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 $\delta_1, \dots, \delta_p$. 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 $\delta_1, \dots, \delta_p$, and then it is natural to expect that the eigenvalues of 'good' estimate $\hat{\Delta}(U)$ of Δ will perform well as estimates of $\delta_1, \dots, \delta_p$.

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 $k = k_1 + k_2$,

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

$k_i > p+1$, $i = 1, 2$, 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 $\delta_1, \dots, \delta_p$ in terms of the loss functions

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

and

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

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

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

The corresponding risk function is denoted by $\mathbf{R}_i(\Delta, \hat{\Delta}(U)) = \mathbf{E}\left[\mathbf{L}_i(\Delta, \hat{\Delta}(U))\right]$ $(i = 1, 2, 3)$ 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 $\mathbf{E}\left[\operatorname{tr}(\Delta^{-1}\hat{\Delta}(U))\right]$. 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*

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

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

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

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

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

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

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

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*

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

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*

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

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

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

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

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

$$\begin{aligned}\mathrm{E}\left[\,\mathrm{tr}\left(\frac{\partial g(U)}{\partial U}V_{(\frac{1}{2})}\right)\right] =& \frac{1}{2}\mathrm{E}\,[U_{ki}\Delta_{jl} + U_{kj}\Delta_{il} + U_{jl}U_{ki} + U_{il}U_{kj}], \\ \mathrm{E}\,[g(U)\,\mathrm{tr}\,(U^{-1}V)] =& \mathrm{E}\,[U_{ij}\Delta_{kl} + U_{ij}U_{kl}],\end{aligned}$$

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

$$\begin{aligned}(k_2 - p - 1)\mathrm{E}\,[U_{ij}U_{kl}] - \mathrm{E}\,[U_{jl}U_{ki}] - \mathrm{E}\,[U_{il}U_{kj}]& \\ = \frac{k_1}{k_2 - p - 1}\{k_1\Delta_{ij}\Delta_{kl} + \Delta_{jl}\Delta_{ki} + \Delta_{il}\Delta_{kj}\}.&\end{aligned} \tag{3.7}$$

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

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

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

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

Thus $\mathrm{E}[U_{ij}U_{kl}]$ 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)
$$
\begin{aligned}
\mathrm{Cov}(U_{ij}, U_{kl}) = &\frac{k_1(k - p - 1)}{(k_2 - p)(k_2 - p - 1)(k_2 - p - 3)}\left\{\frac{2}{k_2 - p - 2}\Delta_{ij}\Delta_{kl}\right. \\
&\left. + \Delta_{jl}\Delta_{ik} + \Delta_{kj}\Delta_{il}\right\},
\end{aligned}
$$

(ii)
$$
\begin{aligned}
\mathrm{E}[UQU] = &\frac{k_1}{(k_2 - p)(k_2 - p - 1)(k_2 - p - 3)}[\{k_1(k_2 - p - 2) + 2\}\Delta Q\Delta \\
&+ (k - p - 1)\{(\Delta Q\Delta)' + \mathrm{tr}(\Delta Q)\Delta\}],
\end{aligned}
$$

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

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

which is equal to $\mathrm{E}_W[WQW]$, 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*

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

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

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

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

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 $L = \mathrm{diag}\,(l_1, \ldots, l_p)$ and $l_1 > \cdots > l_p > 0$, and $\varphi = \mathrm{diag}\,(\varphi_1(L), \cdots, \varphi_p(L))$.

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

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

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

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

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

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

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 $\mathbf{E}[\,\mathrm{tr}(\Delta^{-1}\hat{\Delta})]$ 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*

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

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

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

and

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

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

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

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*

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

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

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

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 $\operatorname{tr} A_{(r)}B_{(1/r)} = \operatorname{tr} AB$ 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

$$\mathbf{E}\left[\frac{1}{2}t^2(z)-t(z)+\frac{a_1t(z)\operatorname{tr}\Delta^{-1}}{\operatorname{tr}U^{-1}}\right]\leq 0.$$

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 $S=k_2U$ converges to Wishart distribution $W_p(k_1,\Delta)$ weakly as n_2 tends to infinity, the estimators in Theorem 3.4.1 with $t^*(z)=t(k_2z)$ 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*

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

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

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

(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

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

since

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

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

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

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

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

By a combiantoric argument it can be seen that

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

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

(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

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

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 $m \leq p-1$,

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

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

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

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

(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 $d_1 \leq \phi_i \leq d_p$, $i = 1, \ldots, p$.

(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 $\mathrm{E}[\operatorname{tr}(\Delta^{-1}U)] = pk_1/(k_2-p-1)$ 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

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

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

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

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

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

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

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

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

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

Combining (3.43) and (3.44) gives that

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

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

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

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

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

which completes the proof.

REMARK 3.5.1. If $\mathbf{E}[M_1(\hat{\Delta}_A) + M_2(\hat{\Delta}_A)] \leq 0$ for general p, then (3.40) becomes the upper bound (due to Gupta and Krishnamoorthy [16]) for the approximate risk of the estimates $\hat{\Delta}_A$. From this, it is seen that the approximate risk given by Muirhead and Verathaworn[42] neglects the terms $M_1(\hat{\Delta}_A)$ and $M_1(\hat{\Delta}_A)$ as long as we restrict our attention to the class of the estimators of the form (3.35). Furthermore we expect that $\hat{\Delta}_{AM}$ 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* $L = \mathrm{diag}(l_1, l_2)$ *and* $l_1 > l_2$, *and let* $\varphi^{PM}(L) = \mathrm{diag}(\varphi_1^{PM}(L), \varphi_2^{PM}(L))$ *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

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

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

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

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

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

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 $M_2(\hat{\Delta}_{PM})$ as $\mathrm{tr}\,(\Delta^{-1}Hh(L)H')$ where $h(L) = \mathrm{diag}\,(h_1(L), h_2(L))$ and

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

Then we can rewrite

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

as

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

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

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

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) $$\mathbf{E}\left[\frac{tr\Delta^{-2}}{(\operatorname{tr} U^{-1})^2}\right] \le \frac{(k_1-p+1)(k_1-p+3)}{(k_2-2)(k_2-4)} \mathbf{E}\left[\frac{\operatorname{tr} U^{-2}}{(\operatorname{tr} U^{-1})^2}\right],$$

(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) $$\mathbf{E}\left[\frac{\operatorname{tr}(U\Delta^{-2})}{\operatorname{tr}U^{-1}}\right] \leq \frac{2}{(k_2 - p - 1)(k_2 - 2)}\left[\frac{(k_1 + k_2)(k_1 - p + 1) + k_1(k_2 - 2)}{k_2} + \frac{(k_1 - p + 1)(k_1 - p + 3)}{k_2 - 4}\right]\mathbf{E}\left[\frac{\operatorname{tr}U^{-2}}{(\operatorname{tr}U^{-1})^2}\right] + \frac{k_1(k_1 - p - 1)}{k_2(k_2 - p - 1)}.$$

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

$$(k_2 - p - 1)\mathrm{E}\left[\frac{\operatorname{tr}(U\Delta^{-2})}{\operatorname{tr} U^{-1}}\right] = k_1 \mathrm{E}\left[\frac{\operatorname{tr}\Delta^{-1}}{\operatorname{tr} U^{-1}}\right] + 2\mathrm{E}\left[\frac{\operatorname{tr}(\Delta^{-1}U^{-1} + \Delta^{-2})}{(\operatorname{tr} U^{-1})^2}\right]. \tag{3.60}$$

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

$$\begin{aligned}&\left[\frac{4a_2^2}{(k_2-p-1)(k_2-2)}\left\{\frac{(k_1+k_2)(k_1-p+1)+k_1(k_2-2)}{k_2}\right.\right. \\ &\qquad\left.\left.+\frac{(k_1-p+1)(k_1-p+3)}{k_2-4}\right\} - \frac{4a_2}{k_2}\right]\mathrm{E}\left[\frac{\operatorname{tr} U^{-2}}{(\operatorname{tr} U^{-1})^2}\right] \\ &\qquad+\frac{2a_2}{k_2}\left\{\frac{a_2k_1(k_1-p-1)}{k_2-p-1} - (k_1-p-1+2\varepsilon)\right\}.\end{aligned} \tag{3.61}$$

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