CHAPTER 2

ESTIMATING MATRIX OF NORMAL MEAN

Assume that

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

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

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

The risk function corresponding to this loss function is

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

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*

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

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 $\mathrm{E}\,|q_{ij}| < \infty$ 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*

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

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*

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

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

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

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) $$\mathrm{tr}\, \nabla_x' X Q = m\, \mathrm{tr}\, Q + \mathrm{tr}\, X' \nabla_x Q',$$

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

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

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

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

(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 $W = X'X$,

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

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

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

(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

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

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*

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

and

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

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

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

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

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

Now using Lemma 2.1.7 leads to

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

which completes the proof.

Next we state calculus lemmas on eigenstructures for the case $m \leq p$. 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 $m \leq p$. 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

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

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 $m \le p$. Let $R = (R_1, R_2, ..., R_m)$, 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) $$D_{\tilde{F}} R_k = f_k^* R_k, \quad f_k^* = (1/2) \sum_{i \ne k} 1/(f_k - f_i),$$

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

$$\varphi_k^{(1)}(F) = \frac{1}{2} \sum_{i \ne k} \frac{\varphi_k(F) - \varphi_i(F)}{f_k - f_i} + \frac{\partial \varphi_k(F)}{\partial f_k}, \quad k = 1, \ldots, m.$$

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

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

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*

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

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

and

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$-4\ \operatorname{tr}\big(D_{\tilde F}h(F)\big)\cdot \tilde F D_{\tilde F}\{\tilde F D_{\tilde F}h(F)\} + 2(m+1)\ \operatorname{tr}\tilde F\{D_{\tilde F}h(F)\}^2. \tag{2.37}$$

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

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

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

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

Noting that $D_F^2 h(F) = D_F RH(F)R'$ 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

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

Similarly, we can get that

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

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

$$\hat{\mathrm{R}} = \sum_{k=1}^{p} \left\{ 2(m-p+1)h_k + 4y_k h_{kk} + 4 \sum_{i > k} \frac{y_k h_k - y_i h_i}{y_k - y_i} + y_k h_k^2 \right\} + pm,$$

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 $F = (f_1, \ldots, f_m)$, $\operatorname{diag}(F) = R'XS^{-1}X'R$, 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 $h_{kk}^{(1)} = \partial^2 h_k^{(1)}/\partial f_k^2$, $k = 1,\dots,p$. We may observe that

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

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

$$4\sum_{k=1}^{p}\sum_{l>k}\frac{f_k h_k^{(1)} - f_l h_l^{(1)}}{f_k - f_l} = -\sum_{k=1}^{p}\frac{2(p-1)b}{u},$$

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

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

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

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

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

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

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

(c) $\quad \gamma_1(t) \geq \gamma_2(t) \geq \cdots \geq \gamma_p(t)$ 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′) $\quad 0 \leq \gamma_k(t) \leq 2(p - m - 1)/(n + 2m - p + 1)$

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

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

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

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

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 $\gamma^{(i)} = (\gamma_1^{(i)}, \ldots, \gamma_p^{(i)})$ 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*

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

(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 $\gamma = (\gamma_1(y_1), \gamma_2(y_2), \ldots, \gamma_p(y_p))$, $Y = \operatorname{diag}(y_1, y_2, \ldots, y_p)$, *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

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

and the estimator

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

is minimax if

$$\text{(iii)}' \qquad 0 \le \gamma_i(t) \le \frac{m-p-1}{n+p+1}, \qquad i = 1, 2, \ldots, p,$$

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

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

and

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

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 $0 \le d_1 \le (m-p-1)/(n-p+3)$ and $0 \le d_2 \le (m-p-1)/(n+p+1)$ 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

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

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 $T(Y) = \operatorname{diag}(t_1(Y), t_2(Y), \dots, t_p(Y))$ and

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

Using (2.56) it can be seen that

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

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

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

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

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

and

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

It follows from these inequalities and some simplification that

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

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

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

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 information 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.