Robust Empirical Bayes Estimation of the Elliptically Countoured Covariance Matrix

Z. Khodadadi
Islamic Azad University, Marvdasht-Branch

B. Tarami
Yasouj University, College of Sciences

Abstract. Let $S$ be the matrix of residual sum of square in linear model $Y = A\beta + e$, where the matrix of errors is distributed as elliptically contoured with unknown scale matrix $\Sigma$. For Stein loss function, $L_1(\hat{\Sigma}, \Sigma) = \text{tr}(\hat{\Sigma}\Sigma^{-1}) - \log|\hat{\Sigma}\Sigma^{-1}| - p$, and squared loss function, $L_2(\hat{\Sigma}, \Sigma) = \text{tr}(\hat{\Sigma}\Sigma^{-1} - I)^2$, we offer empirical Bayes estimators of $\Sigma$, which dominate any scalar multiple of $S$, i.e., $aS$, by an effective amount. In fact, this study somehow shows that improvement of the empirical Bayes estimators obtained under the normality assumption remains robust under elliptically contoured model.

AMS Subject Classification: 62C12.
Keywords and Phrases: Covariance matrix, elliptically contoured, empirical Bayes estimators.

1. Introduction

The problem of empirical Bayes (EB) estimation with the normal covariance matrix $\Sigma$, was considered by [3], who proved that these estimators dominate all scalar multiples of the unbiased estimator. Our objective is to establish the dominance results by Haff [4] remains robust under the elliptically contoured distribution which we refer to it as ECD in this...
paper. Here, we consider the problem of estimation with the elliptically contoured covariance matrix \( \Sigma \); then we get the empirical Bayes estimators which dominate the usual unbiased estimators for each of two invariant loss functions \( L_1 \) and \( L_2 \). The dominance results under \( L_1 \) and \( L_2 \) were first offered by James and Stein. There are many studies to estimate a covariance matrix for the normality assumption; under loss functions \( L_1 \) and \( L_2 \), see [2,3], [6], [9] and [11], [8], [10]. The identity for the ECD which was derived by [7], known in the literature as the "Stein-Haff identity", is applied to compute risk functions.

Let \( Y \) be an \( N \times p \) random matrix with multivariate linear model

\[
Y = A\beta + e
\]

where \( e \) is an \( N \times p \) matrix of random errors, \( A \) is a known full rank \( N \times m \) matrix and \( \beta \) is an \( m \times p \) matrix of unknown parameters. We assume that the error matrix \( e \) has an elliptical density

\[
|\Sigma|^{-\frac{N}{2}} f(\text{tr}(\Sigma^{-1}e'e))
\]

where \( \Sigma \) is a \( p \times p \) unknown positive-definite matrix, \( f(\cdot) \) is a differentiable and nonnegative real-value function. Here \(|B|\), \(\text{tr}(B)\) and \(B^t\) stand for the determinant, the trace and the transpose of a square matrix \( B \), respectively.

Let \( S \) be the matrix of residual sum of squares, i.e.,

\[
S = Y^t(I_N - A(A'^tA)^{-1}A')Y,
\]

and let \( n = N - m \). Under the elliptically assumption, the expected value for various functions of \( S \) have been derived; by [7].

Let \( \hat{\Sigma} \) be an estimator of \( \Sigma \). We assume that the loss function is

\[
L_1(\hat{\Sigma}, \Sigma) = \text{tr}(\Sigma\Sigma^{-1}) - \log|\Sigma\Sigma^{-1}| - p,
\]

or

\[
L_2(\hat{\Sigma}, \Sigma) = \text{tr}(\hat{\Sigma}\Sigma^{-1} - I)^2,
\]
and define the risk function by

\[ R_i(\hat{\Sigma}, \Sigma) = E[L_i(\hat{\Sigma}, \Sigma)|\Sigma], \quad i = 1, 2. \]

Let \( \hat{\Sigma} \) and \( \hat{\Sigma}_* \) be the competing estimators of \( \Sigma \), where \( \hat{\Sigma} \) dominates \( \hat{\Sigma}_* \) (mod \( L_i \)) if \( R_i(\hat{\Sigma}, \Sigma) \leq R_i(\hat{\Sigma}_*, \Sigma)(\forall \Sigma) \).

According to the same notation used in [7], let \( f(0)(x) = f(x) \),

\[ f(k+1)(x) = \frac{1}{2} \int_x^\infty f(k)(t)dt, \quad k = 0, 1 \]

and

\[ E_{\Sigma}^{(k)}[v(S)] = \int v(S) |\Sigma|^{-\frac{N}{2}} f^{(k)}(\text{tr}(\Sigma^{-1}(y - A\beta)(y - A\beta))) dy \]

where \( v(S) \) is an integrable function of \( S \). Also we use the transformation to polar coordinates to get,

\[ E_{\Sigma}^{(k)}[1] = \gamma(k) = \frac{2\pi^{NP/2}}{\Gamma(NP/2)} \int_0^\infty r^{Np-1} f^{(k)}(r^2)dr, \quad k = 0, 1, 2 \]

and assume that \( \gamma(i) < \infty \), for more details see [7].

Following [3], the empirical Bayes estimators have the form

\[ \hat{\Sigma} = a[S + ut(u)C] \]

[5] where \( t(\cdot) \) is a nonnegative and non-increasing function, \( C \) is an arbitrary positive definite matrix, \( u = (\text{tr}S^{-1})^{-1} \) and

\[ 0 \leq a \leq max\{\frac{1}{n\gamma(1)}, \frac{\gamma(1)}{(n + p + 1)\gamma(2)}\}. \]

Without loss of generality, we assume \( C = I \). It should be noted that for \( t \equiv 0 \), we have the obvious estimators, the scalar multiples of \( S \).
1.1 Synopsis

The consideration of the scalar multiples of $S$ shows that the best estimator (mod $L_1$) is the unbiased estimator

$$\hat{\Sigma}_1 = \frac{1}{n\gamma(1)} S$$

and the best estimator (mod $L_2$) is

$$\hat{\Sigma}_2 = \frac{\gamma(1)}{(n + p + 1)\gamma(2)} S.$$ 

The main result of this paper deals with the EB estimators(2). Furthermore, for each loss function, there are conditions under which they dominate the best scalar multiple of $S$.

1.2 Stein-Haff Identity and Its Application

Let $T(S) = (t_{ij}(S))$ be a $p \times p$ matrix whose elements are functions of $S = (s_{ij})$. Denote

$$\{DsT(S)\}_{ij} = \sum_{a=1}^{p} \frac{1}{2} (1 + \delta_{ia}) \frac{\partial t_{aj}(S)}{\partial s_{ia}}$$

where $\delta_{ia}$ is Kronecker’s delta. From Lemma 1 in [11], for suitable choice of a matrix $T(S)$, the Stein-Haff identity is given by

$$E_{\Sigma}^{(k)} \left[ \text{tr} \left( \Sigma^{-1} T(S) \right) \right] = E_{\Sigma}^{(k+1)} \left[ (n - p - 1) \text{tr} \left( S^{-1} T(S) \right) + 2 \text{tr} DsT(S) \right]$$

and from (3) and (4) for a real valued function $h(S)$, we observe that

$$E_{\Sigma}^{(k)} \left[ \text{tr} \left( \Sigma^{-1} h(S) \right) \right] = E_{\Sigma}^{(k+1)} \left[ (n - p - 1) h(S) \text{tr}(S^{-1}) + \text{tr} \left( \frac{\partial h(S)}{\partial S} \right) \right]$$

First, we apply (4) and (5) to calculate risk function $R_1$. It appears that $\alpha_1(\Sigma) = R_1(\hat{\Sigma}, \Sigma) - R_1(\hat{\Sigma}_1, \Sigma)$ has terms under the unusual expectation of the form $h(S)\text{tr}(\Sigma^{-1})$. Theorem 2.3 gives conditions under which $\alpha_1(\Sigma) \leq 0 \ (\forall \Sigma)$. Since $\alpha_2(\Sigma) = R_2(\hat{\Sigma}, \Sigma) - R_2(\hat{\Sigma}_2, \Sigma)$ has terms which are quadratic in $\Sigma^{-1}$, calculating $R_2$ is more difficult than $R_1$. 
2. Main Results

The following four theorems are the main results of this paper. The proofs are postponed to Section 3.

**Theorem 2.1.** Under the loss function $L_1$, the best estimator of the form $\hat{\Sigma}_a = aS$ is given by $a = \frac{1}{n\gamma(1)}$.

**Theorem 2.2.** Under the loss function $L_2$, the best estimator of the form $\hat{\Sigma}_a = aS$ is given by $a = \frac{\gamma(1)}{(n+p+1)\gamma(2)}$.

Our main result concern the EB estimators (2). For comparing $\hat{\Sigma}$ with $\hat{\Sigma}_i, i = 1, 2$, $a$ is replaced by $\frac{1}{n\gamma(1)}$ and $\frac{\gamma(1)}{(n+p+1)\gamma(2)}$.

**Theorem 2.3.** Let $\hat{\Sigma}$ is given by (2), with

i) $a = \frac{1}{n\gamma(1)}$,

ii) $u = (\text{tr}S^{-1})^{-1}$

iii) $t$ is a constant, $0 \leq t \leq \frac{2(n\gamma(0)+p-n)}{n\gamma(0)}$.

Then $\hat{\Sigma}$ dominates $\hat{\Sigma}_1$ (mod $L_1$), i.e., $R_1(\hat{\Sigma}, \Sigma) \leq R_1(\hat{\Sigma}_1, \Sigma)$ ($\forall \Sigma$).

An optimal value of $t$ is $\frac{\gamma(0)+p-n}{n\gamma(0)}$ (as seen from the proof).

**Theorem 2.4.** Let $\hat{\Sigma}$ is given by (2), with

i) $a = \frac{\gamma(1)}{(n+p+1)\gamma(2)}$,

ii) $u = (\text{tr} S^{-1})^{-1}$

iii) $t$ is a constant, $0 \leq t \leq \frac{2[(n-2p+2)-p(n-1)(n-p)]}{p(n-p-1)(n-p+1)}$.

Then $\hat{\Sigma}$ dominates $\hat{\Sigma}_2$ (mod $L_2$), i.e., $R_2(\hat{\Sigma}, \Sigma) \leq R_2(\hat{\Sigma}_2, \Sigma)$ ($\forall \Sigma$).

The choice $\frac{(n-2p+2)-p(n-1)(n-p)}{p(n-p-1)(n-p+1)}$ of $t$ is the optimal choice.

The latter calculations depends on the following lemmas.
Lemma 2.5. [Hisayuki Tsukuma 2005] Let $Q$ be a $p \times p$ matrix of constants. Under the conditions of Lemma 2.1. in [5] (see appendix), we have

i) $E_{\Sigma}^{(0)}[S] = n\gamma(1)\Sigma,$

ii) $E_{\Sigma}^{(0)}[SQS] = \gamma(2)\left\{n^2\Sigma Q\Sigma + n\Sigma Q'\Sigma + n\tr(Q\Sigma)\Sigma\right\}.$

Lemma 2.6. Let $F$ be a $p \times p$ matrix-valued function of $S = (s_{ij})$ and $\phi(S)$ be a scalar function of $S$. Then, we have

i) $D_S(\phi F) = \frac{1}{2}(\frac{\partial \phi}{\partial S} \cdot F) + \phi D_S F,$

and for a matrix $Q_{p \times p}$ of constants,

ii) $D_S(SQ) = (Q + pQ^t)/2,$

iii) $D_S(S^{-2}Q) = -S^{-3}Q - \left\{\tr(S^{-1})S^{-2}Q + \tr(S^{-2})S^{-1}Q\right\}/2,$

iv) $D_S(S^{-1}Q) = -\left\{S^{-2}Q + \tr(S^{-1})S^{-1}Q\right\}/2.$

Proof. (i) For a matrix $F(S)_{p \times p}$ and a scalar $\phi(S)$, from (3) we have

$$[D_S\phi F]_{ij} = \sum_a \frac{1}{2}(1 + \delta_{ia})\frac{\partial}{\partial s_{ia}}(\phi F)_{aj}$$

$$= \sum_a \frac{1}{2}(1 + \delta_{ia})\left\{\frac{\partial \phi}{\partial s_{ia}} F_{aj} + \frac{\partial (F)_{aj}}{\partial s_{ia}} \cdot \phi\right\}$$

$$= \sum_a \frac{1}{2}(1 + \delta_{ia})\frac{\partial \phi}{\partial s_{ia}} F_{aj}$$

$$+ \phi \sum_a \frac{1}{2}(1 + \delta_{ia}) \frac{\partial (F)_{aj}}{\partial s_{ia}}$$

$$= \frac{1}{2}(\frac{\partial \phi}{\partial S} \cdot F)_{ij} + \phi (D_S F)_{ij}$$

which gives Lemma 3.2 (i). □

The following properties of the operator $D_S$ (see the definition in (3)) are required for computations.
Proof. (ii) For a matrix $Q$ of constant, to derive Lemma 3.2 (ii), we note that

$$[D_SSQ]_{ij} = \frac{1}{2} \sum_{a} (1 + \delta_{ia}) \frac{\partial}{\partial s_{ia}} (SQ)_{aj}$$

$$= \frac{1}{2} \sum_{a} \{\delta_{ia} Q_{aj} + \delta_{aa} Q^t_{ij}\}$$

$$= \frac{1}{2} (IQ)_{ij} + \frac{P}{2} Q^t_{ij}.$$

Proof. (iii) Applying

$$\frac{\partial (S^{-1})_{kl}}{\partial s_{ij}} = \{- (S^{-1})_{kl} (S^{-1})_{jl} - (S^{-1})_{li} (S^{-1})_{jk}\},$$

we obtain

$$[D_SS^{-2}Q]_{ij} = \sum_{a,k} \frac{1}{2} (1 + \delta_{ia}) \frac{\partial}{\partial s_{ia}} [(S^{-1})_{ak} (S^{-1}Q)_{kj}]$$

$$= -\frac{1}{2} \sum_{a,k} (S^{-1})_{ia} (S^{-1})_{ak} (S^{-1}Q)_{kj}$$

$$- \frac{1}{2} \left( \sum_{k} (S^{-1})_{ik} (S^{-1}Q)_{kj} \right) \left( \sum_{a} (S^{-1})_{aa} \right)$$

$$- \frac{1}{2} \left( \sum_{a,k} (S^{-1})_{ak} (S^{-1})_{ki} \right) \left( \sum_{m} (S^{-1})_{am} Q_{mj} \right)$$

$$- \frac{1}{2} \left( \sum_{a,k} (S^{-1})_{ak} (S^{-1})_{ak} \right) \left( \sum_{m} (S^{-1})_{mi} Q_{mj} \right)$$

$$= -\frac{1}{2} (S^{-3}Q)_{ij} - \frac{1}{2} \text{tr}(S^{-1}) (S^{-2}Q)_{ij} - \frac{1}{2} (S^{-3}Q)_{ij}$$

$$- \frac{1}{2} \text{tr}(S^{-2}) (S^{-1}Q)_{ij}.$$

(iv). This is given by the similar way in [4]. □

Proof of Theorem 2.1. The proof is similar to that of Theorem 4.1. in [5]; therefore, we state the outline of the proof only. Let $\hat{\Sigma}_{1k} =$
\[ \frac{1}{n\gamma(1)}(1 + k)\mathbf{S}, \ |k| < 1 \text{ and } \mathbf{S}_1 = \frac{1}{n\gamma(1)}\mathbf{S}. \] We want to show that

\[ R_1(\mathbf{S}_1, \mathbf{S}) - R_1(\mathbf{S}_1, \mathbf{S}) = E_{\Sigma}^{(0)} \left[ \frac{k}{n\gamma(1)} \text{tr} (\mathbf{S}\Sigma^{-1}) - p \log(1+k) \right] \geq 0 \quad (\forall \Sigma). \] (6)

From Lemma 2.5(i), we can see that

\[ R_1(\mathbf{S}_1, \mathbf{S}) - R_1(\mathbf{S}_1, \mathbf{S}) = pk - p \log(1+k). \]

Similar to [5], the inequality (6) holds and the proof is complete. \( \square \)

**Proof of Theorem 2.2.** The risk of a scalar multiple of \( \mathbf{S} \) is

\[ R_2(a\mathbf{S}, \Sigma) = E_{\Sigma}^{(0)} \text{tr}(a\mathbf{S}\Sigma^{-1} - \mathbf{I})^2 = \]

\[ a^2 E_{\Sigma}^{(0)} \text{tr}(\mathbf{S}\Sigma^{-1}\mathbf{S}\Sigma^{-1}) - 2aE_{\Sigma}^{(0)} \text{tr}(\mathbf{S}\Sigma^{-1}) + pE_{\Sigma}^{(0)}[1]. \]

Using (1) and Lemma 2.5 (i) and (ii), we have

\[ R_2(a\mathbf{S}, \Sigma) = \gamma(2)a^2p(n^2 + nnp) - 2\gamma(1)nap + p\gamma(0) \]

and the last equality is minimized at \( a = \frac{\gamma(1)}{(n+p+1)\gamma(2)}. \) \( \square \)

**Proof of Theorem 2.3.** Write (2) as

\[ \mathbf{\hat{S}} = a\mathbf{S} + g(\mathbf{S})\mathbf{I} \quad (7) \]

where \( g(\mathbf{S}) = aut(u). \) Taking differentiating of \( u \) with respect to \( \mathbf{S} \) gives

\[ \frac{\partial u}{\partial \mathbf{S}} = -(\text{tr}\mathbf{S}^{-1})^{-2} \frac{\partial (\text{tr}\mathbf{S}^{-1})}{\partial \mathbf{S}} = u^2 \mathbf{S}^{-2}. \]

As a result, we obtain

\[ \frac{\partial g(\mathbf{S})}{\partial \mathbf{S}} = a \left[ \frac{\partial u}{\partial \mathbf{S}} t(u) + u \frac{\partial t(u)}{\partial \mathbf{S}} \right] = a [u^2 t(u) + \frac{\partial t(u)}{\partial u} u^3] \mathbf{S}^{-2}. \] (8)

Let \( t(u) \) be a constant function, the equation (8) implies that

\[ \frac{\partial g(\mathbf{S})}{\partial \mathbf{S}} = au^2 t \mathbf{S}^{-2}. \] (9)
From [5] we have the
\[ \rho_k = \text{tr} \left( \frac{S^{-1}}{\text{tr} S^{-1}} \right)^k, \quad k = 1, 2, \ldots. \] (10)

It is convenient to note that \( \rho_k \) decreases in \( k \) and \( 0 \leq \rho_k \leq 1 \).

Use (7) with \( a = \frac{1}{n\gamma(1)} \) and
\[ \alpha_1 = R_1(\hat{\Sigma}, \Sigma) - R_1(\hat{\Sigma}_1, \Sigma) = E^{(0)}_{\Sigma} g(S) \text{tr} \Sigma^{-1} - \log |I + n\gamma(1)g(S)S^{-1}|. \]

Knowing that \( t(u) \) is a constant, say \( t \), and from (5) with \( h = g \), and applying the expansion
\[ \log |I + \alpha B| = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \alpha^n \text{tr}(B^n) \]
where \( \alpha \) is a real number and \( B_{p \times p} \) a symmetric matrix, we obtain
\[ \alpha_1 = E^{(1)}_{\Sigma} \left[ (n-p-1)g(S)\text{tr}(S^{-1}) + \text{tr} \left( \frac{\partial g(S)}{\partial S} \right) \right] - E^{(0)}_{\Sigma} \left[ \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{i} t^i \rho_i \right]. \]

A simple calculation shows that
\[ \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{i} t^i \rho_i \geq t - \frac{1}{2} t^2, \]
because the terms of the above series are in decreasing magnitude. Substituting (9) into (11), with \( a = \frac{1}{n\gamma(1)} \), we obtain
\[ \alpha_1 \leq \frac{1}{n\gamma(1)} E^{(1)}_{\Sigma} \left[ (n-p-1)t + \rho_2 t \right] + E^{(0)}_{\Sigma} \left[ \frac{1}{2} t^2 - t \right], \]
\[ \leq \frac{1}{n\gamma(1)} (n-p) t E^{(1)}_{\Sigma}[1] + (\frac{1}{2} t^2 - t) E^{(0)}_{\Sigma}[1], \quad \text{(since } \rho_2 \leq 1), \]
\[ \leq \left[ \frac{n-p}{n} - \gamma(0) \right] t + \frac{\gamma(0)}{2} t^2 \quad \text{(from (1)).} \] (12)
A sufficient condition for $\alpha_1(\Sigma) \leq 0 \ (\forall \Sigma)$ is

\[
\left[ \frac{n-p}{n} - \gamma(0) \right] t + \frac{\gamma(0)}{2} t^2 \leq 0.
\]

Finally, it is seen that

\[
0 \leq t \leq \frac{2(n\gamma(0) + p - n)}{n\gamma(0)}
\]

It is also seen, from (12) that (11) is bounded above by

\[
\left[ \frac{n-p}{n} - \gamma(0) \right] t + \frac{\gamma(0)}{2} t^2 ;
\]

hence, the optimal $t$ is $\frac{n\gamma(0) + p - n}{n\gamma(0)}$. \(\square\)

Theorem 2.4. compares $\hat{\Sigma}$ in (7) and $\hat{\Sigma}_2 = aS$, $a = \frac{\gamma(1)}{(n+p+1)\gamma(2)}$. Let

\[
\alpha_2(\Sigma) = R_2(\hat{\Sigma}, \Sigma) - R_2(\hat{\Sigma}_2, \Sigma)
\]

\[
= E_\Sigma^{(0)} \left[ 2a g(S) \text{tr}(S\Sigma^{-2}) \right] - 2g(S)\text{tr}(\Sigma^{-1} + g^2(S)\text{tr}(\Sigma^{-2})
\]

Then identities for the terms in (13) are given in the following Lemma.

**Lemma 2.7.** Under conditions of Lemma 5.2,[5](see appendix)

i) For $h(S) = 2a g(S)$, we have

\[
E_\Sigma^{(0)}[h(S) \text{tr}(\Sigma^{-2}S)] = E_\Sigma^{(2)} \left[ n(n-p-1)h(S) \text{tr}(S^{-1}) 
\right.
\]

\[
- (p+1)\text{tr}\left( \frac{\partial h(S)}{\partial S} \right)
\]

\[
+ 2\text{tr} D_S \left( \frac{\partial h(S)}{\partial S} \cdot S \right)
\]

ii) For $h(S) = -2g(S)$, we obtain

\[
E_\Sigma^{(0)}[h(S) \text{tr}(\Sigma^{-1})] = E_\Sigma^{(1)} \left[ (n-p-1)h(S) \text{tr}(S^{-1}) + \text{tr}(\frac{\partial h(S)}{\partial S}) \right]
\]
iii) For $h(S) = g^2(S)$, we get

$$E_{\Sigma}^{(0)}[h(S) \text{ tr}(\Sigma^{-2})] = E_{\Sigma}^{(2)}[2 \text{ tr}D_S\left(\frac{\partial h(S)}{\partial S}\right) + 2(n - p - 1)$$

$$\text{tr}(S^{-1} \frac{\partial h(S)}{\partial S}) + (n - p - 1)$$

$$(n - p - 2)h(S) \text{ tr}(S^{-2})$$

$$-(n - p - 1)h(S) \text{ tr}(S^{-1})$$

Proof (i). Set $T = S\Sigma^{-1}h(S)$, by using (4), we can see

$$E_{\Sigma}^{(0)}\left[\text{tr}(\Sigma^{-1}S\Sigma^{-1}h(S))\right] =$$

$$E_{\Sigma}^{(1)}\left[(n - p - 1) \text{ tr}(h(S) \Sigma^{-1}) + 2 \text{ tr}D_S(h(S) S\Sigma^{-1})\right].$$

Now applying Lemma 2.6 (i) and (ii) with $Q = \Sigma^{-1}$ to the second term $D_S(S\Sigma^{-1}h(S))$; hence, we get

$$E_{\Sigma}^{(0)}[h(S) \text{ tr}(\Sigma^{-2}S)] = E_{\Sigma}^{(1)}\left[n \text{ tr}(\Sigma^{-1}h(S)) + \text{ tr}(\Sigma^{-1} \frac{\partial h(S)}{\partial S}S)\right].$$

The result obtains by applying (5) to the first term and then (4) to the second.

(ii). In (5), set $k = 0$.

(iii). Put $T = \Sigma^{-1}h(S)$. From (4) and Lemma 2.6. (i), we have

$$E_{\Sigma}^{(0)}\left[\text{tr}(\Sigma^{-2}h(S))\right] = E_{\Sigma}^{(1)}\left[(n - p - 1) \text{ tr}(\Sigma^{-1}S^{-1}h(S)) + \text{ tr}(\frac{\partial h(S)}{\partial S}S^{-1})\right].$$

Applying (4) to the first term with $T = S^{-1}h(S)$ and applying it to the second with $T = \frac{\partial h(S)}{\partial S}$, we obtain

$$E_{\Sigma}^{(0)}\left[\text{tr}(\Sigma^{-2}h(S))\right] = E_{\Sigma}^{(2)}\left[(n - p - 1)^2h(S)\text{ tr}(S^{-2})$$

$$+ 2(n - p - 1) \text{ tr}D_S(S^{-1}h(S))$$

$$+ (n - p - 1) \text{ tr}(S^{-1} \frac{\partial h(S)}{\partial S})$$

$$2\text{ tr}D_{\frac{\partial h(S)}{\partial S}}\right].$$
Our result yields by using Lemma 2.6. (i) with $F = S^{-1}$ and $\phi = h$ and Lemma 2.6. (iv) with $Q = I$ to the second term.

The special of Lemma 2.7. by taking $g(S) = aut$, $u = (\text{tr}(S^{-1}))^{-1}$ and $a = \frac{\gamma(1)}{(n+p+1)\gamma(2)}$ is the following lemma.

\textbf{Lemma 2.8.} For $g(S) = aut$,

\begin{enumerate}[i)]
\item $E_{\Sigma}\left[2a^2 ut \text{tr}(\Sigma^{-2}S)\right] = E_{\Sigma}^{(2)}\left[2n a^2(n - p - 1)t - 2a^2 \rho_2(p + 2)t + 4a^2 t \rho_3 - 2a^2 t\right]$, 
\item $E_{\Sigma}^{(0)}\left[-2aut \text{tr}(\Sigma^{-1})\right] = E_{\Sigma}^{(1)}\left[-2a(n - p - 1)t - 2at \rho_2\right]$, 
\item $E_{\Sigma}^{(0)}\left[a^2 u^2 t^2 \text{tr}(\Sigma^{-2})\right] = E_{\Sigma}^{(2)}\left\{- (n-p-1)a^2 t^2 + 6a^2 \rho_4 t^2 + 4a^2 t^2 \rho_3 (n-p-2) + a^2 t \rho_2 ((n - p - 1)^2 t - (n - p - 1)t - 4t)\right\}$
\end{enumerate}

\textbf{Proof (i).} In Lemma 2.7. (i), we put $h(S) = 2a^2 ut$ and then compute $D_S \left(\frac{\partial h(S)}{\partial S} S\right)$. By applying Lemma 2.6. (i), (iv) with $\phi = u^2$, $F = S^{-1}$ and $Q = I$, we have

\[
\text{tr} \left( D_S \left(\frac{\partial h(S)}{\partial S} S\right) \cdot S \right) = \text{tr} \left\{ 2a^2 t \left[ \frac{1}{2} \frac{\partial^2}{\partial S^2} S^{-1} + u^2 \left\{ -\frac{1}{2} S^{-2} - \frac{1}{2} \text{tr}(S^{-1}) S^{-1} \right\}\right]\right\} 
\]

\[
= \text{tr} \left\{ 2a^2 t u^3 S^{-3} - a^2 t u^2 S^{-2} - a^2 t u^2 \text{tr}(S^{-1}) S^{-1} \right\} 
\]

\[
= 2a^2 t \rho_3 - a^2 t \rho_2 - a^2 t, 
\]

Result (i) is now obtainable by using (9).

(ii). Using (9) and (10), the result obtains from Lemma 2.7. (ii) with $h(S) = -2aut$.

(iii). Take $h(S) = a^2 u^2 t^2$ in Lemma 2.7. (iii); then compute $D_S \left(\frac{\partial h(S)}{\partial S}\right)$. Similarly, using Lemma 2.6. (i), (iii) with $\phi = u^3$, $F = S^{-2}$
and $Q = I$, we have

$$
\text{tr}D_S\left(\frac{\partial h(S)}{\partial S}\right) = 2a^2t^2 \text{tr}\left[\frac{1}{2}\left(\frac{\partial u^3}{\partial S}S^{-2}\right) + u^3D_SS^{-2}\right]
= a^2t^2[3\rho_4 - 2\rho_3 - 2\rho_2].
$$

(14)

Now substitute (14) in Lemma 2.7. (iii) to achieve the result.

□

**Proof of Theorem 2.4.** Each term in (13) obtained from Lemma 2.7 (i), (ii) and (iii); hence, we have

$$
\alpha_2 = E^{(2)}_\Sigma\left[a^2\{2n(n - p - 1) + 2\rho_2(p + 2) + 4\rho_3 - 2\}\right]
+ a^2t^2\left\{-2(n - p - 1) + 6\rho_4 + 4\rho_3(n - p - 2) + \rho_2(n - p - 1)^2\right\}
- \rho_2(n - p - 1) - 4\rho_2\right] + E^{(1)}_\Sigma\left[at\{-2(n - p - 1) - 2\rho_2\}\right].
$$

The above coefficient of $a^2t$ can be written as $2[n(n - p - 1) - 1] + 2\rho_2(p + 2) + 4\rho_3$. Since $\rho_k \leq 1$ thus, the entire quantity is bounded above by

$$
2[n(n - p - 1) - 1] + 2[p + 2] + 4 = 2[(n - 1) - p + 3].
$$

(16)

Since $\rho_k \searrow 0 \leq \rho_k \leq 1$ and $n > p + 1$, the coefficient of $a^2t^2$ is bounded above by

$$
\rho_2\left[2(n - p - 1) + (n - p - 1)^2 - 2\right]
$$

and the above term is bounded by

$$
\rho_2\left[2(n - p - 1) + (n - p - 1)^2\right].
$$

The above coefficient of $\rho_2$ is positive and $\rho_2 \leq 1$; accordingly, the coefficient of $a^2t^2$ is bounded above by

$$
(n - p - 1)(n - p + 1).
$$

(17)

From [1], paper 137, we have $\frac{1}{p} \leq \rho_2 \leq 1$ then in (15), for the coefficient of $at$ we can see

$$
-2(n - p - 1) - 2\rho_2 \leq -\frac{2}{p}.
$$

(18)
Finally, from (15)-(18) a sufficient condition for $\alpha_2(\Sigma) \leq 0$ $\forall \Sigma$ is

$$2a^2t[(n-1)(n-p)+3] \gamma(2) + a^2t^2(n-p-1)(n-p+1) \gamma(2) + at(-\frac{2}{p}) \gamma(1) \leq 0.$$ (19)

Recall that $a = \frac{\gamma(1)}{\gamma(2)(n+p+1)}$ then (19) is equivalent to

$$0 \leq t \leq \frac{2[(n-2p+1) - p(n-1)(n-p)]}{p(n-p-1)(n-p+1)}.$$  

The left hand side of (19) is minimized at

$$t = \frac{(n-2p+1) - p(n-1)(n-p)}{p(n-p-1)(n-p+1)}.$$ □

Appendix

**Lemma 2.1.** Let $Q$ be a $p \times p$ matrix of constants. If

$$\lim_{z \to \pm \infty} |z| f(1)(z^2 + a^2) = 0$$

for any real $a$, then we have

(i) $E^{(0)} \Sigma [S] = n \gamma(1) \Sigma$.

If $\lim_{z \to \pm \infty} |z|^3 f(1)(z^2 + a^2) = 0$ and $\lim_{z \to \pm \infty} |z| f(2)(z^2 + a^2) = 0$

for any real $a$, then we have

(ii) $E^{(0)} \Sigma [SQS] = \gamma(2)(n^2 \Sigma Q \Sigma + n \Sigma Q \Sigma + ntr(Q \Sigma) \Sigma)$,

(iii) $E^{(0)} \Sigma [tr(QS)S] = \gamma(2)(n^2 tr(Q \Sigma) \Sigma + n \Sigma Q \Sigma + n \Sigma Q \Sigma)$.

**Lemma 5.2.** Let $G(S) = (g_{ab}(S)) = G(\sum_{c=1}^n \hat{z}_c z_c)$, be a $p \times p$ matrix whose elements are differentiable with respect to $z_{jk}$ $(j = 1, 2, ..., n, k = 1, 2, ..., p)$. For $-2 \leq i \leq 1$, assume that

(a) $E^{(i)} \Sigma [||g_{ab}(S)||] < \infty$;

(b) $\lim_{z_{jk} \to \pm \infty} |z_{jk}| G(\sum_{c=1}^n \hat{z}_c z_c)(\sum_{c=1}^n \hat{z}_c z_c)^{-1} f^{i+1}(z_{jk} + a^2) = 0_{p \times p}$

for any real $a$.

Then we have

$$E^{(i)} \Sigma [\Sigma^{-1} G(S)] = E^{(i+1)} \Sigma [(n-p-1)S^{-1} G(S) + 2D_a G(S)].$$
References


Zahra Khodadadi
Department of Mathematics
Assistant Professor of Mathematics
Islamic Azad University-Marvdasht Branch
Marvdasht, Iran.
E-mail: zhr_khodadady@yahoo.com

Bahram Tarami
Department of Mathematics
Assistant Professor of Mathematics
College of Sciences, Yasouj University
Yasouj, Iran.
E-mail: tarami@mail.yu.ac.ir