = m \operatorname{tr} Q + \operatorname{tr} X' \nabla_x Q',$$

(ii) $$\nabla_x Q = 2 X D_w Q,$$

(iii) $$\operatorname{tr} X' \nabla_x S Q = 2 \operatorname{tr} S D_w W Q' - (p+1) \operatorname{tr} S Q.$$

PROOF. (i) Let $Q = (q_{ij})$. The left hand side can be expressed as

$$\begin{aligned}\sum_{k_1,k_3=1}^{p}\sum_{k_2=1}^{m}\frac{\partial}{\partial X_{k_2k_1}}X_{k_2k_3}q_{k_3k_1} &= \sum_{k_1,k_3=1}^{p}\sum_{k_2=1}^{m}\left\{\delta_{k_1k_3}q_{k_3k_1} + X_{k_2k_3}\frac{\partial q_{k_3k_1}}{\partial X_{k_2k_1}}\right\}\\ &= m\sum_{k_1,k_3=1}^{p}\delta_{k_1k_3}q_{k_3k_1} + \sum_{k_1,k_3=1}^{p}\sum_{k_2=1}^{m}X_{k_2k_3}\frac{\partial q_{k_3k_1}}{\partial X_{k_2k_1}}\\ &= m\operatorname{tr}Q + \operatorname{tr}X'\nabla_X Q'.\end{aligned}$$

(ii) Using the symmetry of W and the chain rule, the (i, j) element of $\nabla_X Q$ can be expressed as

$$\begin{aligned}\sum_{k_1=1}^{p}\frac{\partial}{\partial X_{ik_1}}q_{k_1j} &= \sum_{k_1=1}^{p}\sum_{k_2\le k_3}^{p}\frac{\partial q_{k_1j}}{\partial W_{k_2k_3}}\frac{\partial}{\partial X_{ik_1}}W_{k_2k_3}\\ &= \sum_{k_1,k_2,k_3=1}^{p}\frac{1+\delta_{k_2k_3}}{2}\frac{\partial q_{k_1j}}{\partial W_{k_2k_3}}\frac{\partial}{\partial X_{ik_1}}W_{k_2k_3}.\end{aligned}\tag{2.3}$$

From $W = X'X$,

$$\begin{aligned}\frac{\partial}{\partial X_{ik_1}}W_{k_2k_3} &= \sum_{k_4=1}^{m}\frac{\partial}{\partial X_{ik_1}}X_{k_4k_2}X_{k_4k_3}\\ &= \sum_{k_4=1}^{m}\{\delta_{ik_4}\delta_{k_1k_2}X_{k_4k_3} + \delta_{ik_4}\delta_{k_1k_3}X_{k_4k_2}\}\\ &= \delta_{k_1k_2}X_{ik_3} + \delta_{k_1k_3}X_{ik_2}.\end{aligned}\tag{2.4}$$

Putting (2.4) into (2.3) and using symmetry of D_W give that

$$\begin{aligned}&\sum_{k_1,k_2,k_3=1}^{p}\frac{1+\delta_{k_2k_3}}{2}\frac{\partial q_{k_1j}}{\partial W_{k_2k_3}}\{\delta_{k_1k_2}X_{ik_3} + \delta_{k_1k_3}X_{ik_2}\}\\ &\qquad = 2\sum_{k_1=1}^{p}\sum_{k_2=1}^{p}\frac{1+\delta_{k_2k_1}}{2}\frac{\partial q_{k_1j}}{\partial W_{k_2k_1}}X_{ik_2}\\ &\qquad = 2(XD_WQ)_{ij}.\end{aligned}$$

(iii) The left hand-side can be expressed as

$$\sum_{k_1,k_3,k_4=1}^{p}\sum_{k_2=1}^{m}X_{k_2k_1}S_{k_3k_4}\frac{\partial}{\partial X_{k_2k_3}}q_{k_4k_1}$$

$$= \sum_{k_1,k_3,k_4=1}^{p} \sum_{k_2=1}^{m} X_{k_2k_1} S_{k_3k_4} \sum_{k_5,k_6=1}^{p} \frac{1+\delta_{k_5k_6}}{2} \frac{\partial q_{k_4k_1}}{\partial W_{k_5k_6}} \frac{\partial}{\partial X_{k_2k_3}} W_{k_5k_6}$$
$$= \sum_{k_1,k_3,k_4,k_5,k_6=1}^{p} \sum_{k_2=1}^{m} X_{k_2k_1} S_{k_3k_4} \frac{1+\delta_{k_5k_6}}{2} \frac{\partial q_{k_4k_1}}{\partial W_{k_5k_6}} (\delta_{k_3k_5} X_{k_2k_6} + \delta_{k_3k_6} X_{k_2k_5})$$
$$= 2 \sum_{k_1,k_3,k_4,k_5=1}^{p} S_{k_3k_4} W_{k_1k_5} \frac{1+\delta_{k_3k_5}}{2} \frac{\partial}{\partial W_{k_3k_5}} q_{k_4k_1}$$
$$= 2 \operatorname{tr} S(WD_W)'Q',$$

where the second equality follows from (2.4) and the last equality follows from symmetry of W and D_W. Using Lemma 2.1.5 and noting that $D_W W = ((p+1)/2)I_p$, we get

$$D_W WQ' = \frac{p+1}{2} Q' + (WD_W)'Q'.$$

Combining these two equalities gives the desired result.

Now we record the calculus lemmas on eigenstructures for the case when $m > p$. For notation, let $A = (a_{ij})$ be a $p \times p$ nonsingular matrix such that $A'SA = I_p$, $A'X'XA = \operatorname{diag}(F)$, $F = (f_1, \cdots, f_p)$, and $f_1 > f_2 > \cdots > f_p$ are the ordered eigenvalues of $X'XS^{-1}$. Furthermore, recall that $D_S = (d_S^{ij})$ with $d_S^{ij} = (1/2)(1+\delta_{ij})(\partial/\partial S_{ij})$ and $S = (S_{ij})$. The following lemma and its proof are obtained from Loh[36] by minor modification.

LEMMA 2.1.7. *Assume that $m > p$. With $A = (a_{ij})$ and $A^{-1} = (a^{ij})$, we have*

(i) $$d_S^{kl} f_i = - a_{li} a_{ki} f_i,$$

(ii) $$d_S^{kl} a^{ij} = \frac{1}{2} a^{ij} a_{ki} a_{li} + \frac{1}{2} \sum_{i' \neq i} a^{i'j} \left(a_{ki} a_{li'} + a_{li} a_{ki'} \right) \frac{f_{i'}}{f_{i'} - f_i},$$

(iii) $$d_W^{kl} f_i = a_{li} a_{ki},$$

(iv) $$d_W^{kl} a^{ij} = \frac{1}{2} \sum_{i' \neq i} a^{i'j} \left(a_{ki} a_{li'} + a_{li} a_{ki'} \right) \frac{1}{f_i - f_{i'}}.$$

PROOF. On differentiating $S = A'^{-1}A^{-1}$ and $W = A'^{-1} \operatorname{diag}(F) A^{-1}$, we have

$$dS = A'^{-1}(dA^{-1}) + (dA'^{-1})A^{-1}$$

and

$$dW = A'^{-1} \operatorname{diag}(F)(dA^{-1}) + (dA'^{-1}) \operatorname{diag}(F)A^{-1} + A'^{-1} \operatorname{diag}(dF)A^{-1}.$$

Multiplying these equations by A' on the left and by A on the right leads to

$$A'dSA = (dA^{-1})A + A'(dA'^{-1}) \tag{2.5}$$

and

$$A'dWA = \operatorname{diag}(F)(dA^{-1})A + A'(dA'^{-1}) \operatorname{diag}(F) + \operatorname{diag}(dF). \tag{2.6}$$

I For obtaining the derivatives with respect to S_{ij}, we may assume that $dW = 0$. From (2.5) and (2.6), we can see that

$$\operatorname{diag}(dF) = (dA^{-1})A \operatorname{diag}(F) - \operatorname{diag}(F)(dA^{-1})A - A'(dS)A \operatorname{diag}(F). \tag{2.7}$$

Then the (i,i) element of (2.7) becomes

$$df_i = -\sum_{k,j} a_{ji}(dS)_{jk} a_{ki} f_i.$$

For $i \neq j$, the (i,j) element of (2.7) provides

$$[(dA^{-1})A]_{ij} f_j - f_i[(dA^{-1})A]_{ij} - [A'(dS)A]_{ij} f_j = 0,$$

since $[\operatorname{diag}(dF)]_{ij} = 0$. This reduces to

$$[(dA^{-1})A]_{ij} = \frac{f_j}{f_j - f_i}[A'(dS)A]_{ij}. \tag{2.8}$$

Furthermore, from (2.5), we can get

$$[(dA^{-1})A]_{ii} = \frac{1}{2}[A'(dS)A]_{ii}. \tag{2.9}$$

Combining (2.8) and (2.9) leads to

$$\begin{aligned}
(dA^{-1})_{ij} &= \sum_{i'=1}^{p} [(dA^{-1})A]_{ii'} a^{i'j} \\
&= \frac{1}{2}[A'(dS)A]_{ii} a^{ij} + \sum_{i' \neq i} [A'(dS)A]_{ii'} a^{i'j} \frac{f_{i'}}{f_{i'} - f_i} \\
&= \sum_{k,l} \left\{ \frac{1}{2} a_{ki}(dS)_{kl} a_{li} a^{ij} + \sum_{i' \neq i} a_{ki}(dS)_{kl} a_{li'} a^{i'j} \frac{f_{i'}}{f_{i'} - f_i} \right\}.
\end{aligned}$$

Noting that

$$(dS)_{kl}\left(\frac{1}{2}(1+\delta_{k'l'})\frac{\partial}{\partial S_{k'l'}}\right)=\frac{1}{2}(\delta_{kk'}\delta_{ll'}+\delta_{kl'}\delta_{lk'}),$$

we can conclude that

$$d_s^{kl}a^{ij}=\frac{1}{2}a^{ij}a_{ki}a_{li}+\frac{1}{2}\sum_{i'\neq i}a^{i'j}(a_{ki}a_{li'}+a_{li}a_{ki'})\frac{f_{i'}}{f_{i'}-f_i}.$$

II For obtaining the derivatives with respect ot W_{ij}, we may assume that $dS=0$. Similarly we get

$$\mathrm{diag}\,(dF)=(dA^{-1})A\,\mathrm{diag}\,(F)-\,\mathrm{diag}\,(F)(dA^{-1})A+A'(dW)A. \tag{2.10}$$

Then (i,i) element of (2.10) becomes

$$df_i=\sum_{j,k}a_{ji}(dW)_{jk}a_{ki}. \tag{2.11}$$

For $i\neq j$, the (i,j) element of (2.10) provides that

$$[(dA^{-1})A]_{ij}f_j-f_i[(dA^{-1})A]_{ij}+[A'(dW)A]_{ij}=0,$$

which reduces to

$$[(dA^{-1})A]_{ij}=\frac{1}{f_i-f_j}[A'(dW)A]_{ij}.$$

From (2.5) and $dS=0$, we can see $[(dA^{-1})A]_{ii}=0$. Hence we can get

$$\begin{aligned}(dA^{-1})_{ij}&=\sum_{i'\neq i}[(dA^{-1})A]_{ii'}a^{i'j}\\&=\sum_{i'\neq i}\sum_{k,l}\left\{a_{ki}(dW)_{kl}a_{li'}a^{i'j}\frac{1}{f_i-f_{i'}}\right\}.\end{aligned}$$

Thus we can conclude

$$d_W^{kl}a^{ij}=\frac{1}{2}\sum_{i'\neq i}a^{i'j}\left(a_{ki}a_{li'}+a_{li}a_{ki'}\right)\frac{1}{f_i-f_{i'}},$$

which completes the proof.

The first part of the following lemma is also taken from Loh[36] by the minor modification.

LEMMA 2.1.8. *Using the notation as in* Lemma 2.1.7, *assume that*

$$\varphi(F) = \operatorname{diag}(\varphi_1(F), \varphi_2(F), \ldots, \varphi_p(F))$$

is differentiable on $\{f_1 > f_2 > \cdots > f_p\}$. *Then we have*

$$\text{(i)} \qquad \operatorname{tr} D_s A'^{-1}\varphi(F)A^{-1} = \sum_{k=1}^{p}\left\{p\varphi_k - f_k\varphi_{kk} - \sum_{l>k}\frac{f_k\varphi_k - f_l\varphi_l}{f_k - f_l}\right\},$$

and

$$\text{(ii)} \qquad \operatorname{tr} D_w A'^{-1}\varphi(F)A^{-1} = \sum_{k=1}^{p}\left\{\varphi_{kk} + \sum_{l>k}\frac{\varphi_k - \varphi_l}{f_k - f_l}\right\},$$

where $\varphi_{kk} = \partial\varphi_k(F)/\partial f_k$, $k = 1, 2, \ldots, p$.

PROOF. (i) The first part of this lemma is taken from Loh[36].

$$\begin{aligned}
\operatorname{tr} D_s A'^{-1}\varphi A^{-1} &= \sum_{i,j,k} d_s^{ij} a^{kj}\varphi_k a^{ki} \\
&= \sum_{i,j,k}\left\{\varphi_k a^{ki} d_s^{ij} a^{kj} + \varphi_k a^{kj} d_s^{ij} a^{ki} + a^{ki}a^{kj}\sum_{i'}\frac{\partial\varphi_k}{\partial f_{i'}} d_s^{ij} f_{i'}\right\} \\
&= \sum_{i,j,k}\left\{2\varphi_k a^{ki} d_s^{ij} a^{kj} + \sum_{i'} a^{ki}a^{kj}\frac{\partial\varphi_k}{\partial f_{i'}} d_s^{ij} f_{i'}\right\}.
\end{aligned}$$

The last equality holds since the symmetry of D_s. Now applying Lemma 2.1.7 gives that

$$\begin{aligned}
\operatorname{tr} D_s A'^{-1}\varphi A^{-1} &= \sum_{i,j,k}\left[\varphi_k a^{ki}\left\{a^{kj}a_{ik}a_{jk} + \sum_{i'\neq k} a^{i'j}(a_{ik}a_{ji'} + a_{jk}a_{ii'})\frac{f_{i'}}{f_{i'} - f_k}\right\}\right. \\
&\quad \left. - \sum_{i'} a^{ki}a^{kj}a_{ii'}a_{ji'}f_{i'}\frac{\partial\varphi_k}{\partial f_{i'}}\right] \\
&= \sum_k\left\{\varphi_k + \sum_{i'\neq k}\frac{f_{i'}\varphi_k}{f_{i'} - f_k} - f_k\frac{\partial\varphi_k}{\partial f_k}\right\} \\
&= \sum_k\left\{p\varphi_k - f_k\varphi_{kk} - \sum_{i'>k}\frac{f_k\varphi_k - f_{i'}\varphi_{i'}}{f_k - f_{i'}}\right\}.
\end{aligned}$$

The last equality holds since

$$\begin{aligned}
\sum_{i'\neq k}\frac{f_{i'}\varphi_k}{f_{i'} - f_k} &= \sum_{i'\neq k}\frac{f_{i'}\varphi_k - f_k\varphi_k + f_k\varphi_k}{f_{i'} - f_k} \\
&= (p-1)\varphi_k - \sum_{i'>k}\frac{f_k\varphi_k}{f_k - f_{i'}} - \sum_{i'<k}\frac{f_k\varphi_k}{f_k - f_{i'}} \\
&= (p-1)\varphi_k - \sum_{i'>k}\frac{f_k\varphi_k - f_{i'}\varphi_{i'}}{f_k - f_{i'}}.
\end{aligned}$$

(ii) Similarly we have

$$
\begin{aligned}
D_W A'^{-1}\varphi(F)A^{-1} &= \sum_{i,j,k} d_W^{ij} a^{kj}\varphi_k a^{ki} \\
&= \sum_{i,j,k}\left\{2\varphi_k a^{ki} d_W^{ij} a^{kj} + \sum_{i'} a^{ki}a^{kj}\frac{\partial \varphi_k}{\partial f_{i'}} d_W^{ij} f_{i'}\right\}.
\end{aligned}
$$

Now using Lemma 2.1.7 leads to

$$
\begin{aligned}
D_W A'^{-1}\varphi(F)A^{-1} &= \sum_{i,j,k}\left[\varphi_k a^{ki}\sum_{i'\neq k} a^{i'j}(a_{ik}a_{ji'} + a_{jk}a_{ii'})\frac{1}{f_k - f_{i'}} + \sum_{i'} a^{ki}a^{kj}a_{ji'}a_{ii'}\frac{\partial \varphi_k}{\partial f_{i'}}\right] \\
&= \sum_{k=1}^{p}\left\{\sum_{i'\neq k}\frac{\varphi_k}{f_k - f_{i'}} + \varphi_{kk}\right\} \\
&= \sum_{k=1}^{p}\left\{\varphi_{kk} + \sum_{i'>k}\frac{\varphi_k - \varphi_{i'}}{f_k - f_{i'}}\right\},
\end{aligned}
$$

which completes the proof.

Next we state calculus lemmas on eigenstructures for the case $m \le p$. For this end, let $\tilde{F} = (\tilde{F}_{ij}) = XS^{-1}X'$ and $D_{\tilde{F}} = (d_{\tilde{F}}^{ij})$ with $(d_{\tilde{F}}^{ij}) = (1/2)(1+\delta_{ij})(\partial/\partial \tilde{F}_{ij})$.

LEMMA 2.1.9. *Assume that $m \le p$. Let Q and T be matrix functions of $\tilde{F}$. Assuming all relevant products and derivatives exist as needed, we have*

(i) $$\nabla_X' Q = 2S^{-1}X'D_{\tilde{F}}Q,$$

(ii) $$\operatorname{tr}\nabla_X' QX = p\operatorname{tr}Q + \operatorname{tr}X\nabla_X' Q,$$

(iii) $$D_{\tilde{F}}QT' = (D_{\tilde{F}}Q)T' + \{Q'D_{\tilde{F}}'\}'T,$$

(iv) $$\operatorname{tr}(Q'D_{\tilde{F}}')'T' = \operatorname{tr}Q'D_{\tilde{F}}T.$$

PROOF. (i) The (i,j) element of $\nabla_X' Q$ is equal to

$$
\sum_{s_1=1}^{m}\sum_{s_2\le s_3}\frac{\partial q_{s_1 j}}{\partial \tilde{F}_{s_2 s_3}}\cdot\frac{\partial \tilde{F}_{s_2 s_3}}{\partial X_{s_1 i}} = \sum_{s_1,s_2,s_3=1}^{m}\frac{1+\delta_{s_2 s_3}}{2}\cdot\frac{\partial q_{s_1 j}}{\partial \tilde{F}_{s_2 s_3}}\cdot\frac{\partial \tilde{F}_{s_2 s_3}}{\partial X_{s_1 i}}, \tag{2.12}
$$

where $Q = (q_{ij})$. From $\tilde{F} = XS^{-1}X'$ and chain rule,

$$
\frac{\partial \tilde{F}_{s_2 s_3}}{\partial X_{s_1 i}} = \sum_{s_4=1}^{p}\{\delta_{s_1 s_2}S^{is_4}X_{s_3 s_4} + \delta_{s_1 s_3}S^{is_4}X_{s_2 s_4}\}, \tag{2.13}
$$

where $S^{-1} = (S^{ij})$. Putting (2.13) into (2.12) and using the symmetry of $\tilde{F}$, we get the desired result.

(ii) From Lemma 2.1.5, we get

$$\operatorname{tr} \nabla_X' Q X = \operatorname{tr} (Q' \nabla_X)' X + \operatorname{tr} (\nabla_X' Q) X.$$

Using that $\operatorname{tr}(AB)'C = \operatorname{tr} ABC'$ for matrices A, B, and C and noting that $\nabla_X X' = pI_m$, we can see that $\operatorname{tr}(Q'\nabla_X)'X = p \operatorname{tr} Q$, which completes the proof.

(iii) See Haff[20].

(iv) The proof follows from the straightforward calculation.

Recall that $F = (f_1, \ldots, f_m)$ and $f_1 > f_2 > \cdots > f_m$ are the ordered eigenvalues of $\tilde{F}$, or equivalently $XS^{-1}X'$.

The following lemma is taken from Stein[48] and Haff[22].

LEMMA 2.1.10. *Assume that $m \leq p$. Let $R = (R_1, R_2, \ldots, R_m)$, where R_k is the normalized column eigenvector corresponding to f_k, and let $\varphi(F) = \operatorname{diag}(\varphi_1(F), \ldots, \varphi_m(F))$ where $\varphi_k(F)(k = 1, \ldots, m)$ is a function from F to $[0, +\infty)$. Assuming that all relevant derivatives exist, we have*

(i) $$D_{\tilde{F}} f_k = R_k R_k',$$

(ii) $$D_{\tilde{F}} R_k = f_k^* R_k, \quad f_k^* = (1/2) \sum_{l \neq k} 1/(f_k - f_l),$$

and

(iii) $$D_{\tilde{F}} [R\varphi(F)R'] = R\varphi^{(1)}(F)R',$$

where $\varphi^{(1)}(F) = \operatorname{diag}(\varphi_1^{(1)}(F), \ldots, \varphi_m^{(1)}(F))$ and

$$\varphi_k^{(1)}(F) = \frac{1}{2} \sum_{l \neq k} \frac{\varphi_k(F) - \varphi_l(F)}{f_k - f_l} + \frac{\partial \varphi_k(F)}{\partial f_k}, \quad k = 1, \ldots, m.$$

PROOF. Taking the differential of $\tilde{F} = R \operatorname{diag}(F) R'$ we obtain

$$d\tilde{F} = (dR) \operatorname{diag}(F) R' + R \operatorname{diag}(F)(dR)' + R \operatorname{diag}(dF) R'.$$

Multiplying on the left by R' and on the right by R, we have

$$R'(d\tilde{F})R = (R'dR)\,\mathrm{diag}\,(F) + \,\mathrm{diag}\,(F)(R'dR)' + \,\mathrm{diag}\,(dF). \tag{2.14}$$

But, taking the differential of $R'R = I$, we have

$$(dR)'R + R'(dR) = 0.$$

This means that $R'dR$ is antisymmetric:

$$(R'dR)' + R'dR = 0. \tag{2.15}$$

Reverting to coordinates, we obtain from (2.14) and (2.15),

$$\begin{aligned}
(R'dR)_{kk} &= 0, \\
(R'dR)_{kl} &= \frac{1}{f_l - f_k}[R'(d\tilde{F})R]_{kl} \quad \text{for } k \neq l, \\
df_k &= [R'(d\tilde{F})R]_{kk}.
\end{aligned}$$

From above equations we may find that

$$(dR)_{kl} = (RR'dR)_{kl} = \sum_{i' \neq l} R_{ki'} \frac{1}{f_l - f_{i'}} [R'(d\tilde{F})R]_{i'l},$$

which gives that

$$d_{F}^{ij} R_{kl} = \frac{1}{2} \sum_{i' \neq l} \frac{R_{ki'}}{f_l - f_{i'}} (R_{ii'}R_{jl} + R_{ji'}R_{il}). \tag{2.16}$$

Also we can see that

$$d_{F}^{ij} f_k = R_{ik}R_{jk}. \tag{2.17}$$

Hence (i) is obtained from (2.17).

(ii) Using (2.16) we can see that the mth entry of $D_{F} R_k$ is given by

$$\begin{aligned}
\sum_u d_{F}^{mu} R_{uk} &= \frac{1}{2} \sum_u \sum_{i' \neq k} \frac{R_{ui'}}{f_k - f_{i'}} (R_{mi'}R_{uk} + R_{ui'}R_{mk}) \\
&= \frac{1}{2} \left[\sum_{i' \neq k} \frac{1}{f_k - f_{i'}} \right] R_{mk},
\end{aligned} \tag{2.18}$$

which completes the proof of (ii).

(iii) The (k,l) element of the matrix $D_F R\varphi R'$ becomes

$$\sum_{a,b} d_F^{ka} R_{ab}\varphi_b R_{lb} = \sum_{a,b} \varphi_b R_{lb} d_F^{ka} R_{ab} + \sum_{a,b} \varphi_b R_{ab} d_F^{ka} R_{lb} + \sum_{a,b} R_{ab} R_{lb} d_F^{ka} \varphi_b .$$

Denote these successive terms by A, B, and C respectively. From (2.18) we can see that

$$A = (1/2) \sum_b \varphi_b R_{lb} R_{kb} \sum_{i' \neq b} \frac{1}{f_b - f_{i'}}.$$

From (2.16) we can see that

$$\begin{aligned} B &= (1/2) \sum_b \sum_{i' \neq b} R_{ki'} R_{li'} \frac{\varphi_b}{f_b - f_{i'}} \\ &= (1/2) \sum_{i'} \sum_{b \neq i'} R_{ki'} R_{li'} \frac{\varphi_b}{f_b - f_{i'}}. \end{aligned}$$

Exchanging b with i', we may get that

$$B = -(1/2) \sum_b R_{kb} R_{lb} \sum_{i' \neq b} \frac{\varphi_{i'}}{f_b - f_{i'}}.$$

Furthermore,

$$\begin{aligned} C &= \sum_{a,b} R_{ab} R_{lb} \sum_m \frac{\partial \varphi_b}{\partial f_m} d_F^{ka} f_m \\ &= \sum_{a,b,m} R_{ab} R_{lb} R_{km} R_{am} \frac{\partial \varphi_b}{\partial f_m} \\ &= \sum_b R_{kb} R_{lb} \frac{\partial \varphi_b}{\partial f_b}. \end{aligned}$$

In summary, it is readily seen that $A + B + C$ is equal to the (k,l) element of $R\varphi^{(1)}R'$, which completes the proof of (iii).

REMARK 2.1.2. If $\varphi(F)$ is smooth enough, $D_F^n[R\varphi(F)R']$ can be obtained by recursion of (iii) of Lemma 2.1.10. We shall use notation $D_F^n[R\varphi(F)R'] = R\varphi^{(n)}(F)R'$ where $\varphi^{(n)}(F) = \operatorname{diag}(\varphi_1^{(n)}(F), \ldots, \varphi_m^{(n)}(F))$.

2.2. DERIVING THE CLASS OF ESTIMATORS

First let us recall the results by Stein[48] concerning the problem of estimating matrix of mean of normal populations with identity covariance matrix. Assume that

$$\tilde{X}; m \times p \;\sim\; N(\tilde{B}, I_m \otimes I_p) \qquad (m > p+1), \tag{2.19}$$

and let the loss be $\operatorname{tr}(\hat{\tilde{B}} - \tilde{B})'(\hat{\tilde{B}} - \tilde{B})$. He introduced the estimators of the form

$$\hat{\tilde{B}}(\tilde{X}) = \tilde{X} + \frac{1}{2}\nabla_{\tilde{X}} h(Y), \tag{2.20}$$

where $Y = (y_1, \ldots, y_p)$, $h(Y)$ is a scalar valued function from R^p or R^m to $[0, \infty)$ as needed, $y_1 > \cdots > y_p$ are the ordered eigenvalues of $\tilde{X}'\tilde{X}$, and $\nabla_{\tilde{X}} h(Y) = (\partial h(Y)/\partial \tilde{X}_{ij})$ for $\tilde{X} = (\tilde{X}_{ij})$. Then using the Normal identity and calculus on eigenstructure the unbiased estimate of the risk of these estimators is obtained in terms of the eigenvalues of $\tilde{X}'\tilde{X}$ and the first and second derivatives of $h(Y)$ with respect to Y. Next consider transformations $\tilde{X} \to \tilde{X}C$ for a $p \times p$ nonsingular matrix C. Then it is readily seen from Lemma 1.1 in Kariya and Sinha[25] that the model (2.19) is transformed into

$$\tilde{\tilde{X}} \;\sim\; N(\tilde{B}C, I_m \otimes (C'C))$$

and the estimator of the form (2.20) is changed into

$$\begin{aligned}\hat{\tilde{B}}C &= \tilde{X}C + \frac{1}{2}\nabla_{\tilde{X}} h(Y) \cdot C \\ &= \tilde{\tilde{X}} + \frac{1}{2}\nabla_{\tilde{\tilde{X}}} h(\tilde{Y}) \cdot C'C\end{aligned}$$

where $\nabla_{\tilde{\tilde{X}}} = (\partial/\partial \tilde{\tilde{X}}_{ij})$ for $\tilde{\tilde{X}} = (\tilde{\tilde{X}}_{ij}) = \tilde{X}C$, $\tilde{Y} = (\tilde{y}_1, \ldots, \tilde{y}_p)$, and $\tilde{y}_1 > \cdots > \tilde{y}_p > 0$ are the ordered eigenvalues of $\tilde{\tilde{X}}'\tilde{\tilde{X}}(C'C)^{-1}$. The last equality above holds since $\nabla_{\tilde{X}} = \nabla_{\tilde{\tilde{X}}} C'$.

Now let us return to the original model (2.1). If $\Sigma = C'C$ is unknown and the estimate of Σ, that is S, is available, it is natural to consider the estimator of the form

$$\hat{B}(X, S) = X + \frac{1}{2}\nabla_X h(F) \cdot S, \tag{2.21}$$

where $F = (f_1, \ldots, f_{min(m,p)})$ and $f_1 > \cdots > f_{min(m,p)} > 0$ are the ordered eigenvalues of $X'XS^{-1}$. It is readily checked that these are invariant under the group of transformations

$$(X, S) \to (OXC, C'SC),$$

where O is an $m \times m$ orthogonal matrix and C is a $p \times p$ nonsingular matrix.

The following representation of the second term on the right hand side of (2.21) gives better understanding of our estimators. For $m > p$, (2.21) can be rewritten as

$$\hat{B}(X, S) = X[I_p + AH(F)A^{-1}], \tag{2.22}$$

where $H(F) = \operatorname{diag}(h_1(F), h_2(F), \ldots, h_p(F))$, $h_k(F) = \partial h(F)/\partial f_k$, $k = 1, 2, \ldots, p$, A is a $p \times p$ nonsingular matrix such that $A'SA = I_P$ and $\operatorname{diag}(F) = A'X'XA$. This can be seen from the following argument. First use (ii) of Lemma 2.1.6, then we have

$$\frac{1}{2}\nabla_X h(F) = XD_W h(F).$$

From ordinary chain rule and (iii) of Lemma 2.1.7, it is seen that the (i, j) element of $D_W h(F)$ can be rewritten as

$$\begin{aligned} d_w^{ij} h(F) &= \sum_{k=1}^{p} h_k d_w^{ij} f_k \\ &= \sum_{k=1}^{p} h_k a_{ik} a_{jk}, \end{aligned}$$

from which it follows that $D_W h(F) = AH(F)A'$. Using the fact $A'SA = I_p$, we can conclude that

$$\frac{1}{2}\nabla_X h(F)S = XAH(F)A^{-1}.$$

For $m \leq p$, we have, from (i) of Lemma 2.1.9 and noting that $h(Y)$ is a scalar, that

$$\frac{1}{2}\nabla_X h(F) = D_F h(F)XS^{-1}.$$

Using ordinary chain rule and (i) of Lemma 2.1.10, we can see that

$$D_F h(F) = RH(F)R'.$$

From these, we can get that

$$\frac{1}{2}\nabla_X h(F)S = RH(F)R'X.$$

Finally it follows that , for $m \leq p$, (2.21) can be expressed as

$$\hat{B}(X,S) = [I_m + RH(F)R']X, \tag{2.23}$$

where $H(F) = \operatorname{diag}(h_1(F), \ldots, h_m(F))$ and $\operatorname{diag}(F) = R'XS^{-1}X'R$ with an $m \times m$ orthogonal matrix R.

It is worth noting that (2.21) becomes

$$\hat{B}(X,S) = \begin{cases} X[I_p + c_1(X'X)^{-1}S + c_2 I_p / \operatorname{tr}(X'X)S^{-1}], & for\ m > p+1; \\ [I_m + c_1(XS^{-1}X')^{-1} + c_2 I_m / \operatorname{tr}(X'X)S^{-1}]X & for\ p > m+1, \end{cases}$$

if we put $h(F) = c_1 \log(\prod_k f_k) + c_2 \log(\sum_k f_k)$. This shows that the estimators of the form (2.21) include the Efron-Morris type estimators.

For the case when $m > p+1$, we can consider different forms of estimators from that given by (2.21).

Recall that $W = X'X$ and let $W = OYO'$ in which $OO' = O'O = I_p$ and $Y = \operatorname{diag}(y_1, \ldots, y_p)$ with $y_1 > \cdots > y_p$ so that y_k is the k-th largest eigenvalues of W. We introduce the forms of estimators

$$\hat{B}(X,S) = X\left[I_p + \frac{1}{\operatorname{tr} S^{-1}} OT(Y)O'\right] \tag{2.24}$$

and

$$\hat{B}(X,S) = X\left[I_p + OT(Y)O'S\right], \tag{2.25}$$

where $T(Y) = \operatorname{diag}(t_1(Y), \ldots, t_p(Y))$ and $t_k(Y)$, $k = 1, \ldots, p$, is an absolutely continuous function of Y to $[0, \infty)$. The first one is a multivariate version of Stein-type estimator considered by Bilodeau and Kariya[5] and the second one includes Efron-Morris type estimator given by Bilodeau and Kariya[5].

2.3. UNBIASED ESTIMATE OF RISK

In this section we shall compute the unbiased estimate of the risk of an almost arbitrary equivariant estimator given by (2.21). First we start with a notation.

Let

$$T(n,m,p;h) = \sum_{k=1}^{p}\Big\{2(m-p+1)h_k + 4f_k h_{kk} + 4\sum_{l>k}\frac{f_k h_k - f_l h_l}{f_k - f_l} + (n+p-3)f_k h_k^2 - 4f_k^2 h_{kk} h_k - 2\sum_{l>k}\frac{f_k^2 h_k^2 - f_l^2 h_l^2}{f_k - f_l}\Big\},$$

where $h_k = \partial h(F)/\partial f_k$ and $h_{kk} = \partial^2 h(F)/\partial f_k^2$, $k = 1,2,\ldots,p$.

Now we have the following

THEOREM 2.3.1. *Assume that* $h(F)$ *satisfies the conditions*

$$\mathbf{E}\, h_k^2 < \infty, \quad \mathbf{E}\,|\partial h_k/\partial X_{ij}| < \infty, \quad \textit{and} \quad \mathbf{E}\,(\partial h_k/\partial S_{ij})^2 < \infty,$$

as well as the regularity conditions of Theorem 2.1 *in* Haff [18]. *Then the unbiased estimate of the risk of the estimator of the form* (2.21) *is given by*

(i) $$\hat{\mathbf{R}}((B,\Sigma),\hat{B}) = pm + T(n,m,p;h) \qquad \textit{for } m > p,$$

and

(ii) $$\hat{\mathbf{R}}((B,\Sigma),\hat{B}) = pm + T(n+m-p,p,m;h) \qquad \textit{for } p \geq m.$$

PROOF. (i) From Lemma 2.1.3, the unbiased risk estimate for (2.21), equivalently (2.22), can be written as

$$2\,\mathrm{tr}\,\nabla_X' XAH(F)A^{-1} + 2\,\mathrm{tr}\,D_S A'^{-1}\,\mathrm{diag}\,(F)H^2(F)A^{-1} + (n-p-1)\,\mathrm{tr}\,\mathrm{diag}\,(F)H^2(F) + pm, \tag{2.26}$$

where $D_S = (d_S^{ij})$ is a $p \times p$ differential operator whose element is given by $(1/2)(1 + \delta_{ij})\partial/\partial S_{ij}$ for $S = (S_{ij})$ and Kronecker's delta δ_{ij}. We shall compute (2.26) term by term. Using (i) of Lemma 2.1.6 and noting that $\nabla_X' X = mI_p$, it can be seen that the first term of (2.26) yields

$$2m\,\mathrm{tr}\,H(F) + 2\,\mathrm{tr}\,X'\nabla_X(AH(F)A^{-1})'. \tag{2.27}$$

Furthermore, from (ii) of Lemma 2.1.6, we get that

$$\begin{aligned} \operatorname{tr} X'\nabla_X (AH(F)A^{-1})' =& 2 \operatorname{tr} W D_w (AH(F)A^{-1})' \\ =& 2 \operatorname{tr} D_w W A H(F) A^{-1} - (p+1) \operatorname{tr} H(F), \end{aligned} \tag{2.28}$$

where $D_w = (d_w^{ij})$ and $d_w^{ij} = (1/2)(1+\delta_{ij})\partial/\partial w_{ij}$ for $W = (w_{ij})$ and a Kronecker's delta δ_{ij}. The last equality of (2.28) holds since $\operatorname{tr} D_w Q_1 Q_2 = \operatorname{tr}(Q_2 D_w Q_1 + Q_1' D_w Q_2')$ for $p \times p$ matrices Q_1 and Q_2 and $D_w W = ((p+1)/2) I_p$. Combining (2.27) with (2.28) and noting that $W = A'^{-1} \operatorname{diag}(F) A^{-1}$ lead to

$$2 \operatorname{tr} \nabla_X' X A H(F) A^{-1} = 2(m-p-1) \operatorname{tr} H(F) + 4 \operatorname{tr} D_w A'^{-1} \operatorname{diag}(F) H(F) A^{-1}. \tag{2.29}$$

Applying (ii) of Lemma 2.1.8 to the second term of (2.29) gives

$$\operatorname{tr} D_w A'^{-1} \operatorname{diag}(F) H(F) A^{-1} = \sum_{k=1}^{p} \left\{ f_k h_{kk} + h_k + \sum_{l>k} \frac{f_k h_k - f_l h_l}{f_k - f_l} \right\}, \tag{2.30}$$

and similarly we can see that the second term of (2.26) provides

$$\operatorname{tr} D_S A'^{-1} \operatorname{diag}(F) H^2(F) A^{-1} = \sum_{k=1}^{p} \left\{ -2 f_k^2 h_k h_{kk} + (p-1) f_k h_k^2 - \sum_{l>k} \frac{f_k^2 h_k^2 - f_l^2 h_l^2}{f_k - f_l} \right\}. \tag{2.31}$$

From (2.29) through (2.30) we complete the proof of (i).

(ii) Using (i) of Lemma 2.1.10, we can see that the estimator given by (2.21) is changed into $[I_m + D_{\tilde{F}} h(F)]X$ where $\tilde{F} = (\tilde{F}_{ij}) = XS^{-1}X'$ and $D_{\tilde{F}} = (d_{\tilde{F}}^{ij})$. Similarly, it follows, from Lemma 2.1.3, that the unbiased estimate of the risk for $X + D_{\tilde{F}} h(F) X$ can be expressed as

$$\begin{aligned} & 2 \operatorname{tr} \nabla_X' D_{\tilde{F}} h(F) X + 2 \operatorname{tr} D_S \{ X' (D_{\tilde{F}} h(F))^2 X \} \\ & \qquad + (n-p-1) \operatorname{tr} \tilde{F} (D_{\tilde{F}} h(F))^2 + pm. \end{aligned} \tag{2.32}$$

We shall calculate (2.32) term by term. From (i) and (ii) of Lemma 2.1.9 and symmetry of $D_{\tilde{F}} h(F)$, it follows that the first term in (2.32) becomes

$$2p \operatorname{tr} D_{\tilde{F}} h(F) + 2 \operatorname{tr} X \nabla_X' D_{\tilde{F}} h(F) = 2p \operatorname{tr} D_{\tilde{F}} h(F) + 4 \operatorname{tr} \tilde{F} D_{\tilde{F}}^2 h(F). \tag{2.33}$$

Next, applying Lemma 2.1.5 to the second term in (2.32) provides that

$$\begin{aligned} 2\ \mathrm{tr}\, D_S\{X'D_{\tilde{F}}h(F)\}(D_{\tilde{F}}h(F))X + 2\ \mathrm{tr}\left[\{X'D_{\tilde{F}}h(F)\}'D_S'\right]'(D_{\tilde{F}}h(F))X \\ = 4\ \mathrm{tr}\,(D_{\tilde{F}}h(F))XD_S\{X'D_{\tilde{F}}h(F)\}. \end{aligned} \tag{2.34}$$

The last equality follows from (iv) of Lemma 2.1.9. Using ordinary chain rule and (iii) of Lemma 2.1.9, we shall show that

$$XD_S\{X'D_{\tilde{F}}h(F)\} = -\tilde{F}D_{\tilde{F}}\left(\tilde{F}D_{\tilde{F}}h(F)\right) + \frac{m+1}{2}\tilde{F}D_{\tilde{F}}\ h(F). \tag{2.35}$$

It can be seen that the (i,j) element of $XD_S\{X'D_{\tilde{F}}h(F)\}$ is

$$\begin{aligned} &\sum_{i_1,i_2=1}^{p}\sum_{i_3=1}^{m} X_{ii_1} d_S^{i_1i_2}\left(X_{i_3i_2} d_{\tilde{F}}^{i_3j}\, h(F)\right) \\ &= \sum_{i_1,i_2=1}^{p}\sum_{i_3,u_1,u_4=1}^{m} X_{ii_1}X_{i_3i_2}\left(d_{\tilde{F}}^{u_1u_4}\, d_{\tilde{F}}^{i_3j}\, h(F)\right) d_S^{i_1i_2}\, \tilde{F}_{u_1u_4} \\ &= \sum_{i_1,i_2,u_2,u_3=1}^{p}\sum_{i_3,u_1,u_4=1}^{m} X_{ii_1}X_{i_3i_2}X_{u_1u_2}X_{u_4u_3}\left(d_{\tilde{F}}^{u_1u_4}\, d_{\tilde{F}}^{i_3j}\, h(F)\right) d_S^{i_1i_2}\, S^{u_2u_3}, \end{aligned} \tag{2.36}$$

where the last equality can be obtained by noting $\tilde{F}_{u_1u_4} = \sum_{u_2,u_3=1}^{p} X_{u_1u_2}S^{u_2u_3}X_{u_4u_3}$. Using Lemma 2.1.4, (2.36) becomes $-\tilde{F}_{iu_1}\tilde{F}_{i_3u_4}d_{\tilde{F}}^{u_1u_4}\, d_{\tilde{F}}^{i_3j}\, h(F)$, which follows that

$$\begin{aligned} XD_S\{X'D_{\tilde{F}}h(F)\} &= -\tilde{F}(\tilde{F}D_{\tilde{F}})'D_{\tilde{F}}h(F) \\ &= -\tilde{F}D_{\tilde{F}}\left(\tilde{F}D_{\tilde{F}}h(F)\right) + \frac{m+1}{2}\tilde{F}D_{\tilde{F}}h(F). \end{aligned}$$

The last equality holds since (iii) of Lemma 2.1.9 and $D_{\tilde{F}}\tilde{F} = \left((m+1)/2\right)I_m$. Hence, putting (2.35) into (2.34), it is seen that the second term in (2.32) can be rewritten as

$$-4\ \mathrm{tr}\left(D_{\tilde{F}}h(F)\right)\cdot \tilde{F}D_{\tilde{F}}\{\tilde{F}D_{\tilde{F}}h(F)\} + 2(m+1)\ \mathrm{tr}\,\tilde{F}\{D_{\tilde{F}}h(F)\}^2. \tag{2.37}$$

Combining (2.33) and (2.37) with (2.32) leads to

$$\begin{aligned} pm +\ \mathrm{tr}\Big[&2pD_{\tilde{F}}h(F) + 4\tilde{F}D_{\tilde{F}}^2h(F) + (n+2m-p+1)\tilde{F}\{D_{\tilde{F}}h(F)\}^2 \\ &- 4(D_{\tilde{F}}h(F))\tilde{F}D_{\tilde{F}}\{\tilde{F}D_{\tilde{F}}h(F)\}\Big]. \end{aligned} \tag{2.38}$$

Furthermore, using (iii) of Lemma 2.1.10 it follows that

$$\begin{aligned} \operatorname{tr} D_{\tilde{F}} h(F) &= \sum_{k=1}^{m} h_k(F), \\ \operatorname{tr} \tilde{F}\{D_{\tilde{F}} h(F)\}^2 &= \sum_{k=1}^{m} f_k h_k^2(F). \end{aligned} \tag{2.39}$$

Noting that $D_{\tilde{F}}^2 h(F) = D_{\tilde{F}} RH(F)R'$ where $H(F) = \operatorname{diag}(h_1(F), \ldots, h_m(F))$ and applying (iii) of Lemma 2.1.10 where $\varphi(F) = \operatorname{diag}(F)H(F)$, we can find that

$$\begin{aligned} \operatorname{tr} \tilde{F} D_{\tilde{F}}^2 h(F) &= \sum_{k=1}^{m} \left[f_k \left\{ h_{kk}(F) + \frac{1}{2} \sum_{l \neq k} \frac{h_k(F) - h_l(F)}{f_k - f_l} \right\} \right] \\ &= \sum_{k=1}^{m} \left[f_k h_{kk}(F) - \frac{m-1}{2} h_k(F) + \sum_{l > k} \frac{f_k h_k(F) - f_l h_l(F)}{f_k - f_l} \right]. \end{aligned} \tag{2.40}$$

Similarly, we can get that

$$\begin{aligned} &\operatorname{tr}(D_{\tilde{F}} h(F)) \tilde{F} D_{\tilde{F}} \{\tilde{F} D_{\tilde{F}} h(F)\} \\ &= \sum_{k=1}^{m} \left[f_k h_k(F) \left\{ h_k(F) + f_k h_{kk}(F) + \frac{1}{2} \sum_{l \neq k} \frac{f_k h_k(F) - f_l h_l(F)}{f_k - f_l} \right\} \right] \\ &= \sum_{k=1}^{m} \left[f_k h_k^2(F) + f_k^2 h_k(F) h_{kk}(F) + \frac{1}{2} \sum_{l > k} \frac{f_k^2 h_k^2(F) - f_l^2 h_l^2(F)}{f_k - f_l} \right]. \end{aligned} \tag{2.41}$$

Finally, putting (2.39), (2.40) and (2.41) into (2.38), we obtain the desired result.

REMARK 2.3.1. We use the same notation as in (2.19) and (2.20). Furthermore, let us define the similar notation in Theorem 2.3.1. Stein[48] showed that, for the case where $\Sigma = I_p$, the unbiased estimate of the risk of the estimator (2.20) is given by

$$\hat{\mathbf{R}} = \sum_{k=1}^{p} \left\{ 2(m-p+1)h_k + 4y_k h_{kk} + 4 \sum_{l > k} \frac{y_k h_k - y_l h_l}{y_k - y_l} + y_k h_k^2 \right\} + pm,$$

where $h_k = \partial h(Y)/\partial y_k$, $k = 1, \ldots, p$. Theorem 2.3.1 is a counterpart of the result of Stein.

Now let us compare the distribution of the eigenvalues of $X'XS^{-1}$ in the cases $m > p$ and $p \geq m$. Assume that $B = 0$ for simplicity. Then, from Muirhead[40], it can be seen that the joint density of the ordered eigenvalues of $X'XS^{-1}$ is , apart from normalizing constants,

$$\prod_{k=1}^{p} \frac{f_k^{(m-p-1)/2}}{(1+f_k)^{(n+m)/2}} \prod_{l>k}^{p}(f_k - f_l) \prod_{k=1}^{p} df_k$$

in the case $m > p$ while it is

$$\prod_{k=1}^{m} \frac{f_k^{(p-m-1)/2}}{(1+f_k)^{(n+m)/2}} \prod_{l>k}^{m}(f_k - f_l) \prod_{k=1}^{m} df_k$$

in the case $p \geq m$. It is easily checked that the second distribution can be obtained from the first one by making the substitutions

$$m \to p, \qquad p \to m, \qquad n \to n+m-p. \tag{2.42}$$

Theorem 2.3.1 tells us that the substitution rule (2.42) is valid to the unbiased estimate of risk and the estimator of the form (2.21) so that the estimator better than the usual estimator X in the case $m > p+1$ results in that in the case $p > m+1$ by using substitution rule (2.42).

2.4. ALTERNATIVE ESTIMATORS

In this section, using Theorem 2.3.1, a systematic search for alternative estimators is carried out.

2.4.1. THE VARIATIONAL FORM OF CERTAIN BAYES ESTIMATOR

Here we derive the variational form of Bayes estimator following an approach due to Haff[23].

First we concentrate on the case $m > p+1$. Let $\pi(\Lambda)$ be an orthogonally invariant prior distribution (i.e., $\pi(H\Lambda H') = \pi(\Lambda)$ for any orthogonal matrix H) where $\Lambda = (B'B)^{(1/2)}\Sigma^{-1}(B'B)^{(1/2)}$. Denote by $g(F|\lambda)$ the conditional density of $F = (f_1, \ldots, f_p)$ given $\lambda = (\lambda_1, \ldots, \lambda_p)$, λ_k $(k = 1, 2 \ldots, p)$, being the k-th largest eigenvalue of Λ. Finally the marginal density of F is denoted by

$$g_\pi(F) = \int g(F|\lambda) d\pi^*(\lambda),$$

where $\pi^*(\lambda) = \int_H \pi(H\Lambda H')dH$. Following argument in Haff[23] the Bayes risk of the estimator $\hat{B}(X,S) = X + (1/2)\nabla_X h(F) \cdot S$ is given by

$$r(\tilde{h}, d\tilde{h}, \pi) = \int \{pm + T(n,m,p;h)\} g_\pi(F) dF,$$

where $\tilde{h} = (h_1, \ldots, h_p)$, $d\tilde{h} = (h_{11}, \ldots, h_{pp})$. Since the loss function is convex, the formal Bayes rule is then unique and is obtained by minimizing the functional $r(\tilde{h}, d\tilde{h}, \pi)$. Theorem 2.1 in Haff[23] tells us that the minimizer $\tilde{h}$ must satisfy the Euler-Lagrange partial differential equations

$$\frac{\partial T}{\partial h_k} = \frac{\partial}{\partial f_k}\frac{\partial}{\partial h_{kk}} T + \left(\frac{\partial T}{\partial h_{kk}}\right)\left(\frac{\partial}{\partial f_k}\log g_\pi(F)\right), \qquad k = 1, 2, \ldots, p,$$

where, by regarding T as $T(n,m,p;h)$, the partial derivatives with respect to h_k are computed. It is readily checked that the solution of this system is given by

$$h_k = -\left[(m-p-1) + 2\sum_{l \neq k} \frac{f_k}{f_k - f_l} - 2f_k \frac{\partial}{\partial f_k}\log g_\pi(F)\right]$$
$$\Big/ \left[f_k\left\{(n+p+1) - 2\sum_{l \neq k}\frac{f_k}{f_k - f_l} + 2f_k\frac{\partial}{\partial f_k}\log g_\pi(F)\right\}\right], \qquad k = 1, 2, \ldots, p.$$

If $g_\pi(F)$ is a constant, then the estimator of B becomes

$$\hat{B}(X,S) = X[I_p + AH(F)A^{-1}], \tag{2.43}$$

where A is a $p \times p$ nonsingular matrix such that $A'SA = I_p$ and $A'X'XA = \operatorname{diag}(F)$, $H(F) = \operatorname{diag}(h_1, \ldots, h_p)$, and

$$h_k = -\left[m - p - 1 + 2\sum_{l \neq k} \frac{f_k}{f_k - f_l}\right] \Big/ \left[f_k\left\{n + p + 1 - 2\sum_{l \neq k} \frac{f_k}{f_k - f_l}\right\}\right], \qquad k = 1, 2, \ldots, p.$$

Next we shall consider the case for $p > m+1$. Recall that $F = (f_1, \cdots, f_m)$, $\operatorname{diag}(F) = R'XS^{-1}X'R$, and R is an $m \times m$ orthogonal matrix . Using substitution rule (2.42), we can see that the variational from of the Bayes estimator, $g_\pi(F)$ being a constant, becomes

$$\hat{B}(X,S) = [I_m + RH(F)R']X, \tag{2.44}$$

where $H(F) = \operatorname{diag}(h_1(F), \ldots, h_m(F))$ with

$$h_k = -\left[p - m - 1 + 2\sum_{l \neq k} \frac{f_k}{f_k - f_l}\right] \Big/ \left[f_k\left\{n + 2m - p + 1 - 2\sum_{l \neq k} \frac{f_k}{f_k - f_l}\right\}\right], \quad k = 1, 2, \ldots, m.$$

REMARK 2.4.1. Let $\tilde{X}'\tilde{X} = O \operatorname{diag}(Y)O'$ with $Y = (y_1, \ldots, y_p)$ and let O be a $p \times p$ orthogonal matrix. In the problem of estimating $\tilde{B}$ based on $\tilde{X}$ in (2.19), Stein[48] proposed the estimator of the form

$$\tilde{X}(I_p + O\phi(Y)O')$$

where $\phi(Y) = \operatorname{diag}(\phi_1(Y), \ldots, \phi_p(Y))$ and

$$\phi_k(Y) = -\left[m - p - 1 + 2\sum_{l \neq k} \frac{y_k}{y_k - y_l}\right] \Big/ y_k, \qquad k = 1, \ldots, p.$$

The estimators (2.43) and (2.44) are counterpart of that of Stein for the case when $\Sigma = I_p$ and the loss function $\operatorname{tr}(\hat{\tilde{B}} - \tilde{B})'(\hat{\tilde{B}} - \tilde{B})$.

REMARK 2.4.2. The estimators proposed are modified in such a way as to make h_k's are increasing sequence, but there may be some reversals of the order. It has not been established that the estimators proposed by (2.43) and (2.44) have a frequentist risk uniformly smaller than the commonly used estimator X.

2.4.2. EFRON-MORRIS TYPE ESTIMATORS

When $m > p+1$, Bilodeau and Kariya[5] obtained the Efron-Morris type estimators. That is of the form

$$\hat{B}(X,S) = X[I_p - a(X'X)^{-1}S - bS/\operatorname{tr}(X'X)],$$

with $a = (m-p-1)/(n+p+1)$ and $b = (p-1)/(n+p+1)$. Unfortunately, this form does not belong to the class given by (2.21) while it belongs to the class given by (2.25). In this section, we derive another Efron-Morris type estimators which are in the class of (2.21).

THEOREM 2.4.1. (i) *For the case where $m > p+1$, the estimator*

$$\hat{B}^{EM}(X,S) = X[I_p - a(X'X)^{-1}S - bI_p/\operatorname{tr}(X'X)S^{-1}] \tag{2.45}$$

is minimax relative to the loss function (2.2) *if $a = (m-p-1)/(n+p+1)$ and $b = (p^2+p-2)/(n-p+3)$.*

(ii) *For the case where $p > m+1$, the estimator*

$$\hat{B}^{EM}(X,S) = [I_m - a(XS^{-1}X')^{-1} - bI_m/\operatorname{tr}XS^{-1}X']X \tag{2.46}$$

is minimax relative to the loss function (2.2) *if $a = (p-m-1)/(n+2m-p+1)$ and $b = (m^2+m-2)/(n-p+3)$.*

PROOF. (i) Let

$$h^{(1)}(F) = -\{\log \prod_{k=1}^{p} f_k^a + \log(\sum_{k=1}^{p} f_k)^b\}, \tag{2.47}$$

where a and b are nonnegative constants. Set $h_k^{(1)} = \partial h_k^{(1)}/\partial f_k$ and $h_{kk}^{(1)} = \partial^2 h_k^{(1)}/\partial f_k^2$, $k = 1,\ldots,p$. We may observe that

$$h_k^{(1)} = -(\frac{a}{f_k} + \frac{b}{u}),$$

$$h_{kk}^{(1)} = (\frac{a}{f_k^2} + \frac{b}{u^2}),$$

$$4\sum_{k=1}^{p}\sum_{i>k}\frac{f_k h_k^{(1)} - f_i h_i^{(1)}}{f_k - f_i} = -\sum_{k=1}^{p}\frac{2(p-1)b}{u},$$

$$2\sum_{k=1}^{p}\sum_{i>k}\frac{f_k^2(h_k^{(1)})^2 - f_i^2(h_i^{(1)})^2}{f_k - f_i} = \sum_{k=1}^{p}\left\{\frac{2(p-1)ab}{u} + \frac{2(p-1)b^2}{pu}\right\},$$

where $u = \sum_{k=1}^{p} f_k$. Use (i) of Theorem 2.3.1 and note that $\sum_{k=1}^{p} f_k^2/u^2 \le \sum_{k=1}^{p}(1/p)$ and $\sum_{k=1}^{p} f_k = \sum_{k=1}^{p} u/p$. Then we get

$$\begin{aligned}
\Delta &= \mathbf{R}((B,\Sigma),X) - \mathbf{R}((B,\Sigma),\hat{B}^{EM}) \\
&\ge -\mathbf{E}\left\{\sum_{k=1}^{p}[\frac{1}{f_k}\big((n+p+1)a^2 - 2(m-p-1)a\big) \right. \\
&\quad \left. + \frac{1}{pu}\{(n-p+3)b^2 + 2\big(pn+2\big)ab - 2(mp-2)b\}]\right\}. \qquad (2.48)
\end{aligned}$$

The first term of the right side in (2.48) is minimized when $a = (m-p-1)/(n+p+1)$, in which the term is negative. For $a = (m-p-1)/(n+p+1)$, the second term is bounded above by

$$(n-p+3)b^2 - \frac{2(n+m)(p^2+p-2)b}{n+p+1} \le (n-p+3)b^2 - 2(p^2+p-2)b.$$

It is minimized when $b = (p^2+p-2)/(n-p+3)$ in which the term is negative. This completes the proof of first part.

(ii) The second part of this theorem can be obtained by using substitution rule (2.42). This completes the proof of the theorem.

2.4.3. ADJUSTED STEIN ESTIMATOR

We consider the approximation to the estimators given by (2.43) and (2.44), i.e., the approximation to the term $\sum_{l \neq k} f_k/(f_k - f_l)$ in these estimators. Since $\sum_{l \neq 1} f_1/(f_1 - f_l) > p - 1$ and $\sum_{l \neq p} f_p/(f_p - f_l) < 0$ for the case $m > p + 1$, so we replace the term by $p - k$ simply. Note that $\sum_{k=1}^{p} \sum_{l \neq k} f_k/(f_k - f_l) = \sum_{k=1}^{p}(p - k)$.

THEOREM 2.4.2. (i) *In the case $m > p + 1$ the estimator*

$$\hat{B}^{AD}(X, S) = X[I_p - AH^{AD}(F)A^{-1}],$$

where $H^{AD}(F) = \mathrm{diag}(d_1/f_1, \ldots, d_p/f_p)$ and $d_1 \geq \cdots \geq d_p$ are nonnegative constants, is minimax with respect to the loss function (2.2) *if $d_k = (m + p - 2k - 1)/(n - p + 2k + 1)$, $k = 1, 2, \ldots, p$.*

(ii) *In the case $p > m + 1$ the estimator*

$$\hat{B}^{AD}(X, S) = [I_m - RH^{AD}(F)R']X,$$

where $H^{AD}(F) = \mathrm{diag}(d_1/f_1, \ldots, d_m/f_m)$ with $d_1 \geq \cdots \geq d_m$, $\mathrm{diag}(F) = R'X'S^{-1}XR$, and R is an $m \times m$ orthogonal matrix, is minimax with respect to the loss function (2.2) *if $d_k = (p + m - 2k - 1)/(n - p + 2k + 1)$, $k = 1, 2, \ldots, m$.*

PROOF. (i) Let

$$h^{(2)}(F) = -\sum_{k=1}^{p} d_k \log f_k$$

in (2.22). Similar calculation to the proof of Theorem 2.4.1 shows that the risk difference between X and the proposed estimator is

$$\begin{aligned}\Delta =& \mathbf{R}((B, \Sigma), X) - \mathbf{R}((B, \Sigma), \hat{B}^{AD}) \\ =& -\mathbf{E}\left[\sum_{k=1}^{p}\left\{-2(m - p - 1)\frac{d_k}{f_k} + (n + p + 1)\frac{d_k^2}{f_k}\right\}\right] \\ &+ \mathbf{E}\left[\sum_{k=1}^{p}\sum_{l>k}\frac{4d_k + 2d_k^2 - (4d_l + 2d_l^2)}{f_k - f_l}\right].\end{aligned} \tag{2.49}$$

Set $y_k = 4d_k + 2d_k^2$. Note that $y_1 \geq \cdots \geq y_p$. Then we get

$$\begin{aligned}\sum_{k=1}^{p}\sum_{t>k}\frac{y_k - y_t}{f_k - f_t} &= \sum_{k=1}^{p}\frac{1}{f_k}\sum_{t>k}\frac{f_k}{f_k - f_t}(y_k - y_t) \\ &\geq \sum_{k=1}^{p}\frac{1}{f_k}\sum_{t>k}(y_k - y_t) = \sum_{k=1}^{p}\frac{(p-k)y_k - \sum_{t>k} y_t}{f_k} \geq 0,\end{aligned} \tag{2.50}$$

since $f_k/(f_k - f_t) > 1$ for $t > k$. From (2.50), (2.49) is bounded below by

$$-\mathbf{E}\left[\sum_{k=1}^{p}\frac{1}{f_k}\left\{(n-p+1+2k)d_k^2 - 2(m+p-1-2k)d_k + \sum_{t>k}(4d_t + 2d_t^2)\right\}\right].$$

Define

$$z_k(d_k) = (n-p+1+2k)d_k^2 - 2(m+p-1-2k)d_k + \sum_{t>k}(4d_t + 2d_t^2), \ k = 1, \ldots, p,$$

where $d_1 \geq d_2 \geq \cdots \geq d_p$. It is sufficient to prove that $z_k(d_k)$ is negative when

$$d_k = \frac{m+p-1-2k}{n-p+1+2k} \ (\text{say} \, d_k^0), \quad k = 1, \ldots, p.$$

We shall fix $d_{k+k'} = d_{k+k'}^0$ $(k' = 1, \ldots, p-k)$ to choose d_k. Then we can see that $z_k(d_k)$ is minimized when $d_k = d_k^0$ and that

$$z_k(d_k^0) < z_k(d_{k+1}^0) = z_{k+1}(d_{k+1}^0) < \cdots < z_p(d_p^0) < 0,$$

since $d_1^0 > \cdots > d_p^0$. This completes the proof.

(ii) The minor modification of the proof of (i) leads to the desired result.

REMARK 2.4.3. Let

$$\hat{B}^{CEM} = X[I_p - \{(m-p-1)/(n+p+1)\}(X'X)^{-1}S].$$

This estimator was proposed in Efron and Morris[14] and called the crude Efron-Morris estimator. Using the notation as in the proof of Theorem 2.4.2, it can be seen that

$$\mathbf{R}((B,\Sigma),X) - \mathbf{R}((B,\Sigma),\hat{B}^{CEM}) = -\mathbf{E}\left[\sum_{k=1}^{p}\frac{1}{f_k}z_p(d_p^0)\right]$$

and

$$\mathbf{R}((B,\Sigma),X) - \mathbf{R}((B,\Sigma),\hat{B}^{AD}) > -\mathbf{E}\left[\sum_{k=1}^{p} \frac{1}{f_k} z_k(d_k^0)\right].$$

These give that

$$\mathbf{R}((B,\Sigma),\hat{B}^{CEM}) - \mathbf{R}((B,\Sigma),\hat{B}^{AD}) > -\mathbf{E}\left[\sum_{k=1}^{p} \frac{1}{f_k}\{z_k(d_k^0) - z_p(d_p^0)\}\right] > 0,$$

since $z_p(d_p^0) \geq z_k(d_k^0)$, $k = 1, 2, \cdots, p$. This concludes that the adjusted estimator $\hat{B}^{AD}$ is better than the crude Efron-Morris estimator $\hat{B}^{CEM}$.

2.4.4. BARANCHIK TYPE ESTIMATORS

The following theorem is a generalization of the results of Baranchik[1] and Lin and Tsai[35] who treated the case of $m > 3$ and $p = 1$, and of $p > 3$ and $m = 1$, respectively.

THEOREM 2.4.3. (i) *Assume that $m > p+1$. Let $A'SA = I_p$,* $\mathrm{diag}(F) = A'X'XA$ *where $F = (f_1, \ldots, f_p)$, and let A is a $p \times p$ nonsingular matrix. Let $\gamma_k(t)$ $(k = 1, 2, \ldots, p)$ be functions satisfying*

(a) $$0 \leq \gamma_k(t) \leq 2(m-p-1)/(n+p+1),$$

(b) $$\gamma_k(t) \text{ is nondecreasing in } t \quad , k = 1, \cdots p,$$

(c) $$\gamma_1(t) \geq \gamma_2(t) \geq \cdots \geq \gamma_p(t) \text{ for } \forall t \geq 0.$$

Then the estimator

$$\hat{B}^{BA}(X,S) = X[I_p - AH^{(\gamma)}(F)A^{-1}],$$

where $\gamma = (\gamma_1, \ldots, \gamma_p)$ and

$$H^{(\gamma)}(F) = diag(\gamma_1(f_1)/f_1, \ldots, \gamma_p(f_p)/f_p),$$

is minimax relative to the loss function (2.2).

(ii) *Assume that $p > m+1$. Let* $\mathrm{diag}(F) = R'XS^{-1}X'R$ *where $F = (f_1, \ldots, f_m)$ and R is an $m \times m$ nonsingular matrix. Let $\gamma_k(t)$ $(k = 1, 2, \ldots, m)$ be functions satisfying*

(a′) $$0 \leq \gamma_k(t) \leq 2(p-m-1)/(n+2m-p+1)$$

as well as the conditions (b) *and* (c) (*replacing* p *by* m). *Then the estimator*

$$\hat{B}^{BA}(X,S) = [I_m - RH^{(\gamma)}(F)R']X,$$

where $\gamma = (\gamma_1, \dots, \gamma_p)$ *and*

$$H^{(\gamma)}(F) = \operatorname{diag}(\gamma_1(f_1)/f_1, \dots, \gamma_m(f_m)/f_m),$$

is minimax relative to the loss function (2.2).

PROOF. (i) Similar to the proof of Theorem 2 in Zheng[54], first we suppose that $\gamma_k(t),\ k = 1,\cdots,p$, are absolutely continuous and have bounded derivatives on $[0,\infty)$. We shall use the notation $\hat{B}(\gamma)$ instead of $\hat{B}^{BA}(X,S)$ for convenience. Using (i) of Theorem 2.3.1 with $h_k = -\gamma_k(f_k)/f_k$, we get that

$$\begin{aligned}
\Delta =& \mathbf{R}((B,\Sigma),X) - \mathbf{R}((B,\Sigma),\hat{B}(\gamma)) \\
=& -\mathbf{E}\left[\sum_{k=1}^{p}\left\{\frac{(n+p+1)\gamma_k(f_k)}{f_k}\left(\gamma_k(f_k) - \frac{2(m-p-1)}{n+p+1}\right)\right\}\right] \\
&+ \mathbf{E}\left[\sum_{k=1}^{p}\left\{4\left(1+\gamma_k(f_k)\right)\frac{\partial\gamma_k(f_k)}{\partial f_k} + 4\sum_{t>k}\left\{1+\frac{1}{2}\left(\gamma_k(f_k)+\gamma_t(l_t)\right)\right\}\right.\right. \\
&\left.\left.\times\left\{\frac{\gamma_k(f_k)-\gamma_t(f_t)}{f_k - f_t}\right\}\right\}\right].
\end{aligned} \tag{2.51}$$

From the above and the conditions on γ_k, we find that $\Delta \geq 0$, which follows that $\hat{B}(\gamma)$ is minimax.

Suppose now that $\gamma_k(t)$'s , $k = 1,\cdots,p$, are general functions satisfying the conditions of the theorem. Let $\gamma^{(i)} = (\gamma_1^{(i)},\dots,\gamma_p^{(i)})$ where $\gamma_k^{(i)}$'s ,$k = 1,\cdots,p$, are functions which are absolutely continuous and have bounded derivatives; such that $\gamma_k^{(i)}(t)$ converges to $\gamma_k(t)$ as $i \to +\infty$. Then the estimator $\hat{B}(\gamma^{(i)})$ converges to $\hat{B}(\gamma)$ (a.e.) as $i \to +\infty$. From (2.51), it follows that $\hat{B}(\gamma^{(i)})$ is minimax and $\| \hat{B}(\gamma^{(i)}) - B \|^2$ has the bounded expectation independent of i. So $\hat{B}(\gamma)$ is minimax. This completes the proof.
(ii) The minor modification in the proof of (i) leads to the desired result.

Using Theorem 2.4.3 we will now give examples of minimax estimators.

EXAMPLE 2.4.1. (i) In the case $m > p+1$, set γ_k, $k = 1, 2, \ldots, p$, all equal to a constant $(m-p-1)/(n+p+1)$. Then we obtain the crude Efron-Morris estimator $\hat{B}^{CEM}$ given in Remark 2.4.3.

(ii) In the case $p > m+1$, set γ_k, $k = 1, 2, \ldots, p$, all equal to a constant $(p-m-1)/(n+2m-p+1)$. Then we obtain the estimator

$$\hat{B}(X, S) = \left[I_m - \frac{p-m-1}{n+2m-p+1} (XS^{-1}X')^{-1} \right] X.$$

EXAMPLE 2.4.2. (i) In the case $m > p+1$ set

$$\gamma_k(t) = \frac{m-p-1}{n+p+1} - \frac{2}{n+p+1} \left[\int_0^1 \frac{(1+t)^{(n+m)/2}}{(1+t\lambda)^{(n+m)/2+1}} \lambda^{(m-p-3)/2} d\lambda \right]^{-1}. \tag{2.52}$$

As in Lin and Tsai[35], it can be seen that these γ_k, $k = 1, 2, \ldots, p$, satisfy the conditions (a), (b), and (c) of Theorem 2.4.3.

(ii) In the case $p > m+1$, set

$$\gamma_k(t) = \frac{p-m-1}{n+2m-p+1} - \frac{2}{n+2m-p+1} \left[\int_0^1 \frac{(1+t)^{(n+m)/2}}{(1+t\lambda)^{(n+m)/2+1}} \lambda^{(p-m-3)/2} d\lambda \right]^{-1}. \tag{2.53}$$

Similarly, it can be seen that these γ_k's satisfy the conditions (a$'$), (b), and (c).

REMARK 2.4.4. When $p > 3$ and $m = 1$ or $m > 3$ and $p = 1$, based on the method of Brown[7] and Brewster and Zidek[6], Kubokawa[32] showed that the Baranchik type estimator using (2.52) or (2.53) beats the crude Stein estimator relative to the loss function (2.2). However, such frequentist risk result in the multivariate case has not been established.

2.5. OTHER CLASS OF MINIMAX ESTIMATORS

For $m > p+1$, we consider other class of minimax estimators which do not belong to (2.21).

Recall that $X'X = OYO'$ in which $OO' = O'O = I_p$ and $Y = \operatorname{diag}(y_1, \ldots, y_p)$ so that $y_1, \ldots, y_p$ are ordered eigenvalues of $X'X$. We introduced two classes of estimators of the form

$$X[I_p + \frac{1}{\operatorname{tr} S^{-1}} OT(Y)O'] \tag{2.54}$$

and

$$X[I_p + OT(Y)O'S] \tag{2.55}$$

where $T(Y) = \operatorname{diag}(t_1(Y), \ldots, t_p(Y))$ and $t_i(Y)$ $(i = 1, \ldots, p)$ are an absolutely continuous function of Y. Using Lemmas in Section 2.1 we can get the unbiased risk estimates for these forms given by (2.54) and (2.55) respectively.

THEOREM 2.5.1 (i) *Assume that* $T(Y)$ *satisfies the regularity conditions needed to establish* Lemma 2.1.3. *The unbiased estimate of the risk of the estimators given by* (2.54) *with respect to the loss function* (2.2) *is*

$$\begin{aligned} pm + \left(1/\operatorname{tr} S^{-1}\right) \sum_{i=1}^{p} \left\{ 2(m-p+1)t_i(Y) + 4y_i \frac{\partial t_i(Y)}{\partial y_i} + 4\sum_{j>i} \frac{y_i t_i(Y) - y_j t_j(Y)}{y_i - y_j} \right\} \\ + \frac{4}{(\operatorname{tr} S^{-1})^3} \operatorname{tr} \left(OYT^2(Y)O'S^{-2}\right) + \frac{n-p-1}{(\operatorname{tr} S^{-1})^2} \operatorname{tr} \left(OYT^2(Y)O'S^{-1}\right). \end{aligned} \tag{2.56}$$

(ii) *The unbiased estimate of risk of the estimators given by* (2.55) *is*

$$pm + \operatorname{tr} OC(Y)O'S \tag{2.57}$$

where $C(Y) = \operatorname{diag}(c_1(Y), \ldots, c_p(Y))$ *and*

$$c_i(Y) = (n+p+1)y_i t_i^2(Y) + 2(m-p+1)t_i(Y) + 4y_i \frac{\partial t_i(Y)}{\partial y_i} + 2\sum_{j \neq i} \frac{y_i t_i(Y) - y_j t_j(Y)}{y_i - y_j}.$$

PROOF. (i) From Lemma 2.1.3, the unbiased estimate of the risk of the estimators given by (2.54) becomes

$$\begin{aligned} pm &+ \frac{2}{\operatorname{tr} S^{-1}} \operatorname{tr}(\nabla_x' XOT(Y)O') + 2\operatorname{tr}(D_s \frac{1}{(\operatorname{tr} S^{-1})^2} OYT^2(Y)O') \\ &+ \frac{n-p-1}{(\operatorname{tr} S^{-1})^2} \operatorname{tr}(OYT^2(Y)O'S^{-1}). \end{aligned} \tag{2.58}$$

Using (i) and (ii) of Lemma 2.1.6 with $Q = OT(Y)O'$ it is seen that

$$\begin{aligned} \operatorname{tr} \nabla_x' XOT(Y)O' &= m \operatorname{tr} OT(Y)O' + \operatorname{tr} X'\nabla_x OT(Y)O' \\ &= m \operatorname{tr} T(Y) + 2 \operatorname{tr} D_w WOT(Y)O' - (p+1) \operatorname{tr} T(Y) \end{aligned} \tag{2.59}$$

where $W = X'X$. The last equality holds since

$$\operatorname{tr} D_w WOT(Y)O' = \operatorname{tr} WD_w OT(Y)O' + ((p+1)/2) \operatorname{tr} T(Y).$$

Using (iii) of Lemma 2.1.10 (replacing F by W) and noting that

$$(1/2) \sum_{i=1}^{p} \sum_{j \neq i} \frac{y_i t_i(Y) - y_j t_j(Y)}{y_i - y_j} = \sum_{i=1}^{p} \sum_{j>i} \frac{y_i t_i(Y) - y_j t_j(Y)}{y_i - y_j}$$

we get

$$\begin{aligned} \operatorname{tr} D_w WOT(Y)O' &= \operatorname{tr} D_w OYT(Y)O' \\ &= \sum_{i=1}^{p} \left\{ y_i \frac{\partial t_i(Y)}{\partial y_i} + t_i(Y) + \sum_{j>i} \frac{y_i t_i(Y) - y_j t_j(Y)}{y_i - y_j} \right\}. \end{aligned} \tag{2.60}$$

From Lemma 2.1.4 we may see that $D_s(1/(\operatorname{tr} S^{-1})^2) = 2(\operatorname{tr} S^{-1})^{-3} S^{-2}$. Using this fact and putting (2.59) and (2.60) into (2.58) give the desired result.

(ii) Similarly we get an unbiased estimate of the risk of the estimators given by (2.55) as

$$pm + 2 \operatorname{tr} \nabla_x' XOT(Y)O'S + 2 \operatorname{tr} D_s SOYT^2(Y)O'S + (n-p-1) \operatorname{tr} OYT^2(Y)O'S. \tag{2.61}$$

From (i) and (iii) of Lemma 2.1.6 we get

$$\operatorname{tr} \nabla_x' XOT(Y)O'S = (m-p-1) \operatorname{tr} OT(Y)O'S + 2 \operatorname{tr} SD_w OYT(Y)O'. \tag{2.62}$$

Using Lemma 2.1.5 and noting that $\operatorname{tr}(AB)'C = \operatorname{tr} ABC'$ for matrices A, B, and C, the third term of (2.61) becomes

$$\begin{aligned} &\operatorname{tr}(D_s SOYT^2(Y)O')S + \operatorname{tr}[\{SOYT^2(Y)O'\}'D_s]'S \\ &= (p+1) \operatorname{tr} OYT^2(Y)O'S \end{aligned} \tag{2.63}$$

Putting (2.62) and (2.63) into (2.61) and using Lemma 2.1.6 straightforward calculation lead to the desired result.

Noting that $\operatorname{tr} AB \le (\operatorname{tr} A)(\operatorname{tr} B)$ for any $p \times p$ positive definite matrices A, B and that the risk of the unbiased estimator X is equal to pm, we obtain the following corollary from (i) of Theorem 2.5.1.

COROLLARY 2.5.1 *The estimator given by* (2.54) *is minimax with respect to the loss function* (2.2) *if*

$$\sum_{i=1}^{p}\Big\{(n-p+3)y_i t_i^2(Y) + 2(m-p+1)t_i(Y) + 4y_i\frac{\partial t_i(Y)}{\partial y_i} + 4\sum_{j>i}\frac{y_i t_i(Y) - y_j t_j(Y)}{y_i - y_j}\Big\} \le 0.$$

2.5.1. OTHER FORMS OF BARANCHIK ESTIMATORS

Here we give a class of minimax estimators derived from (2.54) and (2.55). Define $T(Y)$ by

$$t_i(Y) = -\frac{\gamma_i(y_i)}{y_i}, \qquad i = 1, \cdots, p$$

where $\gamma_i(t)$ is an absolutely continuous and nonnegative function of $t > 0$.

THEOREM 2.5.2. *Assume that* $m > p+1$. *Let* $\gamma_i(t)$, $i = 1, \ldots, p$, *be functions satisfying*

(i) $\qquad \gamma_i(t)$ *is nondecreasing in* t,

(ii) $\qquad \gamma_1(t) \ge \gamma_2(t) \ge \cdots \ge \gamma_p(t) \qquad$ *for* $\forall t > 0$.

Let

$$T^{(\gamma)}(Y) = \operatorname{diag}\Big(\frac{\gamma_1(y_1)}{y_1}, \frac{\gamma_2(y_2)}{y_2}, \ldots, \frac{\gamma_p(y_p)}{y_p}\Big)$$

where $\gamma = (\gamma_1(y_1), \gamma_2(y_2), \cdots, \gamma_p(y_p))$, $Y = \operatorname{diag}(y_1, y_2, \cdots, y_p)$, *and* O *is an orthogonal matrix such that* $X'X = OYO'$. *Then the estimator*

$$\hat{B}_1^{(\gamma)}(X, S) = X[I_p - (1/\operatorname{tr} S^{-1})OT^{(\gamma)}(Y)O']$$

is minimax if

(iii) $$0 \le \gamma_i(t) \le \frac{m-p-1}{n-p+3}, \qquad i = 1, 2, \ldots, p,$$

and the estimator

$$\hat{B}_2^{(\gamma)}(X, S) = X[I_p - OT^{(\gamma)}(Y)O'S]$$

is minimax if

(iii)′ $$0 \le \gamma_i(t) \le \frac{m-p-1}{n+p+1}, \qquad i = 1, 2, \ldots, p,$$

with respect to the loss function (2.2).

PROOF. We shall prove the minimaxity of $\hat{B}_2^{(\gamma)}$. Then the minimaxity of $\hat{B}_1^{(\gamma)}$ is obtained in the same way.

Similar to the proof of Theorem 2 in Zheng[54], first we supposed that $\gamma_i(t)$, $i = 1, 2, \ldots, p$, are absolutely continuous and have bounded derivative on $[0, \infty)$. Using (2.57) we get that

$$\begin{aligned} \Delta(\gamma) =& \mathbf{R}((B, \Sigma), X) - \mathbf{R}((B, \Sigma), \hat{B}_2^{(\gamma)}) \\ =& -\mathbf{E}[\operatorname{tr} OC^{(\gamma)}(Y)O'S] \end{aligned} \tag{2.64}$$

where $C^{(\gamma)}(Y) = diag(c_1^{(\gamma)}(Y), c_2^{(\gamma)}(Y), \ldots, c_p^{(\gamma)}(Y))$ and

$$\begin{aligned} c_i^{(\gamma)}(Y) =& (n+p+1)\frac{\gamma_i^2(y_i)}{y_i} - 2(m-p-1)\frac{\gamma_i(y_i)}{y_i} - 4y_i\frac{\partial \gamma_i(y_i)}{\partial y_i} \\ & - 2\sum_{j \ne i} \frac{\gamma_i(y_i) - \gamma_j(y_j)}{y_i - y_j}, \qquad i = 1, 2, \ldots, p. \end{aligned} \tag{2.65}$$

From the conditions (i), (ii), and (iii)′, it follows that $c_i^{(\gamma)}(Y) \le 0$. So $\Delta(\gamma) \ge 0$ since S is positive definite matrix.

Suppose now that $\gamma_i(t)$, $i = 1, 2, \ldots, p$, are general functions satisfying the conditions (i), (ii), and (iii)′. Let $\gamma^{(k)} = (\gamma_1^{(k)}(t), \gamma_2^{(k)}(t), \ldots, \gamma_p^{(k)}(t))$ where $\gamma_i^{(k)}$, $i = 1, 2, \ldots, p$, are functions which are absolutely continuous and have bounded derivative; such that $\gamma_i^{(k)}(t)$ converges to $\gamma_i(t)$ as $k \to +\infty$. Then $\hat{B}_2^{(\gamma^{(k)})}$ converges to $\hat{B}_2^{(\gamma)}$ (a.e.) as $k \to +\infty$. From (2.64) and (2.65), it follows that $\hat{B}_2^{(\gamma^{(k)})}$ is minimax and $\| \hat{B}_2^{(\gamma^{(k)})} - B \|^2$ has bounded expectation independent of k. So $\hat{B}_2^{(\gamma)}$ is minimax. This completes the proof.

EXAMPLE 2.5.1. Setting $\gamma_i(t) = (m-p-1)/(n-p+3)$ in $B_1^{(\gamma)}(X,S)$ and $\gamma_i(t) = (m-p-1)/(n+p+1)$ in $B_2^{(\gamma)}(X,S)$ yields Stein-type shrinkage estimators

$$B_{s1}(X,S) = X\left[I_p - \frac{m-p-1}{n-p+3}\frac{(X'X)^{-1}}{\operatorname{tr} S^{-1}}\right] \tag{2.66}$$

and

$$B_{s2}(X,S) = X\left[I_p - \frac{m-p-1}{n+p+1}(X'X)^{-1}S\right] \qquad (\text{Efron} - \text{Morris}[15]) \tag{2.67}$$

respectively. Note that both estimators reduce to the usual Stein estimator for the case $p = 1$.

EXAMPLE 2.5.2. Let $\gamma_i(y_i) = d_1(1 + d_1 y_i^{-1})^{-1}$ in $\hat{B}_1^{(\gamma)}(X,S)$ and $\gamma_i(y_i) = d_2(1 + d_2 y_i^{-1})^{-1}$ in $\hat{B}_2^{(\gamma)}(X,S)$ for $0 \leq d_1 \leq (m-p-1)/(n-p+3)$ and $0 \leq d_2 \leq (m-p-1)/(n+p+1)$ respectively. Using Theorem 2.5.2 we get that the estimators

$$\hat{B}_1^{(\gamma)}(X,S) = X\left(I_p - d_1(X'X + d_1 I_p)^{-1}/\operatorname{tr} S^{-1}\right)$$

and

$$\hat{B}_2^{(\gamma)}(X,S) = X\left(I_p - d_2(X'X + d_2 I_p)^{-1} S\right)$$

are minimax.

2.5.2. ADJUSTED ESTIMATOR

Now we give another multivariate extension (being of the form (2.54)) of the Stein estimator.

THEOREM 2.5.3. *Let*

$$\hat{B}(X,S) = X[I_p - OT^{(1)}(Y)O'/\operatorname{tr} S^{-1}] \tag{2.68}$$

where $T^{(1)}(Y) = \operatorname{diag}(d_1/y_1, d_2/y_2, \ldots, d_p/y_p)$ *and* $d_1, d_2, \ldots, d_p$ *are constants with* $d_1 \geq d_2 \geq \cdots \geq d_p > 0$. *Then the estimator* (2.68) *is minimax relative to the loss function* (2.2) *when* $d_i = (m+p-2i-1)/(n-p+3)$, $i = 1, 2, \ldots, p$.

PROOF. From Corollary 2.5.1 it suffices to show that

$$\Delta = \sum_{i=1}^{p} \{(n-p+3)\frac{d_i^2}{y_i} - 2(m-p-1)\frac{d_1}{y_i} - 4\sum_{j>i} \frac{d_i - d_j}{y_i - y_j}\} \leq 0.$$

Noting that $y_i/(y_i - y_j) > 1$ for $j > i$ and $d_1 \geq d_2 \geq \cdots \geq d_p > 0$ we get

$$\begin{aligned}\sum_{i=1}^{p}\sum_{j>i} \frac{d_i - d_j}{y_i - y_j} &= \sum_{i=1}^{p} \frac{1}{y_i} \sum_{j>i} \frac{y_i}{y_i - y_j}(d_i - d_j) \geq \sum_{i=1}^{p} \frac{1}{y_i} \sum_{j>i} (d_i - d_j) \\ &= \sum_{i=1}^{p} \frac{(p-i)d_i - \sum_{j>i} d_j}{y_i} > 0\end{aligned}$$

since $y_i/(y_i - y_j) > 1$. It follows that

$$\Delta \leq \sum_{i=1}^{p} \frac{1}{y_i} \left\{ (n-p+3)d_i^2 - 2(m+p-2i-1)d_i + 4\sum_{j>i} d_j \right\}.$$

Denote the term inside curly bracket of the above inequality by $z_i(d_i)$. Then it suffices to show that $z_i(d_i)$ is negative when $d_i = (m+p-2i-1)/(n-p+3)$ (say d_i^0), $i = 1, 2, \ldots, p$. For fixed d_j $(j = i+1, \ldots, p)$, $z_i(d_i)$ is minimized at $d_i = d_i^0$, which follows that

$$z_i(d_i^0) < z_i(d_{i+1}^0) = z_{i+1}(d_{i+1}^0).$$

Then

$$z_i(d_i^0) < z_{i+1}(d_{i+1}^0) < \cdots < z_p(d_p^0) < 0,$$

since $d_1^0 > \cdots > d_p^0$. This completes the proof.

2.5.3. IMPROVING UPON STEIN TYPE ESTIMATORS

Bilodeau and Kariya[5] gave the estimator (included in (2.55)) which beats the Efron-Morris estimator given by (2.67) for the case $p \geq 2$. In this section we consider two classes of estimators which beat the Stein-type shrinkage estimators given by (2.66) and (2.67) for the case $p \geq 2$, respectively.

Let $\alpha(Y)$ be a real-valued function of $Y = \operatorname{diag}(y_1, y_2, \ldots, y_p)$ satisfying

$$\alpha(Y) \geq 0 \ for \ y_1 \geq y_2 \geq \cdots y_p \geq 0 \tag{2.69}$$

$$E_{B,\Sigma}|\alpha_i(Y)\sqrt{y_i}\,| < \infty \tag{2.70}$$

where $\alpha_i(Y) = \partial\alpha(Y)/\partial y_i$, $i = 1, 2, \ldots, p$.

THEOREM 2.5.4. (i) *Assume that*

$$\sum_{i=1}^{p}\left\{\frac{p+1}{2}\alpha(Y) - y_i\alpha_i(Y)\right\} \geq 0. \tag{2.71}$$

Then ,when $m > p+1$ *and* $p \geq 2$, *the estimators*

$$\begin{aligned}\hat{B}^*_{s1}(X,S) =& X\Bigg[I_p - \frac{1}{(n-p+3)\operatorname{tr}S^{-1}}\Bigg\{(m-p-1)(X'X)^{-1} \\ &+ 2I_p\Big/\left(\frac{\operatorname{tr}X'X}{p(p+1)-2} + \alpha(Y)\right)\Bigg\}\Bigg]\end{aligned}$$

beat the Stein estimator $\hat{B}_{s1}(X,S)$ *given by* (2.66) *with respect to the loss function* (2.2).

(ii) *Assume that*

$$\frac{p+1}{p-1} + 2\alpha_i(Y) \geq 0 \qquad i = 1, 2, \ldots, p \tag{2,72}$$

and that

$$\sum_{i=1}^{p}\left\{\frac{p+1}{2p}\alpha(Y) - y_i\alpha_i(Y)\right\} \geq 0. \tag{2.73}$$

Then, when $m > p+1$ *and* $p \geq 2$, *the estimators*

$$\hat{B}^*_{s2}(X,S) = X\left[I_p - \frac{1}{n+p+1}\left\{(m-p-1)(X'X)^{-1} + 2I_p\Big/\left(\frac{\operatorname{tr}X'X}{p-1} + \alpha(Y)\right)\right\}S\right]$$

beat the Stein estimator $\hat{B}_{s2}(X,S)$ *given by* (2.67) *with respect to the loss function* (2.2).

PROOF. (i) Let

$$g_1(u,\alpha) = \frac{u}{p(p+1)-2} + \alpha(Y)$$

where $u = \operatorname{tr} X'X$. Then $\hat{B}^*_{s1}$ can be written as $X[I_p + OT(Y)O'/\operatorname{tr} S^{-1}]$ where $T(Y) = \operatorname{diag}(t_1(Y), t_2(Y), \ldots, t_p(Y))$ and

$$t_i(Y) = -\frac{1}{n-p+3}\left(\frac{m-p-1}{y_i} + \frac{2}{g_1(u,\alpha)}\right), \qquad i = 1,2,\ldots,p.$$

Using (2.56) it can be seen that

$$\begin{aligned}\Delta =& \mathbf{R}((B,\Sigma),\hat{B}^*_{s1}) - \mathbf{R}((B,\Sigma),\hat{B}_{s1}) \\ =& \frac{4}{n-p+3}\mathbf{E}\left[\sum_{i=1}^{p}\frac{1}{\operatorname{tr} S^{-1}}\left\{-\frac{m}{g_1(u,\alpha)} + \frac{2y_i}{g_1^2(u,\alpha)}\frac{\partial g_1(Y)}{\partial y_i}\right\}\right. \\ &\left. + \frac{4}{(\operatorname{tr} S^{-1})^3}\operatorname{tr}(OQ(Y)O'S^{-2}) + \frac{n-p-1}{(\operatorname{tr} S^{-1})^2}\operatorname{tr}(OQ(Y)O'S^{-1})\right]\end{aligned}$$

where $Q(Y) = \operatorname{diag}(q_1(Y), q_2(Y), \ldots, q_p(Y))$ and

$$q_i(Y) = \frac{1}{n-p+3}\left(\frac{m-p-1}{g_1(u,\alpha)} + \frac{y_i}{g_1^2(u,\alpha)}\right), \qquad i = 1,2,\ldots,p.$$

As $q_i(Y)$ is nonnegative by the condition on $\alpha(Y)$, we have

$$\operatorname{tr}(OQ(Y)O'S^{-2}) \leq \operatorname{tr} Q \cdot (\operatorname{tr} S^{-1})^2$$

and

$$\operatorname{tr}(OQ(Y)O'S^{-1}) \leq \operatorname{tr} Q \cdot \operatorname{tr} S^{-1}.$$

It follows from these inequalities and some simplification that

$$\Delta \leq \frac{4}{n-p+3}\mathbf{E}\left[\frac{1}{\operatorname{tr} S^{-1}}\sum_{i=1}^{p}\left\{-\frac{p+1}{g_1(u,\alpha)} + \frac{y_i}{g_1^2(u,\alpha)} + \frac{2y_i}{g_1^2(u,\alpha)}\frac{\partial g_1(u,\alpha)}{\partial y_i}\right\}\right]. \quad (2.74)$$

Differentiating $g_1(u,\alpha)$ with respect to y_i and noting that $\sum_{i=1}^{p} y_i = u$, we can see that the term inside curly bracket in the right hand side of (2.74) is equal to

$$\sum_{i=1}^{p}\frac{-(p+1)\alpha(Y) + 2y_i\alpha_i(Y)}{g_1^2(u,\alpha)}$$

which follows that $\Delta \leq 0$ by the condition (2.71). This completes the proof of (i).

(ii) Let

$$g_2(u,\alpha) = \frac{u}{p-1} + \alpha(Y) \quad \text{and} \quad t_i(Y) = -\frac{1}{n+p+1}\left(\frac{m-p-1}{y_i} + \frac{2}{g_2(u,\alpha)}\right).$$

Similarly, using (2.57) and straightforward calculation give that

$$\begin{aligned}\Delta =&\mathbf{R}((B,\Sigma),\hat{B}^*_{s2}) - \mathbf{R}((B,\Sigma),\hat{B}_{s2})\\ =&\frac{4}{n+p+1}\mathbf{E}\left[\operatorname{tr}(OQ^*(Y)O'S) - \frac{p+1}{g_2(u,\alpha)}\operatorname{tr} S\right]\end{aligned}$$

where $Q^*(Y) = sdiag(q^*_1(Y), q^*_2(Y), \ldots, q^*_p(Y))$ and

$$q^*_i(Y) = \frac{y_i}{g_2^2(u,\alpha)}\left(\frac{p+1}{p-1} + 2\alpha_i(Y)\right), \qquad i = 1,2,\ldots,p.$$

As $q^*_i(Y)$ is nonnegative by the condition (2.72), we have $\operatorname{tr}(OQ^*(Y)O'S) \leq \operatorname{tr} Q^*(Y) \cdot \operatorname{tr} S$. Using this inequality and noting that $u = \sum_{i=1}^p y_i$, we can see that

$$\Delta \leq \frac{4}{n+p+1}\mathbf{E}\left[\operatorname{tr} S \sum_{i=1}^{p}\left\{\frac{2y_i\alpha_i(Y)}{g_2^2(u,\alpha)} - \frac{(p+1)\alpha(Y)}{pg_2^2(u,\alpha)}\right\}\right].$$

This completes the proof from (2.73).

EXAMPLE 2.5.3. (i) Let $\alpha(Y) = \operatorname{tr} X'X/\{p(p+1)-2\}$. Using Theorem 2.5.4 we get that the estimator

$$\hat{B}^*_{s1}(X,S) = X\left(I_p - \frac{m-p-1}{(n-p+3)\operatorname{tr} S^{-1}}(X'X)^{-1} - \frac{p(p+1)-2}{(n-p+3)\operatorname{tr} S^{-1}\operatorname{tr} X'X}\right)$$

beats the estimator $\hat{B}_{s1}(X,S)$ given by (2.66).

(ii) Let $\alpha(Y) = \operatorname{tr} X'X/(p-1)$. Then the estimator $\hat{B}^*_{s2}$ becomes

$$\hat{B}^*_{s2}(X,S) = X\left(I_p - \frac{m-p-1}{n+p+1}(X'X)^{-1}S - \frac{p-1}{n+p+1}\frac{S}{\operatorname{tr} X'X}\right) \qquad (2.75)$$

which beats the estimator $\hat{B}_{s2}(X,S)$ given by (2.67). The superiority of $\hat{B}^*_{s2}$ over $\hat{B}_{s2}$ is proved by Bilodeau and Kariya[5] while it is also seen from Theorem 2.5.4.

EXAMPLE 2.5.4. Let $\alpha(Y) = \sum_{i<j} c_{ij}/(y_i - y_j)$ where c_{ij} is a nonnegative constant. Using the conclusion of Theorem 2.5.4 we get that the estimator

$$
\begin{aligned}
\hat{B}^*_{s1}(X,S) = & X\left(I_p - \frac{m-p-1}{(n-p+3)\operatorname{tr} S^{-1}}(X'X)^{-1}\right) \\
& - 2X \Big/ \left\{(n-p+3)\operatorname{tr} S^{-1}\left(\frac{\operatorname{tr} X'X}{p(p+1)-2} + \sum_{i<j} \frac{c_{ij}}{y_i - y_j}\right)\right\}
\end{aligned}
$$

beats the estimator $\hat{B}_{s1}(X,S)$.

But the condition (2.72) is crucial so that we can't obtain the estimator of the second form in Theorem 2.5.4 which beats the estimator $\hat{B}_{s2}(X,S)$ for this choice.

2.6. CONCLUDING REMARKS ON THE PROPOSED ESTIMATORS

For the case $m > p+1$ Efron and Morris[15] derived the estimator given by (2.67) from the empirical Bayes argument. Later, using Lemma 2.1.3, Bilodeau and Kariya[5] obtained the estimator given by (2.75) for the same case. Since these two estimators take into account the matrix structure of mean B, they are not trivial extension of the Stein estimator of the normal mean vector. However, their method involves in the direct calculation and the eigenstructure is used implicitly. The representations of the unbiased estimate of risk in Theorems 2.3.1 and 2.5.1 shed new light on this structure and suggest us to utlize the imformation of the ordered eigenvalues of $X'XS^{-1}$ or $X'X$. These unbiased risk estimators facilitate a systematic seacrch for alternatives such as multivariate extensions of Baranchik estimators, adjusted estimators, and the variational form of Bayes estimators for the case $p > m+1$ as well as $m > p+1$. However, from rather complicated nature of the estimators proposed, it appears that the analytic comparison among these estimators is not possible at this point. But it is our belief that by taking into account the terms $\sum_{l>k}(f_k^2 h_k^2 - f_l^2 h_l^2)/(f_k - f_l)$ and $\sum_{l>k}(f_k h_k - f_l h_l)/(f_k - f_l)$ of the unbiased estimate of the risk in Theorem 2.3.1, one should be able to obtain superior alternatives which have substantial improvement in risk over the usual estimator X. Now work is in progress along this direction.

CHAPTER 3

ESTIMATING EIGENVALUES IN THE MULTIVARIATE F-DISTRIBUTION

Suppose that S_1 and S_2 are independent Wishart matrices with $S_i \sim W_p(k_i, \Sigma_i)$ where Σ_i is positive definite matrix and k_i is degrees of freedom, $i = 1, 2$. The eigenvalues $\delta_1, \ldots, \delta_p$ ($\delta_1 \geq \cdots \geq \delta_p > 0$) of $\Sigma_1 \Sigma_2^{-1}$ are important, for example, in the problem of testing $\Sigma_1 = \Sigma_2$ against $\Sigma_1 \neq \Sigma_2$ as the power function of any invariant test statistics under a natural group of transformations depends only on $\delta_1, \ldots, \delta_p$. The literature includes DasGupta[9], Dey[10] , Muirhead and Verathaworn[42] and Leung and Muirhead[33]. Ideally, a decision theoretic approach would specify a loss function in terms of $\delta_1, \ldots, \delta_p$ and risk calculation would be done with respect to the expectation of the joint distribution of the eigenvalues $l_1, \ldots, l_p$ ($l_1 > \cdots > l_p > 0$) of $S_1 S_2^{-1}$. However, this approach seems unfeasible mainly due to the complexity of the distribution of the ordered eigenvalues $l_1, \ldots, l_p$. Instead, following approach by Muirhead and Verathaworn[42], we construct a $p \times p$ positive definite random matrix U with the scale matrix (whose eigenvalues are equal to $\delta_1, \ldots, \delta_p$) and the degrees of freedom k_1 and k_2 as a function of S_1 and S_2 such that the eigenvalues of a $p \times p$ random matrix U have the same distribution as those of $S_1 S_2^{-1}$ and this distribution depends only on $\delta_1, \ldots, \delta_p$. As we restrict our attention to the class of orthogonally invariant estimators $\hat{\Delta}(U)$ of Δ, the eigenvalues of $\hat{\Delta}(U)$ may be interpreted as estimates of $\delta_1, \ldots, \delta_p$, and then it is natural to expect that the eigenvalues of 'good' estimate $\hat{\Delta}(U)$ of Δ will perform well as estimates of $\delta_1, \ldots, \delta_p$.

Let U be a $p \times p$ positive definite random matrix having density function

$$C(\det \Delta)^{-k_1/2}(\det U)^{(k_1-p-1)/2}\det(I_p + \Delta^{-1}U)^{-k/2}, \tag{3.1}$$

where $k = k_1 + k_2$,

$$C = \Gamma_p\left(\frac{1}{2}k\right)/\left\{\Gamma_p\left(\frac{1}{2}k_1\right)\Gamma_p\left(\frac{1}{2}k_2\right)\right\}, \qquad \Gamma_p(a) = \pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma\left(a - \frac{1}{2}(i-1)\right),$$

$k_i > p+1$, $i = 1, 2$, and Δ is a positive definite parameter matrix. Let us denote this distribution by $F_p(k_1, k_2; \Delta)$. This distribution generalizes usual F-distribution in much the same way that the Wishart distribution does the χ^2-distribution. Some of properties of the multivariate F-distribution are similar to those of the Wishart distribution, which are discussed in several papers such as Dawid[8], Khatri[26], Konno[27], Olkin and Rubin[43], Tan[52], Mitra[39], Perlman[45], and De Waal[53].

In Section 3.2, we first derive the first and second order moments for this distribution, which are useful for the statistical inference on the parameters of the multivariate F-distribution.

Next, we consider the problem of estimating the eigenvalues $\delta_1, \ldots, \delta_p$ in terms of the loss functions

$$\mathbf{L}_1(\Delta, \hat{\Delta}(U)) = \mathrm{tr}\,(\Delta^{-1}\hat{\Delta}(U)) - \log\det(\Delta^{-1}\hat{\Delta}(U)) - p, \tag{3.2}$$

and

$$\mathbf{L}_2(\Delta, \hat{\Delta}(U)) = \mathrm{tr}\,(\Delta^{-1}\hat{\Delta}(U) - I_p)^2. \tag{3.3}$$

These loss functions are originally proposed for the estimation problem of the eigenvalues of the normal covariance matrix. Recently, Bilodeau[4] proposed a loss function

$$\mathbf{L}_3(\Delta, \hat{\Delta}(U)) = \mathrm{tr}\,\{(\Delta + U)^{-1}(\hat{\Delta}(U) + U)\} - \log\det\{(\Delta + U)^{-1}(\hat{\Delta}(U) + U)\} - p. \tag{3.4}$$

The corresponding risk function is denoted by $\mathbf{R}_i(\Delta, \hat{\Delta}(U)) = \mathbf{E}\,[\mathbf{L}_i(\Delta, \hat{\Delta}(U))]$ $(i = 1, 2, 3)$ taking expectation with respect to the distribution given by (3.1). Following an approach similar to that of Haff[22] in the problem of estimating the normal covariance matrix, Muirhead and Verathaworn[42] developed an approximation to the Bayes rule under the loss function (3.2). Later, using an approximation to the risk function $\mathbf{R}_1$ (appeared in Muirhead and Verathaworn[42]), Gupta and Krishnamoorthy[16] and Dey[11] proposed new estimators, which are analogous to the Stein's adjusted minimax estimator and Dey and Srinivasan's estimator of the normal covariance matrix, respectively. However, it has not been established that the estimators proposed have a frequentist risk uniformly smaller than the best multiple of U under the loss function (3.2) which is just the unbiased estimator

$$\hat{\Delta}_{UN} = \frac{k_2 - p - 1}{k_1} U. \tag{3.5}$$

On the other hand, Bilodeau[4] obtained the improved estimators under the loss function (3.4), which beat the best multiple of U under the same loss function.

Section 3.1 deals with preliminary lemmas concerning the action of a matrix of differential operator and identity for $\mathbf{E}\,[\,\mathrm{tr}\,(\Delta^{-1}\hat{\Delta}(U))]$. After deriving the exact moments of U in Section 3.2, several new estimators are proposed in Sections 3.3 and 3.4. An improved estimator which modifies all eigenvalues of the estimator $\hat{\Delta}_{UN}(U)$ in the same direction similar to that of Haff[18] for the eigenvalues of the normal covariance matrix is given under the loss function (3.2). Next, it is shown that the estimator similar to that of Perron[46] for the eigenvalues of the normal covariance matrix is better than $\hat{\Delta}_{UN}(U)$ under the loss function (3.2). Finally, for the case where $p = 2$, the estimators of Gupta and Krishnamoorthy[16] and of Perron-type, are minimax under the loss function (3.2). In Section 3.4, it is shown that Haff-type estimator beats the best multiple of U under the loss function (3.3).

3.1. PRELIMINARIES

In this section we state calculus lemmas and the F-identity which is similar to the Wishart-identity in Lemma 2.1.2. For notation, let D be a $p \times p$ matrix of differential operator whose (i,j) element is given by $(1/2)(1+\delta_{ij})\partial/\partial U_{ij}$ for $U = (U_{ij})$ and a Kronecker's delta δ_{ij}. The following lemma describes the action of the operator D on matrix products of U, Δ, and a $p \times p$ matrix Q.

LEMMA 3.1.1. *Let Q be a $p \times p$ matrix. Then we have*

$$\text{(i)} \qquad \operatorname{tr} DQU = \frac{p+1}{2} \operatorname{tr} Q,$$

$$\text{(ii)} \qquad \operatorname{tr} DUQU = (p+1) \operatorname{tr}(QU),$$

$$\text{(iii)} \qquad (QD)'U = \frac{1}{2}\{ \operatorname{tr}(Q) I_p + Q\} \qquad \text{(Haff[22])},$$

$$\text{(iv)} \qquad DU\Delta^{-1}U = \frac{1}{2}(\operatorname{tr} \Delta^{-1}U) I_p + \frac{p+2}{2}\Delta^{-1}U.$$

PROOF. (i) It follows from Haff[18].

(ii) Put $Q = e_i e_j'$ where e_i is the i th unit column vector. The direct calculation shows that

$$\begin{aligned}
\operatorname{tr} DU e_i e_j' U &= \sum_{k=1}^{p} \left\{ \frac{\partial}{\partial U_{kk}} U_{ki} U_{jk} + \frac{1}{2} \sum_{k \neq l} \frac{\partial}{\partial U_{kl}} U_{ki} U_{jl} \right\} \\
&= 2U_{ij} + \sum_{k \neq i} U_{ij} \\
&= (p+1) e_j' U e_i,
\end{aligned}$$

which completes the proof of (ii).

(iii) Direct calculation of the (i,j) element of the right hand side yields that of the left hand side. See also Haff[21] for the proof.

(iv) Using Lemma 2.1.5 (replaced S and D_S by U and D, respectively) and noting that U and Δ are symmetric, we have

$$DU\Delta^{-1}U = (DU\Delta^{-1})U + (\Delta^{-1}UD)'U.$$

From $DU = ((p+1)/2)I_p$ and the third part of this lemma, it follows that

$$(DU\Delta^{-1})'U = \frac{p+1}{2}\Delta^{-1}U$$

and

$$(\Delta^{-1}UD)'U = \frac{1}{2}\left(\operatorname{tr}(\Delta^{-1}U)I_p + \Delta^{-1}U\right).$$

Combining these equations we obtain the desired result.

Next we state the F-identity due to Muirhead and Verathaworn[42]. For notation, let $Q_{(r)} = (\delta_{ij}q_{ij} + r(1-\delta_{ij})q_{ij})$ where $Q = (q_{ij})$ and r is a constant. Furthermore, recall that $U = (U_{ij})$ and $D = (d_{ij})$ with $d_{ij} = (1/2)(1+\delta_{ij})(\partial/\partial U_{ij})$.

LEMMA 3.1.2. *Let U follow the $F_p(k_1, k_2; \Delta)$ distribution defined by* (3.1). *For a suitable choice of a $p \times p$ matrix-valued function $V(U, \Delta)$ and a scalar function $g(U)$, we have*

$$\begin{aligned} k\mathbf{E}\left[g(U)\operatorname{tr}(\Delta+U)^{-1}V\right] = \ & \mathbf{E}\Big[2g(U)\operatorname{tr}(DV) + 2\operatorname{tr}\left(\frac{\partial g(U)}{\partial U_{ij}}V_{\left(\frac{1}{2}\right)}\right) \\ & + (k_1 - p - 1)g(U)\operatorname{tr}(U^{-1}V)\Big], \end{aligned} \tag{3.6}$$

where $\partial g(U)/\partial U = (\partial g(U)/\partial U_{ij})$ and $k = k_1 + k_2$.

PROOF. See Muirhead and Verathaworn[42] for the proof.

3.2. MOMENTS OF THE MULTIVARIATE F-DISTRIBUTION

Using Lemma 3.1.2 we shall compute the first and second order moments of the random matrix U.

THEOREM 3.2.1. *Let U follow the $F_p(k_1, k_2; \Delta)$ distribution and put $\Delta = (\Delta_{ij})$, then*

$$\text{(i)} \qquad \mathbf{E}\,[U_{ij}] = \frac{k_1}{k_2 - p - 1}\Delta_{ij} \qquad \text{if } k_2 - p - 1 > 0,$$

$$\begin{aligned}\text{(ii)} \qquad \mathbf{E}\,[U_{ij}U_{k\ell}] =& \frac{k_1}{(k_2-p)(k_2-p-1)(k_2-p-3)}[\{k_1(k_2-p-2)+2\}\Delta_{ij}\Delta_{k\ell} \\ &+ (k-p-1)(\Delta_{j\ell}\Delta_{ik} + \Delta_{kj}\Delta_{i\ell})] \qquad \text{if } k_2 - p - 3 > 0,\end{aligned}$$

where $k = k_1 + k_2$.

PROOF. (i) In the equation (3.6), set $g(U) = 1$ and $V = (\Delta + U)e_j e_i' U$, where e_i is the ith unit column vector. If $k_2 - p - 1 > 0$, the expectation of each term in (3.6) exists. Use (i) and (ii) of Lemma 3.1.1, then we obtain the desired result.

(ii) Set $g(U) = U_{ij}$ and $V = (\Delta + U)e_\ell e_k' U$ in (3.6). Note that the left hand side in (3.6) becomes $k\mathbf{E}\,[U_{ij}U_{k\ell}]$. Using (i) and (ii) of Lemma 3.1.1, the first term of the right hand side in (3.6) provides

$$\mathbf{E}\,[g(U)\,\mathrm{tr}\,(DV)] = \frac{p+1}{2}\Delta_{k\ell}\mathbf{E}\,[U_{ij}] + (p+1)\mathbf{E}\,[U_{ij}U_{k\ell}].$$

Noting that $\partial g(U)/\partial U = (e_i e_j' + e_j e_i')/(1 + \delta_{ij})$, the second and third terms of the right hand side become

$$\mathbf{E}\left[\mathrm{tr}\left(\frac{\partial g(U)}{\partial U}V_{(\frac{1}{2})}\right)\right] = \frac{1}{2}\mathbf{E}\,\left[U_{ki}\Delta_{j\ell} + U_{kj}\Delta_{i\ell} + U_{j\ell}U_{ki} + U_{i\ell}U_{kj}\right],$$

$$\mathbf{E}\,[g(U)\,\mathrm{tr}\,(U^{-1}V)] = \mathbf{E}\,[U_{ij}\Delta_{k\ell} + U_{ij}U_{k\ell}],$$

respectively. Combining these equations and using the first part of this theorem lead to

$$\begin{aligned}(k_2 - p - 1)&\mathbf{E}\,[U_{ij}U_{k\ell}] - \mathbf{E}\,[U_{j\ell}U_{ki}] - \mathbf{E}\,[U_{i\ell}U_{kj}] \\ &= \frac{k_1}{k_2-p-1}\{k_1\Delta_{ij}\Delta_{k\ell} + \Delta_{j\ell}\Delta_{ki} + \Delta_{i\ell}\Delta_{kj}\}.\end{aligned} \tag{3.7}$$

In the similar way, from $g(U) = U_{ik}$ and $V = (\Delta + U)e_l e_j' U$, we obtain

$$
\begin{aligned}
(k_2 - p - 1)\mathbf{E}[U_{ik}U_{jl}] - \mathbf{E}[U_{ij}U_{kl}] - \mathbf{E}[U_{kj}U_{il}]& \\
= \frac{k_1}{k_2 - p - 1}\{k_1\Delta_{jl}\Delta_{ik} + \Delta_{kl}\Delta_{ji} + \Delta_{il}\Delta_{kj}\},&
\end{aligned}
\tag{3.8}
$$

and, from $g(U) = U_{il}$ and $V = (\Delta + U)e_j e_k' U$, we get

$$
\begin{aligned}
(k_2 - p - 1)\mathbf{E}[U_{il}U_{kj}] - \mathbf{E}[U_{ki}U_{jl}] - \mathbf{E}[U_{kl}U_{ij}]& \\
= \frac{k_1}{k_2 - p - 1}\{k_1\Delta_{kj}\Delta_{il} + \Delta_{jl}\Delta_{ki} + \Delta_{ij}\Delta_{kl}\}.&
\end{aligned}
\tag{3.9}
$$

Thus $\mathbf{E}[U_{ij}U_{kl}]$ is determined by the linear equations of (3.7), (3.8), and (3.9), which completes the proof.

COROLLARY 3.2.1. *If U follows the $F_p(k_1, k_2; \Delta)$ distribution and $k_2 - p - 3 > 0$, then*

(i) $$\mathbf{Cov}(U_{ij}, U_{kl}) = \frac{k_1(k-p-1)}{(k_2-p)(k_2-p-1)(k_2-p-3)}\left\{\frac{2}{k_2-p-2}\Delta_{ij}\Delta_{kl} + \Delta_{jl}\Delta_{ik} + \Delta_{kj}\Delta_{il}\right\},$$

(ii) $$\mathbf{E}[UQU] = \frac{k_1}{(k_2-p)(k_2-p-1)(k_2-p-3)}[\{k_1(k_2-p-2)+2\}\Delta Q\Delta + (k-p-1)\{(\Delta Q\Delta)' + \operatorname{tr}(\Delta Q)\Delta\}],$$

for $k = k_1 + k_2$ and a $p \times p$ matrix Q.

PROOF. From Theorem 3.2.1, direct calculation leads to the result.

REMARK 3.2.1. The results of Theorem 3.2.1 and Corollary 3.2.1 include the moments of the Wishart matrix in Haff[18] as a special case. Put $U^* = k_2 U$ in Corollary 3.2.1 and assume that Q is symmetric, then we get

$$\lim_{k_2 \to \infty} \mathbf{E}[U^*QU^*] = k_1(k_1+1)\Delta Q\Delta + k_1 \operatorname{tr}(\Delta Q)\Delta,$$

which is equal to $\mathbf{E}_W[WQW]$, W having the Wishart distribution $W_p(k_1, \Delta)$, as U^* converges to W weakly.

REMARK 3.2.2. We derive the second order moments of the matrix U by using the identity (3.6), however Professor Sinha pointed out that combining the Wishart moments due to Haff[18] with the inverted Wishart prior on a covariance matrix also gives the same results without using the identity (3.6). Namely, assuming that

$$W \sim W_p(k_1, \Sigma) \qquad \text{and} \qquad \Sigma^{-1} \sim W_p(k_2, \Delta^{-1}),$$

the joint density of W and Σ^{-1} is

$$\frac{2^{-pk/2}(\det \Delta)^{k_2/2}(\det W)^{(k_1-p-1)/2}}{\Gamma_p(k_1/2)\Gamma_p(k_2/2)(\det \Sigma)^{(k-p-1)/2}} \mathrm{etr}\left\{-\frac{1}{2}\Sigma^{-1}(W+\Delta)\right\}(dW)(d\Sigma^{-1}), \tag{3.10}$$

where $k = k_1 + k_2$. By making a transformation $\Phi = (W+\Delta)^{1/2}\Sigma^{-1}(W+\Delta)^{1/2}$, we have the joint density of W and Φ

$$\frac{2^{-pk/2}(\det \Delta)^{k_2/2}(\det W)^{(k_1-p-1)/2}}{\Gamma_p(k_1/2)\Gamma_p(k_2/2)\det (W+\Delta)^{k/2}}(\det \Phi)^{(k-p-1)/2}\mathrm{etr}\left(-\frac{1}{2}\Phi\right)(dW)(d\Phi). \tag{3.11}$$

Furthermore, by integrating out (3.11) with respect to Φ, it is seen that the marginal density of W becomes the $F_p(k_1, k_2; \Delta)$ distribution. From (3.10), it follows that the second order moments of the multivariate F-distribution can be calculated by

$$\int\int W_{ij}W_{kl}\frac{2^{-pk/2}(\det \Delta)^{k_2/2}(\det W)^{(k_1-p-1)/2}}{\Gamma_p(k_1/2)\Gamma_p(k_2/2)(\det \Sigma)^{(k-p-1)/2}} \mathrm{etr}\left\{-\frac{1}{2}\Sigma^{-1}(W+\Delta)\right\}(dW)(d\Sigma^{-1}), \tag{3.12}$$

where $W = (W_{ij})$. First, integrating (3.12) with respect to W (having the Wishart distribution $W_p(k_1, \Sigma)$) gives

$$\begin{aligned}\int k_1(\Sigma_{ik}\Sigma_{jl} + \Sigma_{il}\Sigma_{jk} + k_1\Sigma_{ij}\Sigma_{kl})\frac{2^{-(pk_2/2)}(\det \Delta)^{(k_2/2)}}{\Gamma_p(k_2/2)}&(\det \Sigma^{-1})^{(k_2-p-1)/2} \\ \times\, \mathrm{etr}\left(-\frac{1}{2}\Delta\Sigma^{-1}\right)&(d\Sigma^{-1}).\end{aligned} \tag{3.13}$$

From Haff[18], we get

$$\mathrm{E}\left[\Sigma_{ij}\Sigma_{kl}\right] = \frac{1}{(k_2-p)(k_2-p-1)(k_2-p-3)}\{(k_2-p-2)\Delta_{ij}\Delta_{kl} + \Delta_{ik}\Delta_{jl} + \Delta_{il}\Delta_{kj}\},$$

where Σ^{-1} follows the $W_p(k_2, \Delta^{-1})$ distribution. By exchanging k with j or j with l, similar formula for $\mathrm{E}\left[\Sigma_{ik}\Sigma_{jl}\right]$ or $\mathrm{E}\left[\Sigma_{il}\Sigma_{kj}\right]$ is obtained. This gives (ii) of Theorem 3.2.1.

If U follows the $F_p(k_1, k_2; \Delta)$ distribution, then U^{-1} follows the $F_p(k_2, k_1; \Delta^{-1})$ distribution. Immediately, Theorem 3.2.1 and Corollary 3.2.1 give the following inverse moments.

COROLLARY 3.2.2. *If U follows the $F_p(k_1, k_2; \Delta)$ distribution and $k_1 - p - 3 > 0$, then*

$$\text{(i)} \qquad \mathbf{E}[U^{-1}] = \frac{k_2}{k_1 - p - 1}\Delta^{-1},$$

$$\begin{aligned}\text{(ii)} \qquad \mathbf{E}[U^{ij}U^{kl}] =& \frac{k_2}{(k_1 - p)(k_1 - p - 1)(k_1 - p - 3)}[\{k_2(k_1 - p - 2) + 2\}\Delta^{ij}\Delta^{kl} \\ &+ (k - p - 1)(\Delta^{jl}\Delta^{ik} + \Delta^{kj}\Delta^{il})],\end{aligned}$$

$$\begin{aligned}\text{(iii)} \ \mathbf{Cov}(U^{ij}, U^{kl}) =& \frac{k_2(k - p - 1)}{(k_1 - p)(k_1 - p - 1)(k_1 - p - 3)}\left[\frac{2}{k_1 - p - 1}\Delta^{ij}\Delta^{kl}\right. \\ &\left.+ \Delta^{jl}\Delta^{ik} + \Delta^{kj}\Delta^{il}\right],\end{aligned}$$

$$\begin{aligned}\text{(iv)} \ \mathbf{E}[U^{-1}QU^{-1}] =& \frac{k_2}{(k_1 - p)(k_1 - p - 1)(k_1 - p - 3)}[\{k_2(k_1 - p - 2) + 2\}\Delta^{-1}Q\Delta^{-1} \\ &+ (k - p - 1)\{(\Delta^{-1}Q\Delta^{-1})' + \operatorname{tr}(\Delta^{-1}Q)\Delta^{-1}\}],\end{aligned}$$

where Q is any $p \times p$ matrix, $U^{-1} = (U^{ij})$, and $\Delta^{-1} = (\Delta^{ij})$.

3.3. SOME IDENTITIES

Recall that our goal is to estimate the eigenvalues of Δ using the eigenvalues of the random matrix U. Hence we shall restrict our class of estimators of Δ to the orthogonally invariant estimators of the form

$$\hat{\Delta}(U) = H\varphi(L)H', \tag{3.14}$$

where H is a $p \times p$ orthogonal matrix such that $U = HLH'$ with $L = \operatorname{diag}(l_1, \ldots, l_p)$ and $l_1 > \cdots > l_p > 0$, and $\varphi = \operatorname{diag}(\varphi_1(L), \cdots, \varphi_p(L))$.

In an attempt to obtain the unbiased estimate of the risk of estimators (3.14) with respect to the loss function (3.2), Muirhead and Verathaworn[42] applied Lemma 3.1.2 (being $V = (\Delta + U)^{-1}\Delta^{-1}\hat{\Delta}$ and $g(U) = 1$ in (3.6)) and derived that

$$\begin{aligned}\mathbf{E}[\operatorname{tr}(\Delta^{-1}\hat{\Delta})] =& \mathbf{E}\left[\frac{k_1 - p - 1}{k_2}\operatorname{tr}(U^{-1}\hat{\Delta}) + \frac{2}{k_2}\operatorname{tr}(D\hat{\Delta})\right. \\ &\left.+ \frac{2}{k_2}\operatorname{tr}(\Delta^{-1}UD\hat{\Delta})\right].\end{aligned} \tag{3.15}$$

Unlike the problem of estimating the normal covariance, it is impossible to get rid of unknown parameter Δ completely in (3.15). At this point they used a heuristic approximation to the right hand side in (3.15), namely they replaced U in the right hand side by its expectation $(k_1/(k_2-p-1))\Delta$, which gives the approximation

$$\mathbf{E}[\operatorname{tr}(\Delta^{-1}\hat{\Delta})] \approx \mathbf{E}\left[\frac{2(k-p-1)}{k_2(k_2-p-1)}\operatorname{tr}(D\hat{\Delta}) + \frac{k_1-p-1}{k_2}\operatorname{tr}(U^{-1}\hat{\Delta})\right].$$

From this, they obtained the unbiased estimate of the approximate risk (omitting constants) given by

$$\hat{R}^* = \frac{2(k-p-1)}{k_2(k_2-p-1)}\operatorname{tr}(D\hat{\Delta}) + \frac{k_1-p-1}{k_2}\operatorname{tr}(U^{-1}\hat{\Delta}) - \log\det\hat{\Delta}. \tag{3.16}$$

Gupta and Krishnamoorthy[16] and Dey[11] also employed this approximate risk to look for new estimators. However, this approximate risk is not helpful to find orthogonally invariant minimax estimators. To this end, we apply Lemma 3.1.2 recursively and obtain more precise formula for $\mathbf{E}[\operatorname{tr}(\Delta^{-1}\hat{\Delta})]$ where $\hat{\Delta}$ belongs to (3.14).

THEOREM 3.3.1. *Assume that the third order moment of U exists and that $\hat{\Delta}(U)$ satisfies* (3.14). *Then*

$$\begin{aligned}\mathbf{E}[\operatorname{tr}(\Delta^{-1}\hat{\Delta})] = \mathbf{E}\Bigg[&\frac{k_1-p-1}{k_2}\operatorname{tr}(U^{-1}\hat{\Delta}) + \frac{2(k-1)}{k_2(k_2-1)}\operatorname{tr}(D\hat{\Delta}) \\ &+ \frac{2}{k_2(k_2-1)}\operatorname{tr}(\Delta^{-1}U)\operatorname{tr}(D\hat{\Delta}) + M_1(\hat{\Delta}) + M_2(\hat{\Delta})\Bigg],\end{aligned} \tag{3.17}$$

where D is $p\times p$ differentiation operator matrix whose elements are given by $(1/2)(1+\delta_{ij})(\partial/\partial U_{ij})$,

$$M_1(\hat{\Delta}) = \frac{4}{k_2(k_2-1)(k_2-2)}\left\{(k-1)\operatorname{tr}(UD^2\hat{\Delta}) + \operatorname{tr}(U)\operatorname{tr}(D^2\hat{\Delta}) + 2\operatorname{tr}(U^2D^3\hat{\Delta})\right\}$$

and

$$\begin{aligned}M_2(\hat{\Delta}) = \frac{4}{k_2(k_2-1)(k_2-2)}\Big\{&\operatorname{tr}(\Delta^{-1}U)\operatorname{tr}(UD^2\hat{\Delta}) + \operatorname{tr}(\Delta^{-1}U^2)\operatorname{tr}(D^2\hat{\Delta}) \\ &+ 2\operatorname{tr}(U^2\Delta^{-1}UD^3\hat{\Delta})\Big\}.\end{aligned}$$

PROOF. We will evaluate $\mathbf{E}[\operatorname{tr}(\Delta^{-1}UD\hat{\Delta})]$ in (3.15). Putting $V=(\Delta+U)\Delta^{-1}UD\hat{\Delta}$ in (3,6), we obtain

$$\begin{aligned}k\mathbf{E}[\operatorname{tr}(\Delta^{-1}UD\hat{\Delta})] = &\mathbf{E}[2\operatorname{tr}(DU\Delta^{-1}UD\hat{\Delta}) + 2\operatorname{tr}(DUD\hat{\Delta}) \\ &+ (k_1-p-1)\operatorname{tr}(D\hat{\Delta} + \Delta^{-1}UD\hat{\Delta})].\end{aligned} \tag{3.18}$$

From Lemma 2.1.5, the first term of the right hand side in (3.18) becomes

$$\operatorname{tr}\{(DU\Delta^{-1}U)D\hat{\Delta}\} + \operatorname{tr}[\{(U\Delta^{-1}U)'D'\}'D\hat{\Delta}].$$

Applying (iv) of Lemma 3.1.1 to the first term above and using the fact that U, Δ, D, and $D\hat{\Delta}$ are symmetric, we have

$$\begin{aligned}\operatorname{tr}(DU\Delta^{-1}UD\hat{\Delta}) =& \frac{p+2}{2}\operatorname{tr}(\Delta^{-1}UD\hat{\Delta}) + \frac{1}{2}\operatorname{tr}(\Delta^{-1}U)\operatorname{tr}(D\hat{\Delta}) \\ &+ \operatorname{tr}(U\Delta^{-1}UD^2\hat{\Delta}).\end{aligned}$$

Similarly we get

$$\operatorname{tr}(DUD\hat{\Delta}) = \frac{p+1}{2}\operatorname{tr}(D\hat{\Delta}) + \operatorname{tr}(UD^2\hat{\Delta}).$$

Substituting these two equations in (3.18) and some simplification give

$$\begin{aligned}\mathbf{E}[\operatorname{tr}(\Delta^{-1}UD\hat{\Delta})] =& \frac{1}{k_2-1}\mathbf{E}[k_1\operatorname{tr}(D\hat{\Delta}) + 2\operatorname{tr}(UD^2\hat{\Delta}) \\ &+ \operatorname{tr}(\Delta^{-1}U)\operatorname{tr}(D\hat{\Delta}) + 2\operatorname{tr}(U\Delta^{-1}UD^2\hat{\Delta})].\end{aligned} \tag{3.19}$$

Furthermore, an application of Lemma 3.1.2 (being $V = (\Delta+U)\Delta^{-1}U(D^2\hat{\Delta})U$ in (3.6)) to the last term of the right hand side in (3.19) yields

$$\begin{aligned}k\mathbf{E}[\operatorname{tr}(U\Delta^{-1}UD^2\hat{\Delta})] =& \mathbf{E}[2\operatorname{tr}\{DU\Delta^{-1}U(D^2\hat{\Delta})U\} + 2\operatorname{tr}\{DU(D^2\hat{\Delta})U\} \\ &+ (k_1-p-1)\operatorname{tr}(UD^2\hat{\Delta} + U\Delta^{-1}UD^2\hat{\Delta})].\end{aligned} \tag{3.20}$$

Use Lemma 2.1.5 and note that $(D^2\hat{\Delta})U$ and $U\Delta^{-1}U$ are symmetric, the first term of the right hand side in (3.20) becomes

$$\operatorname{tr}\{(DU\Delta^{-1}U)(D^2\hat{\Delta})U\} + \operatorname{tr}[U\Delta^{-1}UD\{(D^2\hat{\Delta})U\}]. \tag{3.21}$$

From Lemma 2.1.5 and (iii) of Lemma 3.1.1., we get

$$D[(D^2\hat{\Delta})U] = (D^3\hat{\Delta})U + \frac{1}{2}(\operatorname{tr}D^2\hat{\Delta})I_p + \frac{1}{2}D^2\hat{\Delta}.$$

Applying (iv) of Lemma 3.1.1 to the first term of (3.21) and putting above equation in the second term, some simplification leads to

$$\begin{aligned}\operatorname{tr}\{DU\Delta^{-1}U(D^2\hat{\Delta})U\} =& \frac{p+3}{2}\operatorname{tr}(U\Delta^{-1}UD^2\hat{\Delta}) + \frac{1}{2}\operatorname{tr}(\Delta^{-1}U)\operatorname{tr}(UD^2\hat{\Delta}) \\ &+ \frac{1}{2}\operatorname{tr}(\Delta^{-1}U^2)\operatorname{tr}(D^2\hat{\Delta}) + \operatorname{tr}(U^2\Delta^{-1}UD^3\hat{\Delta}).\end{aligned}$$

Similarly we get

$$\operatorname{tr}\{DU(D^2\hat{\Delta})U\} = \frac{p+2}{2}\operatorname{tr}(UD^2\hat{\Delta}) + \frac{1}{2}\operatorname{tr}(U)\operatorname{tr}(D^2\hat{\Delta}) + \operatorname{tr}(U^2D^3\hat{\Delta}).$$

Putting these two equations into (3.20), we get

$$\begin{aligned}\mathbf{E}[\operatorname{tr}(U\Delta^{-1}UD^2\hat{\Delta})] = &\frac{1}{k_2-2}\mathbf{E}[(k_1+1)\operatorname{tr}(UD^2\hat{\Delta}) + \operatorname{tr}(U)\operatorname{tr}(D^2\hat{\Delta}) \\ &+ 2\operatorname{tr}(U^2D^3\hat{\Delta}) + \operatorname{tr}(\Delta^{-1}U)\operatorname{tr}(UD^2\hat{\Delta}) \\ &+ \operatorname{tr}(\Delta^{-1}U^2)\operatorname{tr}(D^2\hat{\Delta}) + 2\operatorname{tr}(U^2\Delta^{-1}UD^3\hat{\Delta})].\end{aligned} \tag{3.22}$$

From (3.15), (3.19), and (3.22), we finally get (3.17), which completes the proof.

Note that the right hand side of (3.17) in Theorem 3.3.1 contains the differentiation D up to the third degree. The following corollary is obtained from (3.15) and (3.19), which contains the diffferentiation D up to the second degree.

COROLLARY 3.3.1. *Assume that the second order moment of U exists and that $\hat{\Delta}(U)$ satisfies* (3.14). *Then*

$$\begin{aligned}\mathbf{E}[\operatorname{tr}(\Delta^{-1}\hat{\Delta})] = \mathbf{E}\Bigg[&\frac{k_1-p-1}{k_2}\operatorname{tr}(U^{-1}\hat{\Delta}) + \frac{2(k-1)}{k_2(k_2-1)}\operatorname{tr}(D\hat{\Delta}) \\ &+ \frac{2}{k_2(k_2-1)}\operatorname{tr}(\Delta^{-1}U)\operatorname{tr}(D\hat{\Delta}) \\ &+ \frac{4}{k_2(k_2-1)}\operatorname{tr}(UD^2\hat{\Delta}) + \frac{4}{k_2(k_2-1)}\operatorname{tr}(U\Delta^{-1}UD^2\hat{\Delta})\Bigg].\end{aligned} \tag{3.23}$$

REMARK 3.3.1. Formula (3.17) can be generalized up to k-th order (k is a positive integer) provided k-th order moments exist. But we don't develop it here.

3.4. IMPROVING UPON THE UNBIASED ESTIMATOR

Here we give two types of estimators which beat the unbiased estimator (3.5) with respect to the loss function (3.2).

3.4.1. HAFF TYPE ESTIMATORS

Consider the estimator of the form

$$\hat{\Delta}_H(U) = a_1\{U + zt(z)I_p\} \tag{3.24}$$

where a_1 is a constant and $t(z)$ is an absolutely continuous, nonincreasing, and nonnegative function of $z = 1/\operatorname{tr} U^{-1}$. It is analogous to the empirical Bayes estimator of the normal covariance matrix by Haff[19]. Note that these estimators belong to the class (3.14).

To prove that $\hat{\Delta}_H$ beats $\hat{\Delta}_{UN}$ under certain conditions on t, we need the following lemma.

LEMMA 3.4.1. *Let U have the $F_p(k_1, k_2; \Delta)$ distribution. Then we have an inequality*

$$\mathbf{E}\left[\frac{t(z)\operatorname{tr}\Delta^{-1}}{\operatorname{tr}U^{-1}}\right] \le \mathbf{E}\left[\frac{t(z)(k_1 - p + 1)}{k_2 - 2}\right],$$

where equality holds iff $p = 1$ and $t(z)$ is a constant.

PROOF. Put $g(U) = t(z)/\operatorname{tr}(U^{-1})$ and $V(\Delta, U) = (\Delta + U)\Delta^{-1}$ in (3.6). Then, noting that

$$\operatorname{tr}(DU\Delta^{-1}) = \frac{p+1}{2}\operatorname{tr}\Delta^{-1} \qquad \text{(see Haff[18])},$$

the first term of right hand side in (3.6) is equal to

$$(p+1)\mathbf{E}\left[t(z)\operatorname{tr}\Delta^{-1}/\operatorname{tr}(U^{-1})\right].$$

Using $(\partial/\partial U)\operatorname{tr}U^{-1} = -U^{-2}_{(2)}$ (see Haff [19]), we get

$$\frac{\partial}{\partial U}g(U) = \frac{t(z)U^{-2}_{(2)}}{(\operatorname{tr}U^{-1})^2} + \frac{t'(z)U^{-2}_{(2)}}{(\operatorname{tr}U^{-1})^3}.$$

From these and the equation $\operatorname{tr} A_{(\tau)}B_{(1/\tau)} = \operatorname{tr} AB$ for any $p\times p$ matrices A and B, direct calculation shows that (3.6) provides

$$k_2 \mathbf{E}\left[\frac{t(z)\operatorname{tr}\Delta^{-1}}{\operatorname{tr}U^{-1}}\right] = \mathbf{E}\left[\frac{2t(z)\operatorname{tr}(U^{-1}\Delta^{-1}+U^{-2})}{(trU^{-1})^2} + \frac{2t'(z)\operatorname{tr}(U^{-2}+U^{-1}\Delta^{-1})}{(\operatorname{tr}U^{-1})^3} + t(z)(k_1-p-1)\right]. \tag{3.25}$$

Note that $t(z)\geq 0$. Applying $\operatorname{tr}(U^{-1}\Delta^{-1}) \leq (\operatorname{tr}U^{-1})(\operatorname{tr}\Delta^{-1})$ and $\operatorname{tr}U^{-2}/(\operatorname{tr}U^{-1})^2 \leq 1$ to the first term of right hand side in (3.25) and noting that the second term of right hand side in (3.25) is less than zero because of $t'(z)\leq 0$, we get the desired result.

THEOREM 3.4.1. *For $p\geq 2$ and $k_i > p+1 (i=1,2)$, the estimators of the form* (3.24) *given by $a_1 = (k_2-p-1)/k_1$ and $t(z)$ an absolutely continuous and nonincreasing function bounded by*

$$0\leq t(z) \leq \frac{2(p-1)(k_1+k_2-p-1)}{k_1(k_2-2)}, \tag{3.26}$$

beat the unbiased estimator $\hat{\Delta}_{UN}$ under the loss (3.2) .

PROOF. Put

$$\alpha_1(\Delta) = \mathbf{R}_1(\Delta,\hat{\Delta}_H) - \mathbf{R}_1(\Delta,\hat{\Delta}_{UN}).$$

Noting that $\log|I+A| \geq \operatorname{tr} A - (1/2)\operatorname{tr} A^2$ for any positive definite matrix A, a condition for $\alpha_1(\Delta)\leq 0$ may be written as

$$\mathbf{E}\left[\frac{1}{2}t^2(z) - t(z) + \frac{a_1 t(z)\operatorname{tr}\Delta^{-1}}{\operatorname{tr}U^{-1}}\right] \leq 0.$$

Using Lemma 3.4.1 it is seen that the condition (3.26) is sufficient for $\alpha_1(\Delta)\leq 0$.

REMARK 3.4.1. Since $S = k_2 U$ converges to Wishart distribution $W_p(k_1,\Delta)$ weakly as n_2 tends to infinity, the estimators in Theorem 3.4.1 with $t^*(z) = t(k_2 z)$ reduces to the estimators of the covariance matrix Δ given by

$$\hat{\Delta}_H = (1/k_1)(S + ut^*(u)I_p),$$

where $t^*(u)$ is an absolutely continuous, nonincreasing function of $z = 1/\operatorname{tr}S^{-1}$ bounded by

$$0\leq t^*(z) \leq \frac{2(p-1)}{k_1}.$$

Theorem 3.4.1 implies that $\hat{\Delta}_H$ dominates $\hat{\Delta}_{UN} = S/k_1$ which was obtained by Haff[19].

3.4.2. PERRON TYPE ESTIMATOR

For notation, let $L_i = \operatorname{diag}(l_1, \ldots, l_{i-1}, 0, l_{i+1}, \ldots, l_p)$, $i = 1, \ldots, p$, and

$$\operatorname{tr}_m(L) = \begin{cases} 1 & \text{if } m = 0, \\ \sum_{1 \le i_1 < \cdots < i_m \le p} \prod_{j=1}^{m} l_{i_j} & \text{if } m = 1, \ldots, p, \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, let

$$w_{im} = \operatorname{tr}_{m-1}(L_i) \operatorname{tr}_{m-1}^{-1}(L) - \operatorname{tr}_m(L_i) \operatorname{tr}_m^{-1}(L)$$

and let $d_1, \ldots, d_p$ be nonnegative constants with $d_1 \le \cdots \le d_p$. Consider the estimator of the form $\hat{\Delta}_P = H\varphi(L)H'$ with $\varphi(L) = \operatorname{diag}(\varphi_1(L), \ldots, \varphi_p(L))$ and

$$\varphi_i(L) = l_i \sum_{m=1}^{p} w_{im} d_m. \tag{3.27}$$

We shall record the computational lemma from Perron[46].

LEMMA 3.4.2.

(i) $\operatorname{tr}_m(L) = l_i \operatorname{tr}_{m-1}(L_i) + \operatorname{tr}_m(L_i)$.

(ii) $\sum_i \operatorname{tr}_m(L_i) = (p-m) \operatorname{tr}_m(L)$.

(iii) *Setting*

$$L_{ij} = \operatorname{diag}(l_1, \ldots, l_{i-1}, 0, l_{i+1}, \ldots, l_{j-1}, 0, l_{j+1}, \ldots, l_p) \quad \textit{for } i \ne j,$$

$$\sum_i \sum_{j \ne i} \operatorname{tr}_m(L_{ij}) = (p-m)(p-m-1) \operatorname{tr}_m(L).$$

(iv) $\operatorname{tr}_m(L_i) - \operatorname{tr}_m(L_j) = (l_j - l_i) \operatorname{tr}_{m-1}(L_{ij}) \quad \textit{for } i \ne j.$

(v) $\operatorname{tr}_m^2(L) - \operatorname{tr}_{m-1}(L) \operatorname{tr}_{m+1}(L) \ge 0 \quad \textit{for } m = 1, \ldots, p-1.$

PROOF. (i) may be obtained from the combinatoric calculation.

(ii) Observe that

$$\begin{aligned}\sum_i \operatorname{tr}_m(L_i) &= \sum_i \{ \operatorname{tr}_m(L) - l_i \operatorname{tr}_{m-1}(L_i)\} \\ &= p \operatorname{tr}_m(L) - \sum_i l_i \operatorname{tr}_{m-1}(L_i) \\ &= (p-m) \operatorname{tr}_m(L),\end{aligned}$$

since

$$\begin{aligned}\sum_i l_i \operatorname{tr}_{m-1}(L_i) &= \sum_i l_i \sum_{\substack{\lambda_1 < \cdots < \lambda_{m-1} \\ \lambda_j \neq i}} \prod_j l_{\lambda_j} \\ &= \sum_{i<\lambda_1<\cdots<\lambda_{m-1}} l_i \prod_j l_{\lambda_j} + \sum_{\lambda_1<i<\lambda_2<\cdots<\lambda_{m-1}} l_i \prod_j l_{\lambda_j} \\ &\quad + \cdots + \sum_{\lambda_1<\cdots<\lambda_{m-1}<i} l_i \prod_j l_{\lambda_j} \\ &= m \sum_{\lambda_1<\cdots<\lambda_m} \prod_j l_{\lambda_j} \\ &= m \operatorname{tr}_m(L).\end{aligned}$$

(iii) Using (i) of this lemma, it can be seen that

$$\begin{aligned}\sum_i \sum_{j\neq i} \operatorname{tr}_m(L_{ij}) &= \sum_i (p-m-1) \operatorname{tr}_m(L_i) \\ &= (p-m)(p-m-1) \operatorname{tr}_m(L),\end{aligned}$$

which completes the proof of (iii).

(iv) Using (i) of this lemma and the direct calculation give the desired result.

(v) If $m = p$ the proof is trivial. In order to complete the proof for $m \leq p-1$, define

$$\begin{aligned}A(k_2) =& \{(\alpha_{11}, \ldots, \alpha_{1k_1}, \alpha_{21}, \ldots, \alpha_{2k_2}) : k_1 = 2(m-k_2), 1 \leq \alpha_{11} < \cdots < \alpha_{1k_1} \leq p, \\ & 1 \leq \alpha_{21} < \cdots < \alpha_{2k_2} \leq p, \{\alpha_{11}, \ldots, \alpha_{1k_1}\} \cap \{\alpha_{21}, \ldots, \alpha_{2k_2}\} = \phi\}.\end{aligned}$$

By a combiantoric argument it can be seen that

$$\operatorname{tr}_m^2(L) - \operatorname{tr}_{m-1}(L) \operatorname{tr}_{m+1}(L) = \sum_{k_2 = 0 \vee (2m-p)}^{m} \sum_{A(k_2)} \prod_{r=1}^{2} \prod_{s=1}^{k_r} l_{\alpha_{rs}}^{r},$$

which is always greater or equal to zero. This completes the proof of (v).

LEMMA 3.4.3.

(i) $\tilde{W} = (w_{ij})$ is doulbly stochastic.

(ii)
$$l_i \frac{\partial}{\partial l_i}\left[\frac{\operatorname{tr}_m(L_i)}{\operatorname{tr}_m(L)}\right] = -\frac{\operatorname{tr}_m(L_i)}{\operatorname{tr}_m(L)}\left(1 - \frac{\operatorname{tr}_m(L_i)}{\operatorname{tr}_m(L)}\right) \le 0.$$

(iii)
$$\sum_{i>j} \frac{l_i w_{im}(L) - l_j w_{jm}(L)}{l_i - l_j} = p - m.$$

(iv) *If* $d_1 < \cdots < d_p$, *then*

$$d_1 < \phi_1 < \phi_2 < \cdots < \phi_p < d_p$$

where $\phi_i = \sum_{m=1}^{p} w_{im} d_m$.

(v) *If* $l_1 > \cdots > l_p$, *then* $\varphi_1 > \cdots > \varphi_p$ *where* φ_i*'s are given by* (3.27).

PROOF. (i) From (i) of Lemma 3.4.2 it may be seen that

$$\sum_i w_{im}(L) = \sum_i \frac{\operatorname{tr}_{m-1}(L_i)}{\operatorname{tr}_{m-1}(L)} - \sum_i \frac{\operatorname{tr}_m(L_i)}{\operatorname{tr}_m(L)} = (p - m + 1) - (p - m) = 1.$$

Also,

$$\begin{aligned}\sum_m w_{im} &= \sum_m \left[\frac{\operatorname{tr}_{m-1}(L_i)}{\operatorname{tr}_{m-1}(L)} - \frac{\operatorname{tr}_m(L_i)}{\operatorname{tr}_m(L)}\right] \\ &= \frac{\operatorname{tr}_0(L_i)}{\operatorname{tr}_0(L)} - \frac{\operatorname{tr}_p(L_i)}{\operatorname{tr}_p(L)} = 1,\end{aligned}$$

since $\operatorname{tr}_0(\cdot) = 1$ and $\operatorname{tr}_p(L_i) = 0$. Finally, using (i) and (v) of Lemma 3.4.2 and some calculation show that

$$\begin{aligned}w_{im}(L) &= \frac{\operatorname{tr}_{m-1}(L_i)\operatorname{tr}_m(L) - \operatorname{tr}_m(L_i)\operatorname{tr}_{m-1}(L)}{\operatorname{tr}_{m-1}(L)\operatorname{tr}_m(L)} \\ &= \frac{l_i\{\operatorname{tr}^2_{m-1}(L_i) - \operatorname{tr}_m(L_i)\operatorname{tr}_{m-2}(L_i)\}}{\operatorname{tr}_{m-1}(L)\operatorname{tr}_m(L)} \ge 0.\end{aligned}$$

These complete the proof of (i).

(ii) is obtained from the direct calculation.

(iii) If $m = p$ then $l_i w_{im}(L) = l_j w_{jm}(L)$. For $m \le p-1$,

$$\sum_i \sum_{j<i} \frac{l_i w_{im}(L) - l_j w_{jm}(L)}{l_i - l_j}$$

$$= \sum_i \sum_{j<i} \left[l_i \left\{ \frac{\mathrm{tr}_{m-1}(L_i)}{\mathrm{tr}_{m-1}(L)} - \frac{\mathrm{tr}_m(L_i)}{\mathrm{tr}_m(L)} \right\} - l_j \left\{ \frac{\mathrm{tr}_{m-1}(L_j)}{\mathrm{tr}_{m-1}(L)} - \frac{\mathrm{tr}_m(L_j)}{\mathrm{tr}_m(L)} \right\} \right] \Big/ (l_i - l_j)$$

$$= \sum_i \sum_{j<i} \left[\left\{ \frac{l_i\, \mathrm{tr}_{m-1}(L_i) - l_j\, \mathrm{tr}_{m-1}(L_j)}{(l_i - l_j)\, \mathrm{tr}_{m-1}(L)} \right\} - \left\{ \frac{l_i\, \mathrm{tr}_m(L_i) - l_j\, \mathrm{tr}_m(L_j)}{(l_i - l_j)\, \mathrm{tr}_m(L)} \right\} \right]$$

$$= \sum_i \sum_{j<i} \left[\left\{ \frac{\mathrm{tr}_m(L_j) - \mathrm{tr}_m(L_i)}{(l_i - l_j)\, \mathrm{tr}_{m-1}(L)} \right\} - \left\{ \frac{\mathrm{tr}_{m+1}(L_j) - \mathrm{tr}_{m+1}(L_i)}{(l_i - l_j)\, \mathrm{tr}_m(L)} \right\} \right]$$

(by (i) of Lemma 3.4.2)

$$= \sum_i \sum_{j<i} \left[\frac{\mathrm{tr}_{m-1}(L_{ij})}{\mathrm{tr}_{m-1}(L)} - \frac{\mathrm{tr}_m(L_{ij})}{\mathrm{tr}_m(L)} \right]$$

(by (v) of Lemma 3.4.2)

$$= \frac{1}{2} \sum_i \sum_{j \neq i} \left[\frac{\mathrm{tr}_{m-1}(L_{ij})}{\mathrm{tr}_{m-1}(L)} - \frac{\mathrm{tr}_m(L_{ij})}{\mathrm{tr}_m(L)} \right]$$

$$= \frac{1}{2}[(p-m+1)(p-m) - (p-m)(p-m-1)]$$

(by (iii) of Lemma 3.4.2)

$$= p - m,$$

which completes the proof of (iii).

(iv) Observe that

$$\phi_i = \sum_{m=1}^{p} w_{im} d_m$$
$$= d_1 + \sum_{m=1}^{p-1} \frac{\mathrm{tr}_m(L_i)}{\mathrm{tr}_m(L)} (d_{m+1} - d_m).$$

Since $\mathrm{tr}_m(L_i) < \mathrm{tr}_m(L_j)$ if $i < j$, it can be seen that ϕ_i is nondecreasing in i . Moreover, ϕ_i is convex combination of $d_1, \ldots, d_p$, therefore $d_1 \le \phi_i \le d_p$, $i = 1, \ldots, p$.

(v) If $i < j$, then

$$\varphi_i - \varphi_j = l_i w_{i1} d_1 - l_j w_{j1} d_1 + \sum_{m=2}^{p} (l_i w_{im} - l_j w_{jm}) d_m$$
$$= \frac{\mathrm{tr}_2(L_j) - \mathrm{tr}_2(L_i)}{\mathrm{tr}(L)} d_1 + \sum_{m=2}^{p} (l_i w_{im} - l_j w_{jm}) d_m.$$

The remainder of the proof is to see that $l_i w_{im} - l_j w_{jm} \geq 0$ for $m = 2, \ldots, p$. If $m = p$ then $l_i w_{im} = l_j w_{jm}$. For $m \leq p-1$, similar argument as in the proof of (iii) of this lemma leads that

$$\begin{aligned}\frac{l_i w_{im} - l_j w_{jm}}{l_i - l_j} &= \frac{\operatorname{tr}_{m-1}(L_{ij})}{\operatorname{tr}_{m-1}(L)} - \frac{\operatorname{tr}_m(L_{ij})}{\operatorname{tr}_m(L)} \\ &= w_{jm}(L_i)\frac{\operatorname{tr}_{m-1}(L_i)}{\operatorname{tr}_{m-1}(L)} + w_{im}(L)\frac{\operatorname{tr}_m(L_{ij})}{\operatorname{tr}_m(L_i)} \geq 0,\end{aligned}$$

which completes the proof of (v).

The next lemma describes the action of D^2 over the estimators proposed in this section.

LEMMA 3.4.4. *Put* $D^2[H\varphi(L)H'] = H\varphi^{(2)}(L)H'$ *where*

$$\varphi^{(2)}(L) = \operatorname{diag}(\varphi_1^{(2)}(L), \ldots, \varphi_p^{(2)}(L)).$$

If $\varphi_i = l_i \sum_m w_{im} d_m$, *then* $\varphi_i^{(2)} \leq 0$, $i = 1, \ldots, p$.

PROOF. Applying Lemma 2.1.10 gives that

$$\varphi_i^{(1)} = \frac{1}{2}\sum_{j\neq i}\sum_m \frac{l_i w_{im} d_m - l_j w_{jm} d_m}{l_i - l_j} + \sum_m w_{im} d_m + l_i \frac{\partial}{\partial l_i}\sum_m w_{im} d_m.$$

Observe that

$$\begin{aligned}\sum_m \sum_{j\neq i} \frac{l_i w_{im} d_m - l_j w_{jm} d_m}{l_i - l_j} &= \sum_m d_m \sum_{j\neq i}\left\{\frac{\operatorname{tr}_{m-1}(L_{ij})}{\operatorname{tr}_{m-1}(L)} - \frac{\operatorname{tr}_m(L_{ij})}{\operatorname{tr}_m(L)}\right\} \\ &= \sum_{j\neq i}\left\{\sum_{m=1}^{p-1} \frac{\operatorname{tr}_m(L_{ij})}{\operatorname{tr}_m(L)}(d_{m+1} - d_m) + d_1\right\} \\ &= \sum_{m=1}^{p-1}(d_{m+1} - d_m)\sum_{j\neq i}\frac{\operatorname{tr}_m(L_{ij})}{\operatorname{tr}_m(L)} + (p-1)d_1 \\ &= \sum_{m=1}^{p-1}\frac{(p-m-1)\operatorname{tr}_m(L_i)}{\operatorname{tr}_m(L)}(d_{m+1} - d_m) + (p-1)d_1.\end{aligned}$$

The last equality follows from (ii) of Lemma 3.4.2. From (ii) of Lemma 3.4.3 and noting that

$$\sum_{m=1}^{p} w_{im} d_m = d_1 + \sum_{m=1}^{p-1} \operatorname{tr}_m(L_i)\operatorname{tr}_m^{-1}(L)(d_{m+1} - d_m),$$

it can be seen that

$$\sum_{m=1}^{p} l_i \frac{\partial}{\partial l_i} w_{im} d_m = -\sum_{m=1}^{p} w_{im} d_m + d_1 + \sum_{m=1}^{p-1} \frac{\operatorname{tr}_m^2(L_i)}{\operatorname{tr}_m^2(L)}(d_{m+1} - d_m).$$

These imply that

$$\varphi_i^{(1)} = \frac{p+1}{2} d_1 + \sum_{m=1}^{p-1} \left\{ \frac{(p-m-1)\operatorname{tr}_m(L_i)}{\operatorname{tr}_m(L)} + \frac{\operatorname{tr}_m^2(L_i)}{\operatorname{tr}_m^2(L)} \right\} (d_{m+1} - d_m). \tag{3.28}$$

Again applying Lemma 2.1.10 to (3.28), the straightforward calculation leads to

$$\begin{aligned}\varphi_i^{(2)} =& \frac{1}{2} \sum_{j \neq i} \sum_{m=1}^{p-1} \left\{ \frac{(p-m-1)\{\operatorname{tr}_m(L_i) - \operatorname{tr}_m(L_j)\}}{\operatorname{tr}_m(L)(l_i - l_j)} + \frac{\operatorname{tr}_m^2(L_i) - \operatorname{tr}_m^2(L_j)}{\operatorname{tr}_m^2(L)(l_i - l_j)} \right\} \\ & \times (d_{m+1} - d_m) + \sum_{m=1}^{p-1} \frac{\partial}{\partial l_i} \left\{ \frac{(p-m-1)\operatorname{tr}_m(L_i)}{\operatorname{tr}_m(L)} + \frac{\operatorname{tr}_m^2(L_i)}{\operatorname{tr}_m^2(L)} \right\} (d_{m+1} - d_m).\end{aligned}$$

Note that

$$\frac{\operatorname{tr}_m(L_i) - \operatorname{tr}_m(L_j)}{l_i - l_j} \leq 0,$$

and

$$\frac{\operatorname{tr}_m^2(L_i) - \operatorname{tr}_m^2(L_j)}{l_i - l_j} \leq 0.$$

As $\operatorname{tr}_m(L) = l_i \operatorname{tr}_{m-1}(L_i) + \operatorname{tr}_m(L_i)$, we have

$$\frac{\partial}{\partial l_i} \frac{1}{\operatorname{tr}_m(L)} = -\frac{\operatorname{tr}_{m-1}(L_i)}{\operatorname{tr}_m^2(L)} \leq 0.$$

From these and $d_{m+1} > d_m$, we can conclude that $\varphi_i^{(2)} \leq 0$.

THEOREM 3.4.2. *Assume that* $k_i > 3$ *for* $i = 1, 2$. *Let*

$$\hat{\Delta}_P(U) = H\varphi^P(L)H',$$

where H *is an* $m \times m$ *orthogonal matrix,* $U = HLH'$ *so that* $L = \operatorname{diag}(l_1, \ldots, l_p)$ *are eigenvalues of* U, *and* $\varphi^P(L) = \operatorname{diag}(\varphi_1^P, \ldots, \varphi_p^P)$ *with* φ_i^P *given by* (3.27) *and* $d_i = (k_2 - p - 1)/(k_1 + p - 2i + 1)$, $i = 1, \ldots, p$. *Then the estimator* $\hat{\Delta}_P(U)$ *beats the unbiased estimator* (3.5) *with respect to the loss function* (3.2).

PROOF. Using Lemma 2.1.10, (i) and (iii) of Lemma 3.4.3, the direct calculation leads to

$$\operatorname{tr}(D\hat{\Delta}_P) = \sum_m \left\{ (p-m+1)d_m + \sum_i l_i \frac{\partial}{\partial l_i} w_{im} d_m \right\}. \tag{3.29}$$

From (ii) of Lemma 3.4.2, we get

$$\begin{aligned}\sum_m l_i \frac{\partial}{\partial l_i} w_{im} d_m &= l_i \frac{\partial}{\partial l_i}\left[d_1 + \sum_{m=1}^{p-1} \frac{\operatorname{tr}_m(L_i)}{\operatorname{tr}_m(L)}(d_{m+1}-d_m)\right] \\ &= -\sum_{m=1}^{p-1} \frac{\operatorname{tr}_m(L_i)}{\operatorname{tr}_m(L)}\left(1 - \frac{\operatorname{tr}_m(L_i)}{\operatorname{tr}_m(L)}\right)(d_{m+1}-d_m) < 0, \end{aligned} \tag{3.30}$$

which implies that

$$\operatorname{tr}(D\hat{\Delta}_P) \le \sum_m (p-m+1)d_m. \tag{3.31}$$

Note that $\mathbf{E}[\operatorname{tr}(\Delta^{-1}U)] = pk_1/(k_2-p-1)$ from Theorem 3.2.1 and that

$$\log\det(\hat{\Delta}_p) \ge \sum_m \{\log l_m + \log d_m\}$$

from the convexity of logarithemic function and the Jensen's inequality. Applying Corollary 3.3.1 and using (3.31) give that

$$\begin{aligned}\mathbf{R}_1(\Delta,\hat{\Delta}_P) - \mathbf{R}_1(\Delta,\hat{\Delta}_{UN}) \le & \frac{k_1-p-1}{k_2}\sum_m d_m + \frac{2(k-p-1)}{k_2(k_2-p-1)}\sum_m (p-m+1)d_m \\ & -\sum_m \log d_m + p\log\left(\frac{k_2-p-1}{k_1}\right) - p \\ & + \frac{4}{k_2(k_2-1)}\mathbf{E}[\operatorname{tr}\{UD^2\hat{\Delta} + U\Delta^{-1}UD^2\hat{\Delta}\}]. \end{aligned}$$

Note that the last term of the right hand side involving the expectation is negative by Lemma 3.4.4. So we can see that

$$\begin{aligned}\mathbf{R}_1(\Delta,\hat{\Delta}_P) - \mathbf{R}_1(\Delta,\hat{\Delta}_{UN}) \le & \frac{k_1-p-1}{k_2}\sum_m d_m + \frac{2(k-p-1)}{k_2(k_2-p-1)}\sum_m (p-m+1)d_m \\ & -\sum_m \log d_m + p\log\left(\frac{k_2-p-1}{k_1}\right) - p, \end{aligned} \tag{3.32}$$

with $d_m = (k_2-p-1)/(k_1+p-2m+1)$, $m = 1,\ldots,p$. The remainder of the proof proceeds as in Gupta and Krishnamoorthy[16]. Lemma 3.1 in Gupta and Krishnamoorthy[16] tells us that

$$\sum_{m=1}^{p} \log(k+p-2m+1) < p\log k$$

for any positive integers k and p, $k > p$. From this it follows that the right hand side of (3.32) is less than zero if

$$\frac{k_1 - p - 1}{k_2} \sum_{m=1}^{p} \frac{k_2 - p - 1}{k_1 + p - 2m + 1} + \frac{2(k - p - 1)}{k_2} \sum_{m=1}^{p} \frac{p - m + 1}{k_1 + p - 2m + 1} \leq p. \tag{3.33}$$

Writing the left hand side of (3.33) as

$$\begin{aligned} &\frac{k_2 - p - 1}{k_2} \sum_{m=1}^{p} \frac{(k_1 + p - 2m + 1) - 2(p - m + 1)}{k_1 + p - 2m + 1} + \frac{2(k - p - 1)}{k_2} \sum_{m=1}^{p} \frac{p - m + 1}{k_1 + p - 2m + 1} \\ &\qquad = \frac{p(k_2 - p - 1)}{k_2} + \frac{2k_1}{k_2} \sum_{m=1}^{p} \frac{p - m + 1}{k_1 + p - 2m + 1} \end{aligned}$$

shows that (3.33) holds if and only if

$$\sum_{m=1}^{p} \frac{p - m + 1}{k_1 + p - 2m + 1} \leq \frac{p(p+1)}{2k_1},$$

equivalently

$$\sum_{m=1}^{p} \left\{ \frac{p - m + 1}{k_1 + p - 2m + 1} - \frac{p - m + 1}{k_1} \right\} \leq 0.$$

Let

$$a_m = \frac{p - m + 1}{k_1 + p - 2m + 1} - \frac{p - m + 1}{k_1}$$

and $x_m = a_m + a_{p-m+1}$, $m = 1, \ldots, p$. Note that $x_m < 0$ if $m \leq p/2$ and that $a_{(p+1)/2} = 0$ if p is odd. So we can see that

$$\sum_{m=1}^{p} a_m = \begin{cases} \sum_{m=1}^{p/2} x_m < 0 & \text{if p is even} \\ \sum_{m=1}^{(p-1)/2} x_m < 0 & \text{if p is odd}. \end{cases}$$

This completes the proof.

3.5. ORTHOGONALLY INVARIANT MINIMAX ESTIMATORS

Muirhead and Verathaworn[42] derived the best upper triangular invariant estimate of the form $\hat{\Delta}_M(U) = T'GT$ where T is upper triangular with the positive diagonal elements, $U = T'T$, $G = \operatorname{diag}(d_1^M, \ldots, d_p^M)$, and

$$d_i^M = \frac{(k_2 - p - 1 + i)(k_2 - p - 2 + i)}{(k_1 + 1 - i)(k_2 - p - 1 + i) + (p - i)(k - p - 1)}, \qquad i = 1, \ldots, p. \tag{3.34}$$

Since the group of upper triangular matrices is solvable, $\hat{\Delta}_M$ is minimax and beats the unbiased estimator $\hat{\Delta}_{UN}$ in terms of the exact risk $\mathbf{R}_1$. However it is not orthogonally invariant so that the eigenvalues of $\hat{\Delta}_M$ may not be taken as estimates of those of Δ. Later, Gupta and Krshinamoorthy[16] considered the estimator

$$\hat{\Delta}_{AM} = H\varphi^{AM}(L)H', \tag{3.35}$$

where $\varphi^{AM}(L) = \operatorname{diag}(\varphi_1^{AM}(L), \varphi_2^{AM}(L), \ldots, \varphi_p^{AM}(L))$, $\varphi_i^{AM}(L) = d_i^M l_i$, and d_i^M is given by (3.34). Their Monte-Carlo study indicated that it is minimax. However, it has not been established that the proposed estimator has a frequentist risk uniformly smaller than the minimax risk.

In this section, following Konno[31], we shall prove the minimaxity of the estimator $\hat{\Delta}_{AM}$ when $p = 2$ by showing that $\hat{\Delta}_{AM}$ has smaller risk uniformly than $\hat{\Delta}_M$. Furthermore, we consider the estimator of the form

$$\hat{\Delta}_{PM} = H\varphi^{PM}(L)H', \tag{3.36}$$

where $\varphi^{PM} = \operatorname{diag}(\varphi_1^{PM}, \varphi_2^{PM})$ and

$$\begin{aligned}
\varphi_1^{PM} &= \left(\frac{l_1}{l_1 + l_2} d_1^M + \frac{l_2}{l_1 + l_2} d_2^M\right) l_1, \\
\varphi_2^{PM} &= \left(\frac{l_2}{l_1 + l_2} d_1^M + \frac{l_1}{l_1 + l_2} d_2^M\right) l_2.
\end{aligned}$$

The estimator (3.36) is a special case ($p = 2$) of the Perron-type estimator. We shall show that this estimator is also minimax.

THEOREM 3.5.1. *Assume that* $k_i > 3$ *for* $i = 1, 2$ *and that the third order moment of* U *exists. When* $p = 2$, *the estimator*

$$\hat{\Delta}_{AM}(U) = H\varphi^{AM}(L)H',$$

where $\varphi^{AM}(L) = \operatorname{diag}(d_1^M l_1, d_2^M l_2)$ *and*

$$d_i^M = \frac{(k_2 - 3 + i)(k_2 - 4 + i)}{(k_1 + 1 - i)(k_2 - 3 + i) + (2 - i)(k - 3)}, \qquad i = 1, 2, \tag{3.37}$$

beats the minimax estimator $\hat{\Delta}_M$ *with respect to the loss function* (3.2). *So it is minimax.*

PROOF. First, assume that $p \geq 2$. Consider the estimators of the form

$$\hat{\Delta}_A(U) = H\varphi(L)H' \tag{3.38}$$

where $U = HLH'$, $\varphi(L) = \operatorname{diag}(\varphi_1, \ldots, \varphi_p)$, and $\varphi_i = d_i l_i$ ($i = 1, \ldots, p$) with nonnegative constants $d_1 \leq \cdots \leq d_p$. From (3.17) in Theorem 3.3.1, the risk of $\hat{\Delta}_A$ can be wriiten as

$$\begin{aligned} \mathbf{R}_1(\Delta, \hat{\Delta}_A) = &\mathbf{E}\left[\frac{2(k-1)}{k_2(k_2-1)} \operatorname{tr}(D\hat{\Delta}_A) + \frac{2}{k_2(k_2-1)} \operatorname{tr}(\Delta^{-1}U) \operatorname{tr}(D\hat{\Delta}_A)\right. \\ &\left. + \frac{k_1 - p - 1}{k_2} \operatorname{tr}(U^{-1}\hat{\Delta}_A) + M_1(\hat{\Delta}_A) + M_2(\hat{\Delta}_A) + C_p\right], \end{aligned} \tag{3.39}$$

where $M_1(\hat{\Delta}_A)$ and $M_2(\hat{\Delta}_A)$ are given in Theorem 3.3.1 and

$$C_p = -\mathbf{E}\left[\sum_{i=1}^{p} \log(d_i l_i)\right] + \log\det\Delta - p.$$

Note that C_p is a constant term which is independent of parameter Δ. Using (iii) of Lemma 2.1.10, we get

$$\operatorname{tr}(D\hat{\Delta}_A) \leq \sum_{i=1}^{p}(p - i + 1)d_i.$$

Putting above inequality in (3.39) and noting that $\mathbf{E}[\operatorname{tr}(\Delta^{-1}U)] = pk_1/(k_2 - p - 1)$, some simplification shows that $\mathbf{R}_1(\Delta, \hat{\Delta}_A)$ is bounded above by

$$\frac{2(k-p-1)}{k_2(k_2-p-1)} \sum_{i=1}^{p}(p-i+1)d_i + \frac{k_1-p-1}{k_2} \sum_{i=1}^{p} d_i + \mathbf{E}[M_1(\hat{\Delta}_A) + M_2(\hat{\Delta}_A)] + C_p. \tag{3.40}$$

Since $M_1(\hat{\Delta}_A)$ and $M_2(\hat{\Delta}_A)$ involve in higher order derivatives of D, it is difficult to evaluate their expectations for general p. From now on, we assume that $p = 2$. Then it is sufficient to show that

$$\mathbf{R}_1(\Delta, \hat{\Delta}_{AM}) \leq C_2 + 2$$

in order to prove minimaxity. Note that the right hand side is the risk of minimax estimator $\hat{\Delta}_M$ when $p = 2$ and $d_i = d_i^M$ for $i = 1, 2$.

Now we evaluate $M_1(\hat{\Delta}_A)$ and $M_2(\hat{\Delta}_A)$ in (3.40) respectively. From (iii) of Lemma 2.1.10, we have

$$\begin{aligned}
\varphi_1^{(2)}(L) &= \frac{(d_1 - d_2)(l_1 - 2l_2)}{2(l_1 - l_2)^2}, \\
\varphi_2^{(2)}(L) &= \frac{(d_1 - d_2)(2l_1 - l_2)}{2(l_1 - l_2)^2}, \\
\varphi_1^{(3)}(L) &= \frac{(d_1 - d_2)(-3l_1 + 5l_2)}{4(l_1 - l_2)^3}, \\
\varphi_2^{(3)}(L) &= \frac{(d_1 - d_2)(5l_1 - 3l_2)}{4(l_1 - l_2)^3}.
\end{aligned} \tag{3.41}$$

The term inside the curly bracket of $M_1(\hat{\Delta}_A)$ in Theorem 3.3.1 can be rewritten as

$$(k-1)\{l_1\varphi_1^{(2)}(L) + l_2\varphi_2^{(2)}(L)\} + (l_1 + l_2)\{\varphi_1^{(2)}(L) + \varphi_2^{(2)}(L)\} + 2\{l_1^2\varphi_1^{(3)}(L) + l_2^2\varphi_2^{(3)}(L)\}, \tag{3.42}$$

and, from (3.41), (3.42) becomes

$$\frac{(d_1 - d_2)[(k-1)(l_1 + l_2)(l_1 - l_2)^2 + 2l_1l_2(l_1 + l_2)]}{2(l_1 - l_2)^3}$$

which follows

$$\mathbf{E}[M_1(\hat{\Delta}_A)] \leq \frac{2(k-1)(d_1 - d_2)}{k_2(k_2 - 1)(k_2 - 2)}, \tag{3.43}$$

since $d_1 - d_2 < 0$ and $(l_1 + l_2)/(l_1 - l_2) > 1$.

The term inside the curly bracket of $M_2(\hat{\Delta}_A)$ can be rewritten as $\mathrm{tr}(\Delta^{-1}Hh(L)H')$ where $h(L) = \mathrm{diag}(h_1(L), h_2(L))$ and

$$h_i(L) = \{l_1\varphi_1^{(2)}(L) + l_2\varphi_2^{(2)}(L)\}l_i + \{\varphi_1^{(2)}(L) + \varphi_2^{(2)}(L)\}l_i^2 + 2\varphi_i^{(3)}(L)l_i^3, \qquad i = 1, 2.$$

Putting (3.41) into this, we get

$$h_1(L) = \frac{(d_1 - d_2)\{(l_1 - l_2)^2 l_1 + l_1 l_2^2 + l_2^3\}l_1}{2(l_1 - l_2)^3} \leq \frac{(d_1 - d_2)l_1}{2}$$

and

$$h_2(L) = \frac{(d_1 - d_2)\{l_1^3 + 2l_1^2 l_2 - 2l_1 l_2^2 + l_2^3\} l_2}{2(l_1 - l_2)^3} \leq \frac{(d_1 - d_2) l_2}{2}.$$

From these, we obtain that

$$\begin{aligned} \mathbf{E}\,[M_2(\hat{\Delta}_{\boldsymbol{A}})] \leq & \frac{2(d_1 - d_2)}{k_2(k_2 - 1)(k_2 - 2)} \mathbf{E}\,[\,\mathrm{tr}\,(\Delta^{-1} U)] \\ = & \frac{4(d_1 - d_2) k_1}{k_2(k_2 - 1)(k_2 - 2)(k_2 - 3)}. \end{aligned} \tag{3.44}$$

Combining (3.43) and (3.44) gives that

$$\mathbf{E}\,[M_1(\hat{\Delta}_{\boldsymbol{A}}) + M_2(\hat{\Delta}_{\boldsymbol{A}})] \leq \frac{2(k-3)(d_1 - d_2)}{k_2(k_2 - 2)(k_2 - 3)}.$$

From this and hazardous calculation, (3.40) with $p = 2$ provides that

$$\begin{aligned} \frac{2(k-3)}{k_2(k_2 - 3)}(2d_1 + d_2) + & \frac{k_1 - 3}{k_2}(d_1 + d_2) + \mathbf{E}\,[M_1(\hat{\Delta}_{\boldsymbol{A}}) + M_2(\hat{\Delta}_{\boldsymbol{A}})] + C_2 \\ & \leq \frac{k_1(k_2 - 2) + k - 3}{(k_2 - 2)(k_2 - 3)} d_1 + \frac{k_1 - 1}{k_2 - 2} d_2 + C_2. \end{aligned} \tag{3.45}$$

Finally, letting $d_i = d_i^{\boldsymbol{M}}$ $(i = 1, 2)$ given by (3.37), we get that

$$\mathbf{R}_1(\Delta, \hat{\Delta}_{\boldsymbol{AM}}) \leq \mathbf{R}_1(\Delta, \hat{\Delta}_{\boldsymbol{M}}),$$

which completes the proof.

REMARK 3.5.1. If $\mathbf{E}\,[M_1(\hat{\Delta}_{\boldsymbol{A}}) + M_2(\hat{\Delta}_{\boldsymbol{A}})] \leq 0$ for general p, then (3.40) becomes the upper bound (due to Gupta and Krishnamoorthy [16]) for the approximate risk of the estimates $\hat{\Delta}_{\boldsymbol{A}}$. From this, it is seen that the approximate risk given by Muirhead and Verathaworn[42] neglects the terms $M_1(\hat{\Delta}_{\boldsymbol{A}})$ and $M_1(\hat{\Delta}_{\boldsymbol{A}})$ as long as we restrict our attention to the class of the estimators of the form (3.35). Furthermore we expect that $\hat{\Delta}_{\boldsymbol{AM}}$ is minimax for higher dimensions as pointed out in Gupta and Krishnamoorthy [16]. But we can't give analytic proof since evaluation of $\mathbf{E}\,[M_1 + M_2]$ is much more difficult than that done in Theorem 3.5.1.

REMARK 3.5.2. In Gupta and Krishnamoorthy[16], several competitors among possible choice of d_i's are considered . The risk of the estimators of the form (3.35) is bounded above by the left hand side of (3.45). By differentiating (3.45) with respect to d_1 or d_2, then it is easily seen that upper bound of the risk of the estimators of the form (3.35) is minimized when $d_i = d_i^M$ $(i = 1, 2)$ so that it seems that d_i^M's are the best constants for the estimators of the form (3.35) when $p = 2$.

REMARK 3.5.3. From the Monte-Carlo study of Gupta and Krishnamoorthy[16], it is found that the estimator $\hat{\Delta}_{AM}$ performs better than the approximate Bayes estimator due to Muirhead and Verathaworn[42] when $p = 2$.

THEOREM 3.5.2. *Assume that* $p = 2$ *and* $k_i > 3$ *for* $i = 1, 2$, *and that the third order moment of* U *exists. Let* $U = HLH'$ *with* $L = \operatorname{diag}(l_1, l_2)$ *and* $l_1 > l_2$, *and let* $\varphi^{PM}(L) = \operatorname{diag}(\varphi_1^{PM}(L), \varphi_2^{PM}(L))$ *where*

$$\begin{aligned} \varphi_1^{PM}(L) =& \left(\frac{l_1}{l_1 + l_2} d_1^M + \frac{l_2}{l_1 + l_2} d_2^M \right) l_1, \\ \varphi_2^{PM}(L) =& \left(\frac{l_2}{l_1 + l_2} d_1^M + \frac{l_1}{l_1 + l_2} d_2^M \right) l_2, \end{aligned} \tag{3.46}$$

and $d_i^M (i = 1, 2)$ *is given by* (3.37). *Then the estimator*

$$\hat{\Delta}_{PM}(U) = H\varphi^{PM}(L)H'$$

beats the minimax estimator $\hat{\Delta}_M$ *with respect to the loss function* (3.2). *So it is minimax.*

PROOF. Note that (3.46) is a special case , i.e., $p = 2$, of (3.27). Use Theorem 3.3.1 and write the risk of $\hat{\Delta}_{PM}$ as

$$\begin{aligned} \mathbf{R}_1(\Delta, \hat{\Delta}_{PM}) =& \mathbf{E} \left[\frac{2(k-1)}{k_2(k_2 - 1)} \operatorname{tr}(D\hat{\Delta}_{PM}) + \frac{2}{k_2(k_2 - 1)} \operatorname{tr}(\Delta^{-1}U) \operatorname{tr}(D\hat{\Delta}_{PM}) \right. \\ &+ \frac{k_1 - 3}{k_2} \operatorname{tr}(U^{-1}\hat{\Delta}_{PM}) + M_1(\hat{\Delta}_{PM}) + M_2(\hat{\Delta}_{PM}) \\ &\left. - \log\det(\Delta^{-1}\hat{\Delta}_{PM}) - 2 \right]. \end{aligned}$$

From (3.29) and (3.30), we can see that

$$\operatorname{tr}(D\hat{\Delta}_{PM}) = 2d_1^M + d_2^M + \frac{2l_1 l_2}{(l_1 + l_2)^2}(d_1^M - d_2^M).$$

From these, we can see that

$$\frac{2(k-1)}{k_2(k_2-1)}\operatorname{tr}(D\hat{\Delta}_{PM}) \le \frac{2(k-1)}{k_2(k_2-1)}\sum_{i=1}^{2}(3-i)d_i^{M}$$

and

$$\begin{aligned}\frac{2}{k_2(k_2-1)}\operatorname{tr}(\Delta^{-1}U)\operatorname{tr}(D\hat{\Delta}_{PM}) =& \frac{4k_1}{k_2(k_2-1)(k_2-3)}\sum_{i=1}^{2}(3-i)d_i^{M} \\ &+\frac{4l_1l_2(d_1^M-d_2^M)}{k_2(k_2-1)(l_1+l_2)^2}\operatorname{tr}(\Delta^{-1}U).\end{aligned}$$

From these and noting that $\mathbf{E}[\operatorname{tr}(\Delta^{-1}U)] = 2k_1/(k_2-3)$ by Theorem 3.2.1, we can see that

$$\begin{aligned}\mathbf{R}_1(\Delta,\hat{\Delta}_{PM}) \le& \frac{2(k-3)}{k_2(k_2-3)}\sum_{i=1}^{2}(3-i)d_i^{M} + \frac{k_1-3}{k_2}\sum_{i=1}^{2}d_i^{M} \\ &+\mathbf{E}\Bigg[M_1(\hat{\Delta}_{PM})+M_2(\hat{\Delta}_{PM}) \\ &+\frac{4l_1l_2(d_1^M-d_2^M)}{k_2(k_2-1)(l_1+l_2)^2}\operatorname{tr}(\Delta^{-1}U) \\ &-\log\det(\Delta^{-1}\hat{\Delta}_{PM})-2\Bigg].\end{aligned}\tag{3.47}$$

Now we shall evaluate the expectation of the terms inside the large bracket of (3.47). We shall use notation φ_i and d_i instead of φ_i^{PM} and d_i^{M}, $i=1,2$, for convenience. From (iii) of Lemma 2.1.10, we may see that

$$\begin{aligned}\varphi_1^{(2)} =& \left\{\frac{1}{2(l_1+l_2)}+\frac{2l_2^2}{(l_1+l_2)^3}\right\}(d_1-d_2) \\ \varphi_2^{(2)} =& \left\{\frac{1}{2(l_1+l_2)}+\frac{2l_1^2}{(l_1+l_2)^3}\right\}(d_1-d_2) \\ \varphi_1^{(3)} =& -\left\{\frac{3}{2(l_1+l_2)^2}+\frac{6l_2^2}{(l_1+l_2)^4}\right\}(d_1-d_2) \\ \varphi_2^{(3)} =& -\left\{\frac{3}{2(l_1+l_2)^2}+\frac{6l_1^2}{(l_1+l_2)^4}\right\}(d_1-d_2).\end{aligned}\tag{3.48}$$

From (3.42) and (3.48), the term inside the curly bracket of $M_1(\hat{\Delta}_{PM})$ becomes

$$\left[\frac{k-1}{2}+\frac{2l_1l_2}{(l_1+l_2)^2}\left\{k-\frac{12l_1l_2}{(l_1+l_2)^2}\right\}\right](d_1-d_2) \le \frac{k-1}{2}(d_1-d_2),\tag{3.49}$$

since $k > 6$, $d_1 < d_2$, and $l_1 l_2/(l_1+l_2)^2 < 1/2$. Write the term inside the curly bracket of $M_2(\hat{\Delta}_{\boldsymbol{PM}})$ as $\operatorname{tr}(\Delta^{-1}Hh(L)H')$ where $h(L) = \operatorname{diag}(h_1(L), h_2(L))$ and

$$h_i(L) = \{l_1\varphi_1^{(2)}(L) + l_2\varphi_2^{(2)}(L)\}l_i + \{\varphi_1^{(2)}(L) + \varphi_2^{(2)}(L)\}l_i^2 + 2\varphi_i^{(3)}(L)l_i^3, \quad i = 1, 2.$$

Then we can rewrite

$$M_2(\hat{\Delta}_{\boldsymbol{PM}}) + \frac{4l_1l_2(d_1-d_2)}{k_2(k_2-1)(l_1+l_2)^2}\operatorname{tr}(\Delta^{-1}U)$$

as

$$\frac{4}{k_2(k_2-1)(k_2-2)}\operatorname{tr}(\Delta^{-1}Hg(L)H')$$

where $g(L)$ is a diagonal matrix whose elements are given by

$$\begin{aligned} g_1(L) =& h_1(L) + \frac{(k_2-2)l_1^2l_2(d_1-d_2)}{(l_1+l_2)^2}, \\ g_2(L) =& h_2(L) + \frac{(k_2-2)l_1l_2^2(d_1-d_2)}{(l_1+l_2)^2}. \end{aligned} \tag{3.50}$$

Putting (3.48) into (3.50) and some simplification lead to

$$g_1(L) = \left[\frac{l_1}{2} + \frac{1}{(l_1+l_2)^4}\left\{(k_2-1)l_1^4l_2 + 2(k_2-5)l_1^3l_2^2 + (k_2+3)l_1^2l_2^3\right\}\right](d_1-d_2)$$

and

$$g_2(L) = \left[\frac{l_2}{2} + \frac{1}{(l_1+l_2)^4}\left\{(k_2+3)l_1^3l_2^2 + 2(k_2-5)l_1^2l_2^3 + (k_2-1)l_1l_2^4\right\}\right](d_1-d_2).$$

Since $k_2 > 3$, $d_1 < d_2$, and $l_1 > l_2$, we get that

$$g_i(L) \le \frac{l_i}{2}(d_1-d_2), \quad i = 1, 2. \tag{3.51}$$

Again using $\mathbf{E}[\operatorname{tr}(\Delta^{-1}U)] = 2k_1/(k_2-3)$ and combining (3.49) and (3.51) provide that

$$\begin{aligned} &\mathbf{E}\left[M_1(\hat{\Delta}_{\boldsymbol{PM}}) + M_2(\hat{\Delta}_{\boldsymbol{PM}}) + \frac{4l_1l_2(d_1-d_2)}{k_2(k_2-1)(l_1+l_2)^2}\operatorname{tr}(\Delta^{-1}U)\right] \\ &\qquad \le \left[\frac{2(k-1)}{k_2(k_2-1)(k_2-2)} + \frac{4k_1}{k_2(k_2-1)(k_2-2)(k_2-3)}\right](d_1-d_2) \\ &\qquad = \frac{2(k-3)(d_1-d_2)}{k_2(k_2-2)(k_2-3)}. \end{aligned} \tag{3.52}$$

From the convexity of the logarithmic function and the Jensen's inequality, we can see that

$$\log\det(\hat{\Delta}_{PM}) \geq \sum_{i=1}^{2}\{\log l_i + \log d_i\}. \tag{3.53}$$

Now we set $d_i = d_i^M$ $(i = 1, 2)$ again. From (3.47), (3.52), and (3.53), some simplification shows that the risk of the estimator $\hat{\Delta}_{PM}$ is bounded above by

$$\frac{k_1(k_2-2)+k-3}{(k_2-2)(k_2-3)} d_1^M + \frac{k_1-1}{k_2-2} d_2^M + C_2,$$

where

$$C_2 = -\mathbf{E}\,[\sum_{i=1,2} \log(d_i^M l_i)] + \log\det(\Delta) - 2$$

and d_i^M, $i = 1, 2$ is given by (3.3). This completes the proof.

REMARK 3.5.4. As far as we consider the estimator of the form (3.14) with (3.27), i.e., Perron-type, we can see that $\mathbf{E}\,[M_1 + M_2] \leq 0$ for general p. However, to prove the minimaxity in Theorem 3.5.2, we need sharper bound of $\mathbf{E}\,[M_1 + M_2]$ for $p = 2$. It seems difficult to obtain similar bound of $\mathbf{E}\,[M_1 + M_2]$ for general p.

REMARK 3.5.5. Although both $\hat{\Delta}_{AM}$ and $\hat{\Delta}_{PM}$ are minimax, it seems that $\hat{\Delta}_{PM}$ is preferable to $\hat{\Delta}_{AM}$ since possibly $\hat{\Delta}_{AM}$ may violate the desirable order $\varphi_1^{AM} \geq \varphi_2^{AM}$, nor does $\hat{\Delta}_{PM}$.

REMARK 3.5.6. To apply Theorem 3.3.1, we assume the existence of the third order moment of the matrix U. But the minimaxity of $\hat{\Delta}_{AM}$ and $\hat{\Delta}_{PM}$ seems still true provided the moment of the matrix U exists.

REMARK 3.5.7. Sharma and Krishnamoorthy[47] and Takemura[51] independently obtained an orthogonally invariant minimax estimator for a normal covariance matrix by averaging the minimax estimator of James and Stein[24] over $p \times p$ orthogonal matrix with respect to Haar measure. Following their approach for $p = 2$, the estimator of the form

$$\hat{\Delta}_{ST}(U) = H\varphi^{ST}(L)H'$$

where $\varphi^{ST}(L) = \operatorname{diag}(\varphi_1^{ST}(L), \varphi_2^{ST}(L))$ and

$$\varphi_1^{ST}(L) = \left(\frac{\sqrt{l_1}}{\sqrt{l_1}+\sqrt{l_2}}d_1^M + \frac{\sqrt{l_2}}{\sqrt{l_1}+\sqrt{l_2}}d_2^M\right) l_1$$

and

$$\varphi_2^{ST}(L) = \left(\frac{\sqrt{l_2}}{\sqrt{l_1}+\sqrt{l_2}}d_1^M + \frac{\sqrt{l_1}}{\sqrt{l_1}+\sqrt{l_2}}d_2^M\right) l_2$$

must be minimax. But we can't give direct proof as in Theorem 3.5.1 or 3.5.2.

3.6. IMPROVED ESTIMATORS UNDER THE SQUARED ERROR LOSS

In this section, we will look briefly at the problem of estimating the eigenvalues of Δ using the loss function (3.3). It is shown in Leung and Muirhead[34] that, for the loss function (3.3) and $k_2 > p+3$, the best estimator of the form aU is given by $\hat{\Delta}_B = a_2 U$ where

$$a_2 = \frac{(k_2-p)(k_2-p-3)}{(k_1+p+1)(k_2-p-1)+pk_1+2}. \tag{3.54}$$

We consider alternative estimators of the form

$$\hat{\Delta}_H = a_2(U + ztI_p)$$

where $z = 1/\operatorname{tr} U^{-1}$ and t is a nonnegative function. Here we lack the generality of t under the loss function (3.3). Now our goal is to find a sufficient condition under which the estimators $\hat{\Delta}_H$ beats $\hat{\Delta}_B$ under the loss function (3.3). Put

$$\begin{aligned}\alpha_2(\Delta) =& \mathbf{R}_2(\hat{\Delta}_H, \Delta) - \mathbf{R}_2(\hat{\Delta}_B, \Delta)\\ =& 2a_2^2 t\mathbf{E}\left[\frac{\operatorname{tr}(U\Delta^{-2})}{\operatorname{tr} U^{-1}}\right] - 2a_2 t\,\mathbf{E}\left[\frac{\operatorname{tr}\Delta^{-1}}{\operatorname{tr} U^{-1}}\right] + a_2^2 t^2\mathbf{E}\left[\frac{\operatorname{tr}\Delta^{-2}}{(\operatorname{tr} U^{-1})^2}\right].\end{aligned} \tag{3.55}$$

To evaluate (3.55), we need the following lemma given by the application of the identity (3.6).

LEMMA 3.6.1. *Let U have the $F_p(k_1, k_2; \Delta)$ distribution with $k_2 > p+3$. Then the following inequalities hold:*

(i) $$\mathbf{E}\left[\frac{tr\Delta^{-2}}{(\operatorname{tr} U^{-1})^2}\right] \leq \frac{(k_1-p+1)(k_1-p+3)}{(k_2-2)(k_2-4)}\mathbf{E}\left[\frac{\operatorname{tr} U^{-2}}{(\operatorname{tr} U^{-1})^2}\right],$$

(ii) $$\mathbf{E}\left[\frac{\operatorname{tr}\Delta^{-1}}{\operatorname{tr}U^{-1}}\right] \geq \frac{2}{k_2}\mathbf{E}\left[\frac{trU^{-2}}{(\operatorname{tr}U^{-1})^2}\right] + \frac{k_1 - p - 1 + 2\varepsilon}{k_2}$$

where

$$\varepsilon = \frac{p(k_1 - p - 1) + 2}{p^2(pk_2 - 2)}, \tag{3.56}$$

(iii) $$\mathbf{E}\left[\frac{\operatorname{tr}(U\Delta^{-2})}{\operatorname{tr}U^{-1}}\right] \leq \frac{2}{(k_2 - p - 1)(k_2 - 2)}\left[\frac{(k_1 + k_2)(k_1 - p + 1) + k_1(k_2 - 2)}{k_2}\right.$$
$$\left. + \frac{(k_1 - p + 1)(k_1 - p + 3)}{k_2 - 4}\right]\mathbf{E}\left[\frac{\operatorname{tr}U^{-2}}{(\operatorname{tr}U^{-1})^2}\right] + \frac{k_1(k_1 - p - 1)}{k_2(k_2 - p - 1)}.$$

PROOF . For (i). Take $g(U) = (\operatorname{tr}U^{-1})^{-2}$ and $V = (\Delta + U)\Delta^{-2}$ in the F identity (3.6). From similar calculation in the proof of Lemma 3.4.1 we may see that (3.6) provides

$$\mathbf{E}\left[\frac{\operatorname{tr}\Delta^{-2}}{(\operatorname{tr}U^{-1})^2}\right] \leq \frac{k_1 - p + 3}{k_2 - 4}\mathbf{E}\left[\frac{\operatorname{tr}(U^{-1}\Delta^{-1})}{(\operatorname{tr}U^{-1})^2}\right]. \tag{3.57}$$

Hence it suffices to show that

$$\mathbf{E}\left[\frac{\operatorname{tr}(U^{-1}\Delta^{-1})}{(\operatorname{tr}U^{-1})^2}\right] \leq \frac{k_1 - p + 1}{k_2 - 2}\mathbf{E}\left[\frac{\operatorname{tr}U^{-2}}{(\operatorname{tr}U^{-1})^2}\right]. \tag{3.58}$$

Set $G = U^{-1}$. Then it is seen that G follows the $F_p(k_2, k_1; \Delta^{-1})$ distribution. Applying Lemma 2.1 (being $h(G) = (\operatorname{tr}G)^{-2}$ and $V(G, \Delta) = (\Delta^{-1} + G)G^2$ in (3.6)) with the distribution of G instead of U and noting that

$$\operatorname{tr}D\Delta^{-1}G^2 = \frac{p+2}{2}\operatorname{tr}(\Delta^{-1}G) + \frac{1}{2}(\operatorname{tr}\Delta^{-1})(\operatorname{tr}G)$$

and

$$\operatorname{tr}G^3 = \frac{2p+3}{2}\operatorname{tr}G^2 + \frac{1}{2}(\operatorname{tr}G)^2,$$

similar argument leads to (3.58), which completes the proof of (i).

For (ii). Let $t(z)$ be a constant in (3.25). Then the remainder of the proof is to evaluate $\mathbf{E}[\operatorname{tr}(U^{-1}\Delta^{-1})/(\operatorname{tr}U^{-1})^2]$. Using the fact $p\operatorname{tr}U^{-2} \geq (\operatorname{tr}U^{-1})^2$ and making a transformation $T = \Delta^{-1/2}U\Delta^{-1/2}$, we can see that the term is bounded below by

$$\frac{1}{p}\mathbf{E}\left[\frac{\operatorname{tr}(T^{-1}\Delta^{-2})}{\operatorname{tr}(T^{-1}\Delta^{-1})^2}\right] \geq \frac{1}{p}\mathbf{E}\left[\frac{1}{\operatorname{tr}T^{-1}}\right]. \tag{3.59}$$

Noting that T^{-1} has the $F_p(k_2, k_1; I_p)$ distribution and using Lemma 3.4 in Leung and Muirhead[33], right hand side of (3.59) is bounded below by $\{p(k_1-p-1)+2\}/\{p^2(pk_2-2)\}$, which completes the proof of (ii).

For (iii). Set $g(U) = (\operatorname{tr} U^{-1})^{-1}$ and $V(U, \Delta) = (\Delta + U)\Delta^{-2}U$ in (3.6). From (iv) of Lemma 3.4.1 and similar argument in the proof of Lemma 3.4.1, we may see that (3.6) gives

$$(k_2 - p - 1)\mathbf{E}\left[\frac{\operatorname{tr}(U\Delta^{-2})}{\operatorname{tr} U^{-1}}\right] = k_1\mathbf{E}\left[\frac{\operatorname{tr}\Delta^{-1}}{\operatorname{tr} U^{-1}}\right] + 2\mathbf{E}\left[\frac{\operatorname{tr}(\Delta^{-1}U^{-1} + \Delta^{-2})}{(\operatorname{tr} U^{-1})^2}\right]. \tag{3.60}$$

First put (3.25) ($t(z)$ being a constant) into the first term in right hand side of (3.60) and use (3.57) and (3.58). Then we may get the desired result.

THEOREM 3.6.1. Assume that $k_2 > p + 3$ and $p \geq 2$. Let

$$\begin{aligned}\beta =& \frac{2(k_2-2)(k_2-4)}{(k_1-p+1)(k_1-p+3)}\left[\frac{k_1-p+1+2\varepsilon}{a_2k_2} - \frac{k_1(k_1-p-1)}{k_2(k_2-p-1)}\right.\\ &- \frac{2}{(k_2-p-1)(k_2-2)}\left\{\frac{(k_1+k_2)(k_1-p+1)+k_1(k_2-2)}{k_2}\right.\\ &\left.\left.+\frac{(k_1-p+1)(k_1-p+3)}{k_2-4}\right\}\right]\end{aligned}$$

where a_2 and ε are defined by (3.54) and (3.56) respectively. If $\beta > 0$, then the estimators of the form

$$\hat{\Delta}_H = a_2(U + tI_p/\operatorname{tr} U^{-1}),$$

where $0 \leq t \leq \beta$ beat $\hat{\Delta}_B$ under the loss function (3.3).

PROOF. From (ii) and (iii) of Lemma 3.6.1, the coefficient of t in (3.55) is bounded above by

$$\begin{aligned}&\left[\frac{4a_2^2}{(k_2-p-1)(k_2-2)}\left\{\frac{(k_1+k_2)(k_1-p+1)+k_1(k_2-2)}{k_2}\right.\right.\\ &\qquad\left.\left.+\frac{(k_1-p+1)(k_1-p+3)}{k_2-4}\right\} - \frac{4a_2}{k_2}\right]\mathbf{E}\left[\frac{\operatorname{tr} U^{-2}}{(\operatorname{tr} U^{-1})^2}\right]\\ &\qquad+\frac{2a_2}{k_2}\left\{\frac{a_2k_1(k_1-p-1)}{k_2-p-1} - (k_1-p-1+2\varepsilon)\right\}.\end{aligned} \tag{3.61}$$

Noting that $\varepsilon \geq 0$ and that the term inside the second curly bracket of (3.61) is bounded by

$$(k_1 - p - 1)\left\{\frac{(k_2 - p - 1)^2 - (k_2 - p + 1)}{(k_2 - p - 1)^2} - 1\right\} < 0,$$

it is seen that (3.61) can be bounded above by

$$-\frac{a_2^2\beta(k_1 - p + 1)(k_1 - p + 3)}{(k_2 - 2)(k_2 - 4)}\mathrm{E}\left[\frac{\operatorname{tr} U^{-2}}{(\operatorname{tr} U^{-1})^2}\right]. \tag{3.62}$$

Using (i) of Lemma 3.6.1 and (3.62), straightforward calculation shows that the sufficient condition for $\alpha_2(\Delta) \leq 0$ becomes $t^2 - \beta t \leq 0$, which completes the proof.

REMARK 3.6.1. Similar to Remark 3.4.1, we can see that Theorem 3.6.1 implies Theorem 4.6 in Haff[19].

Unfortunately β is not always positive when $p = 2$. We carry out numerical calculation to see which k_1 and k_2 satisfy $\beta > 0$. It indicates that β is monotonically decreasing in k_1 for each fixed k_2, which follows that β has just one sign change. Table 1 shows that ,for example, β is not positive for $k_1 \geq 43$ when $k_2 = 15$. It also shows that the minimum of k_1 such that β doesn't take positive value for each k_2 first goes down and then goes up as k_2 increases. When $p \geq 3$, our numerical calculation shows that β is always positive.

Table 1. For fixed k_2, minimum of k_1 under which β is not positive.

k_2	10	11	12	13	14	15	16	17	18	19	20	30	40	50
k_1	298	74	53	46	44	43	43	43	44	46	47	64	83	102

REFERENCES

[1] Baranchik, A. J. (1970). A family of minimax estimators of the mean of a multivariate normal distribution. *Ann. Math. Statist.* **41** 642-645.

[2] Berger, J. O. (1980). Improving on inadmissible estimators in continuous exponential families with applications to simultaneous estimation of gamma scale parameters. *Ann. Statist.* **8** 545-571.

[3] Berger, J. O. (1985). *Statistical Decision Theory and Beyesian Analysis, 2nd ed.* Springer, New York.

[4] Bilodeau, M.(1990). Estimation of the eigenvalues of $\Sigma_1\Sigma_2^{-1}$, Rapports de recherche du départment de mathématiques et de statistique, **D.M.S. No.** 90-5, Universite de Montreal.

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

[6] Brewster, J.F. and Zidek, J.V. (1974). Improving upon equivariant estimators. *Ann. Statist.* **2** 21-38.

[7] Brown, L.D. (1968). Inadmissibility of the usual estimators of scale parameters in the problem with unknown location parameters. *Ann. Math. Statist.* **39** 29-48.

[8] Dawid, A. P. (1981). Some matrix-variate distribution theory: Notational condsiderations and a Bayesian application. *Biometrika* **68** 265-274.

[9] DasGupta, A. (1986). A general theorem on decision theory for nonnegative functionals: With application. *Ann. Statist.* **17** 1360-1374.

[10] Dey, D. K. (1988). Simultaneous estimation of eigenvalues. *Ann. Inst. Statist. Math.* **40** 137-147.

[11] Dey, D. K. (1989). On estimation of the scale matrix of the multivariate F distribution. *Commun. Statist. -Theory Meth.* **18** 1373-1383.

[12] Dey, D. K. and Srinivasan, C. (1985). Estimation of a covariance matrix under Stein's loss. *Ann. Statist.* **13** 1581-1591.

[13] Eaton, M. (1987). *Group invariance application in statistis.* NSF-CBMS regional conference series in probability and statistcs.

[14] Efron, B. and Morris, C. (1972). Empirical Bayes on vector observations: an extension of Stein's method. *Biometrika* **59** 335-347.

[15] Efron, B. and Morris, C. (1976). Families of minimax estimators of the mean of a multivariate normal distribution. *Ann. Statist.* 4 11-21.

[16] Gupta, A. K. and Krishnamoorthy, K. (1990). Improved estimators of eigenvalues of $\Sigma_1\Sigma_2^{-1}$. *Statiatics and Decisions* 8 247-263.

[17] Haff, L.R. (1977). Minimax estimators for a multinormal precision matrix. *J. Multivariate Anal.* 7 374-385.

[18] Haff, L. R. (1979). An identity for the Wishart distribution with applications. *J. Multivariate Anal.* 9 531-544.

[19] Haff, L. R. (1980). Empirical Bayes estimation of the multivariate normal covariance matrix, *Ann. Statist.* 8 586-597.

[20] Haff, L. R. (1981). Further identities for the Wishart distribution with applications in regression. *Cand. J. Statist.* 9 215-224.

[21] Haff, L.R. (1982). Identities for the inverse Wishart distribution with computational results in linear and quadratic distribution. *Sankhyā*, **B**. 44 245-258.

[22] Haff, L. R. (1982). Solutions of the Euler-Lagrange equations for certain multivariate normal estimation problem. Unpublished manuscript.

[23] Haff, L. R. (1988). The variational form of certain Bayes estimators. Unpublished manuscript.

[24] James, W., and Stein, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* 1 361-379 Univ. of California Press.

[25] Kariya, T. and Sinha, B.K. (1989). ***Robustness of Statistical Tests***, Academic Press INC., San Diego.

[26] Khatri, C. G. (1959). On the mutual independence of certain statistics. ***Ann. Math. Statist.*** **30** 1258-1262.

[27] Konno, Y. (1988). Exact moments of the multivariate F and Beta distributions, *J. Japan. Statist. Soc.*, **18**, 123-130.

[28] Konno, Y. (1990). Families of minimax estimators of matrix of normal means with unknown covariance matrix. *J. Japan Statist. Soc.* **20** 191-201.

[29] Konno,Y. (1991). On estimation of a matrix of normal means with unknown covariance matrix. *J. Multivariate Anal.* **36** 44-55.

[30] Konno, Y. (1991). A note on estimating eigenvalues of scale matrix of the multivariate F-distribution. ***Ann. Inst. Statist. Math.*** **43** 157-165.

[31] Konno, Y. (1991). On estimating eigenvalues of the scale matrix of the multivariate F distribution. To appear in *Sankhyā.*

[32] Kubokawa, T. (1991). An approach to improving the James-Stein estimator. *J. Multivariate Anal.* **36** 121-126.

[33] Leung, P. L. and Muirhead, R. J. (1987). Estimation of parameter matrices and eigenvalues in MANOVA and canonical correlation analysis, ***Ann. Statist.*** **15** 1651-1666.

[34] Leung, P. L. and Muirhead R. J. (1988). On estimating characteristic roots in a two-sample problem, *Statistical Decision Theory and Related Topics* IV (S. S. Gupta and J. O. Berger, eds.), 397-402.

[35] Lin, P.- E., and Tsai, H.- L. (1973). Generalized Bayes minimax estimators of the multivarite normal mean with unknown covariance matrix. ***Ann. Statist.*** **1** 142-145.

[36] Loh, W.-L. (1988). Estimating covariance matrices. Ph.D. thesis. Stanford University.

[37] Loh, W.-L. (1991). Estimating covariance matrices. ***Ann. Statist*** **19** 283-296.

[38] Loh, W.-L. (1991). Estimating covariance matrices II. *J. Multivariate Anal.* **36** 163-174.

[39] Mitra, S. K. (1970). A density-free approach to the matrix variate Beta distribution *Sankhyā. Series A* **32** 81-88.

[40] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*, Wiley, New York.

[41] Muirhead, R. J. (1987). Development in eigenvalue estimation. *Advances in Multivariate Analysis* (A.K. Gupta, ed.), 277-288.

[42] Muirhead, R. J., and Verathaworn, T. (1985). On estimating the latent roots of $\Sigma_1\Sigma_2^{-1}$, *Multivariate Analysis* - VI (P. R. Krishnaiah, ed.), 431-447.

[43] Olkin, I. and Rubin, H. (1964). Multivariate Beta distributions and independence properties of the Wishart distribution. *Ann. Math. Statist.* **35** 261-269.

[44] Olkin, I., and Selliah, J. B. (1977). Estimating covariances in a multivariate normal distribution, *Statistical Decision Theory and Related Topics* II (S. S. Gupta, and D. S. Moore, eds.), 313-326, Academic Press, New York.

[45] Perlman, M.D. (1977). A note on the matrix-variate F distribution. *Sankhyā* **A 39** 290-298.

[46] Perron, F. (1989). Minimax estimators of a covariance matrix. Rapports de recherche du départment de mathématiques et de statistique, **D.M.S. No.** 89-23, Universite de Montreal.

[47] Sharma, D. and Krishnamoorthy, K. (1983). Orthogonal equivariant minimax estimators of bivariate normal covariance matrix and precision matrix. *Calcutta Statist. Assoc. Bull.* **32** 23-45.

[48] Stein, C. (1973). Estimation of the mean of a multivariate normal distribution. In *Proc. Prague Symp. Asymptotic Statist.* 345-381.

[49] Stein, C. (1977). Lectures on multivariate estimation theory. (In Russian.) In: investigation on statistical estimation theory. I. (Zapiski Nauchych Seminarov LOMI im. V. A. Steklova AN SSSR vol. 74). Leningrad, 4-65.

[50] Strawderman, W.E. (1972). Proper Bayes minimax estimators of the multivariate normal mean vector for the case of common unknown variance. *Ann. Statist.* **1** 1189-1194.

[51] Takemura, A. (1984). An orthogonally invariant minimax estimator of the covariance matrix of a multivariate normal population. *Tsukuba J. Math.* **8** 367-376.

[52] Tan, W.Y. (1969). Note on the multivariate and the generalized multivariate Beta distributions. *J. Amer. Statist. Assoc.* **64** 230-241.

[53] Waal, D. J. de (1970). Distributions connected with a multivariate Beta statistic. *Ann. Math. Statist.* **41** 1091-1095.

[54] Zheng, Z. (1986). On estimation of matrix of normal mean. *J. Multivariate Anal.* **18** 70-82.

[55] Zidek, J.V. (1978). Deriving unbiased risk estimators of multinormal mean and regression coefficient estimators using zonal polynomials. *Ann. Statist.* **6** 769-782.