ORTHOGONAL EQUIVARIANT MINIMAX ESTIMATORS OF BIVARIATE NORMAL COVARIANCE MATRIX AND PRECISION MATRIX

DIVAKAR SHARMA AND K. KRISHNAMOORTHY

Department of Mathematics,
Indian Institute of Technology, Kanpur.

ABSTRACT : The scale and orthogonal equivariant minimax estimators are obtained for the bivariate normal covariance matrix and precision matrix under Selliah's (1964) and Stein's (1961) loss functions. These new estimators are better than Selliah's and Stein's minimax estimators. An unbiased estimator of the risk of the new estimator is obtained under Selliah's loss function using Haff's (1979) identity for the Wishart distribution. Simulation results seem to indicate that the new estimators dominate the corresponding Haff's [(1979), (1980)] estimators. We also prove that, for $p=2$, Haff's estimators are not minimax.

1. Introduction

Let $S \sim W_p(\Sigma, n)$, where Σ is unknown and $n>p+1$. The problems considered here are, estimating Σ under the loss functions

$$L_1(\Sigma, \hat{\Sigma})=\text{tr}\,(\hat{\Sigma}\,\Sigma^{-1}-I)^2 \qquad \dots (1.1)$$

$$L_2(\Sigma, \hat{\Sigma})=\text{tr}\,(\hat{\Sigma}\Sigma^{-1})-\log|\hat{\Sigma}\,\Sigma^{-1}|-p \qquad \dots (1.2)$$

and Σ^{-1} under the loss functions

$$L^{(1)}(\Sigma^{-1}, \hat{\Sigma}^{-1})=\text{tr}\,(\hat{\Sigma}^{-1}\,\Sigma-I)^2 \qquad \dots (1.3)$$

$$L^{(2)}(\Sigma^{-1}, \hat{\Sigma}^{-1})=\text{tr}\,(\hat{\Sigma}^{-1}\,\Sigma)-\log|\hat{\Sigma}^{-1}\,\Sigma|-p. \qquad \dots (1.4)$$

The usual estimator for Σ is $\frac{S}{n}$. Selliah (1964) considered the loss function L_1 and proved the inadmissibility of any constant multiple of S, in particular $\frac{S}{n}$, by obtaining a better estimator. Selliah's estimator is equivariant under the group of lower triangular matrices and though minimax suffers from the fact that certain elements of Σ are highly over estimated while some others are highly under estimated. This is, as pointed out by Eaton (1970) because "the standard orthonormal basis in $\mathbb{R}^p$ plays a vital role under the action of elements of G_A^{*}. A way out is to give equal importance to all orthonormal bases. In otherwords we will like to have a minimax estimator which is also orthogonal equivariant. In section 2 we derive such an estimator for $p=2$, which is also scale equivariant and better than Selliah's estimator.

We also derive, for $p=2$, a scale and orthogonal equivariant estimator of Σ for the loss L_2, which is better than Stein's (1961) minimax estimator.

We obtain minimax estimator of Σ^{-1} for $L^{(1)}$ and $L^{(2)}$ similar to those of Selliah and Stein for Σ, and show that no constant multiple of S^{-1} is minimax for the losses $L^{(1)}$ and $L^{(2)}$. We then find scale and orthogonal equivariant estimators better than these minimax estimators.

The estimators considered in this paper are of the form

$$aS+b\,(\,|S|^{1/2}/\text{tr}\,(S^{1/2}))S^{1/2} \text{ for } \Sigma \text{ and} \qquad \ldots \quad (1.5)$$

$$aS^{-1}+(b/\text{tr}\,(S^{1/2}))S^{-1/2} \text{ for } \Sigma^{-1}\ (a>0,\ b\geqslant 0). \qquad \ldots \quad (1.6)$$

An unbiased estimator of the risk of (1.5) under L_1 is obtained by using the identity derived by Haff (1979). Using this unbiased estimator, we prove that among the class of estimators (1.5), the estimator with $a=(n^2+n+2)/(n^3+5n^2+6n+4)$ and $b=2n/(n^3+5n^2+6n+4)$ is the unique minimax estimator. Simulation results seem to indicate that our estimators of Σ and Σ^{-1} for the losses L_1, L_2 and $L^{(1)}$ respectively are better than the corresponding estimators of Haff (1980) and Haff (1979). We also prove that Haff's estimators of Σ are not minimax for L_1 and L_2.

2. The scale and orthogonal equivariant minimax estimator for L_1

Selliah (1964) considered the group G_A of lower triangular matrices A acting on $\{S : S>0\}$ as $S\to ASA'$. This induces the transformation $\Sigma\to A\ \Sigma\ A'$ and with the loss function L_1 mentioned above and $\hat{\Sigma}(S)\to A\hat{\Sigma}(S)A'$, we see the problem remains invariant. It is easy to verify that the equivariant estimator is of the form $TD_\delta T'$ where T is a lower triangular matrix satisfying $TT'=S$, and D_δ is a diagonal matrix. As the group of lower triangular matrices acts transitively on the parameter space, the risk function is a constant and the best choice $D_\delta{}^*$ has its diagonal elements satisfying

$$(n+p-1)(n+p+1)\delta_1^*+(n+p-3)\delta_2^*+\ \dots\ +(n-p+1)\delta_p^*=n+p-1$$
$$(n+p-3)\delta_1^*+(n+p-3)(n+p-1)\delta_2^*+\ \dots\ +(n-p+1)\delta_p^*=n+p-3$$
$$\vdots$$
$$(n-p+1)\delta_1^*+(n-p+1)\delta_2^*+\ \dots\ +(n-p+1)(n-p+3)\delta_p^*=n-p+1.$$

TD_δ^*T', being different from the lower triangular equivariant estimator S/n, is better than the latter and from Kiefer's (1957) result is also minimax. For the special case of $p=2$,

$$\delta_1^*=\frac{n^2+n+2}{n^3+5n^2+6n+4} \quad \text{and} \quad \delta_2^*=\frac{n^2+3n+2}{n^3+5n^2+6n+4}$$

and the minimax risk is

$$\frac{2(3n^2+5n+4)}{n^3+5n^2+6n+4}.$$

To derive an orthogonal equivariant estimator, we state below a Lemma which, though perhaps well known, has not been stated explicitly in literature.

Lemma 2.1. *Let L be a loss function satisfying $L(\Sigma, \hat{\Sigma}) \geqslant 0$, $L(\Sigma, \hat{\Sigma})=L(B\Sigma B', B\hat{\Sigma}B')$ for all nonsingular B. Then, if $\varphi(S)$ is minimax estimator with constant risk so is $\varphi_B(S)=B^{-1}\varphi(BSB')B'^{-1}$ for any nonsingular B.*

Proof. Risk $R(\Sigma, \varphi_B(S))$ of $\varphi_B(S)$ at Σ is

$$E_\Sigma L(\Sigma, B^{-1}\varphi(BSB')B'^{-1})=E_{B\Sigma B'}L(\Sigma, B^{-1}\varphi(S)B'^{-1})$$
$$=E_{B\Sigma B'}\, L(B\Sigma B', \varphi(S))=E_\Sigma L(\Sigma, \varphi(S))=R(\Sigma, \varphi(S)),$$

which proves the assertion.

We notice that the loss function L_1 satisfies the conditions stated in Lemma 2.1 and so, in particular.

$$\varphi_{1\Gamma'}(S)=\Gamma(T_\Gamma D_\delta^* T_\Gamma')\Gamma' \text{ where } T_\Gamma T_\Gamma'=\Gamma' S\Gamma, \qquad \dots \quad (2.1)$$

is a constant risk minimax estimator of Σ for any orthogonal Γ. We can construct from the estimators $\varphi_{1\Gamma'}(S)$ an orthogonal and scale equivariant minimax estimator $\psi_1(S)$ which is better than any $\varphi_{1\Gamma'}(S)$, making use of the following lemma.

Lemma 2.2. *Let v be the invariant probability measure over the orthogonal group $O(p)$ and define*

$$\psi_1(S)=\int_{O(p)}\varphi_{1\Gamma'}(S)d\nu(\Gamma). \qquad \ldots (2.2)$$

Then, $\psi_1(S)$ is an orthogonal and scale equivariant estimator, is minimax and better than any $\varphi_{1\Gamma'}(S)$. In particular, it is better than $TD_\delta^* T'$.

Proof. As shown in Sharma (1980), $L_1(\Sigma, \hat{\Sigma})$ is a strictly convex function of $\hat{\Sigma}$. Jensen's inequality implies that

$$L_1(\Sigma, \psi_1(S)) \leqslant \int_{O(p)} L_1(\Sigma, \varphi_{1\Gamma'}(S))d\nu(\Gamma), \qquad \ldots (2.3)$$

with strict inequality holding with a positive probability under the distribution of S. Taking expectation in (2.3) we get,

$$R_1(\Sigma, \psi_1(S)) < R_1(\Sigma, \varphi_{1\Gamma'}(S))$$

for any orthogonal Γ. Thus, $\psi_1(S)$ is better than $\varphi_{1\Gamma'}(S)$ and so is also minimax.

To prove the orthogonal equivariance, we notice that for any orthogonal matrix η,

$$\psi_1(\eta S\eta')=\int_{O(p)} \varphi_{1\Gamma'}(\eta S\eta')d\nu(\Gamma)$$

$$=\int_{O(p)} \eta'\varphi_{1(\eta\Gamma)'}(S)\eta d\nu(\eta\Gamma)$$

$$=\int_{O(p)} \eta'\varphi_{1(\eta\Gamma)'}(S)\eta d\nu(\eta\Gamma) \text{ (since, } \nu \text{ is invariant)}$$

$$=\eta'\psi_1(S)\eta'.$$

$\psi_1(S)$ is also clearly scale equivariant.

Derivation of $\psi_1(S)$ for a general p is difficult, we obtain $\psi_1(S)$ for $p=2$. Let S have the spectral decomposition $\alpha' D_\lambda \alpha$ where α is orthogonal matrix and $D_\lambda=\text{diag}\ (\lambda_1, \lambda_2)$ with $\lambda_1 \geqslant \lambda_2$. As $\psi_1(S)$ is orthogonal and scale equivariant

$$\psi_1(S)=\lambda_1\alpha'\psi_1\left(\begin{pmatrix}1 & 0\\ 0 & c\end{pmatrix}\right)\alpha \text{ where } c=\lambda_2/\lambda_1 \qquad \ldots (2.4)$$

and so it suffices to find $\psi\left(\begin{pmatrix}1 & 0\\ 0 & c\end{pmatrix}\right)$. Straight-forward calculation leads us to

$$\varphi_{1\Gamma'}(S)=\delta_1^*\begin{pmatrix}1 & 0\\ 0 & c\end{pmatrix}+(\delta_2^*-\delta_1^*)\ \frac{c}{1+(c-1)\gamma_{21}^2}\begin{pmatrix}\gamma_{12}^2 & \gamma_{12}\gamma_{22}\\ \gamma_{12}\gamma_{22} & \gamma_{22}^2\end{pmatrix} \qquad \ldots (2.5)$$

where γ_{ij} is the (i, j) element of Γ. To evaluate $\psi_1(S)$ we must find the integrals

$$\int_{O(2)} \frac{\gamma_{12}\gamma_{22}}{1+(c-1)\gamma_{21}^2}\, d\nu(\Gamma) \text{ and } \int_{O(2)} \frac{\gamma_{i2}^2}{1+(c-1)\gamma_{21}^2}\, d\nu(\Gamma),\ i=1, 2. \qquad \dots (2.6)$$

Let us represent an orthogonal matrix by

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \text{ with } -\pi < \theta < \pi,$$

and take the uniform distribution for θ. This amounts to the invariant probability measure over 0(2), (for example, see Lehmann (1959), p. 335). The values of the integrals in (2.6) are seen to be

0, $(c-1)^{-1}(1-c^{-1/2})$ and $c^{-1/2}-(c-1)^{-1}(1-c^{-1/2})$ respectively.

After some simplification, we get

$$\psi_1\left(\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}\right)=\frac{1}{1+\sqrt{c}}\begin{pmatrix} \delta_1^*+\delta_2^*\sqrt{c} & 0 \\ 0 & c(\delta_1\sqrt{c}+\delta_2) \end{pmatrix}. \qquad \dots (2.7)$$

Substituting the values of δ_1^* and δ_2^* and making use of the relation (2.4), $\psi_1(S)$ seen to be

$$a_0 S+b_0 \frac{|S|^{1/2}}{\operatorname{tr}(S^{1/2})} S^{1/2}, \qquad \dots (2.8)$$

where $a_0=\dfrac{n^2+n+2}{n^3+5n^2+6n+4}$ and $b_0=\dfrac{2n}{n^3+5n^2+6n+4}$.

3. An unbiased estimator and the upper bound of the risk of $aS+b(|S|^{1/2}/\operatorname{tr}(S^{1/2}))S^{1/2}$ under the loss L_1

Consider the class of estimators

$$\psi(S)=aS+b(|S|^{1/2}/\operatorname{tr}(S^{1/2}))S^{1/2} (a > 0, b \geqslant 0),$$

for $p=2$. Now, our interest is to find if there is any other estimator in this class $\psi(S)$ better than $\psi_1(S)$, that is, to find optimal values of a and b which minimizes $\sup_{\Sigma} R_1(\Sigma, \psi(S))$. As evaluating $R_1(\Sigma, \psi(S))$ is complicated, we obtain an unbiased estimator of $R_1(\Sigma, \psi(S))$ using Haff's (1979) identity for the Wishart distribution. From this unbiased estimator, we obtain an upper bound of $R_1(\Sigma, \psi(S))$ which is a function of a and b only. Also, we show that $\psi_1(S)$ is the unique minimax estimator among the class of estimators $\psi(S)$.

Haff's identity is stated as follows :

Theorem 3.1. *For a matrix* $T_{p\times p}=(T_{ij}(S))$ *and a scalar* $h(S)$ *assume that*

(i) *the function* $T_{ij}(S)$ and $h(S)$ *satisfy the conditions of the* Stoke's *theorem on all regions*

$$\mathscr{R}=R(\rho_1, \rho_2)=\{S : S \geqslant 0, 0 < \rho_1 < \| S \| < \rho_2\},$$

(ii) *on* $b_1(\mathscr{R})=\{S : S\geqslant 0, \| S \|=\rho_1\}$

$$\sup_{S\varepsilon b_1(\mathscr{R})} | h(S) | \; \| T \|=o\,(\rho_1^{1-np/2}) \text{ as } \rho_1\to 0+,$$

(iii) *on* $b_2(\mathscr{R})=\{S : S\geqslant 0, \| S \|=\rho_2\}$

$$\sup_{S\varepsilon b_2(\mathscr{R})} | h(S) | \; \| T \|=o[\rho_2^{1-np/2}e^{m\rho_2}] \text{ as } \rho_2\to\infty$$

for arbitrary $m>0$.

Then, we have

$$E(h(S)tr\, T\,\Sigma^{-1})=2E(h(S)D^*(T_{(1/2)}))+2E\left(tr\,\frac{\partial h}{\partial S}\,T_{(1/2)}\right)$$
$$+(n-p-1)\,E\,(h(S)\,tr\,S^{-1}T) \qquad \dots \quad (3.1)$$

where, for $M=(m_{ij})_{p\times p}$ *matrix,* $\|M\|$ *is defined as* $\|M\|=(\Sigma\Sigma\, m_{ij}^2)^{1/2}$, $D^*(M)=\Sigma\Sigma\frac{\partial m_{ij}}{\partial s_{ij}}$ *and*

$$M_{(t)}=(m'_{ij})_{p\times p},\quad m'_{ij}=m_{ij},\; i=j,\quad m'_{ij}=tm_{ij},\; i\neq j. \qquad \dots \quad (3.2)$$

The risk of $\psi(S)$ under the loss L_1 is

$$R_1(\Sigma, \psi(S))=E\text{tr}[a\, S\Sigma^{-1}-I+b(|S|^{1/2}/\text{tr}(S^{1/2}))S^{1/2}\Sigma^{-1}]^2$$
$$=R_1^*(\Sigma, aS)+[b^2E(|S|/\text{tr}(S^{1/2}))^2\text{tr}(S^{1/2}\Sigma^{-1}S^{1/2}\Sigma^{-1})$$
$$+2abE(|S|^{1/2}/\text{tr}(S^{1/2}))\text{tr}(S\Sigma^{-1}S^{1/2}\Sigma^{-1})$$
$$-2bE(|S|^{1/2}/\text{tr}(S^{1/2}))\text{tr}(S^{1/2}\Sigma^{-1})]. \qquad \dots \quad (3.3)$$

The following results are needed for applying the identity (3.1) to (3.3)

(i) $D^*(S^{1/2}_{(1/2)})=\frac{\mathrm{tr}(S^{1/2})}{2|S|^{1/2}}+1/\mathrm{tr}(S^{1/2})$

(ii) $D^*((S^{1/2}\Sigma^{-1}S^{1/2})_{(1/2)})=\mathrm{tr}(\Sigma^{-1})+\mathrm{tr}(S^{1/2}\Sigma^{-1})/\mathrm{tr}(S^{1/2})$

(iii) $D^*((S\Sigma^{-1}S^{1/2})_{(1/2)})=2\mathrm{tr}\,(S^{1/2}\Sigma^{-1})+\mathrm{tr}(S\Sigma^{-1})/2\mathrm{tr}(S^{1/2})$

(iv) $\partial|S|/\partial S=|S|\,S^{-1}_{(2)}$ and $\partial\mathrm{tr}\,(S^{1/2})/\partial S=(1/2)S^{-1/2}_{(2)}$. ... (3.4)

(Results (i), (ii) and (iii) are proved in the Appendix for the special case $p=2$ and (iv) for a general p).

First, we will find an unbiased estimator for the term $E(|S|/(\mathrm{tr}(S^{1/2}))^2)\mathrm{tr}(S^{1/2}\Sigma^{-1}S^{1/2}\Sigma^{-1})$, which occurs in (3.3), by applying the identity (3.1). Take $h(S)=|S|/(\mathrm{tr}(S^{1/2}))^2$ and $T=S^{1/2}\Sigma^{-1}S^{1/2}$. From (3.1), it follows that

$$E(|S|/(\mathrm{tr}(S^{1/2}))^2)\mathrm{tr}(S^{1/2}\Sigma^{-1}S^{1/2}\Sigma^{-1})$$
$$=2E[(|S|/(\mathrm{tr}(S^{1/2}))^2)D^*((S^{1/2}\Sigma^{-1}S^{1/2})_{(1/2)})]$$
$$+2E(\mathrm{tr}(\partial(|S|/(\mathrm{tr}(S^{1/2}))^2)/\partial S)(S^{1/2}\Sigma^{-1}S^{1/2})_{(1/2)})$$
$$+(n-p-1)E(|S|/(\mathrm{tr}(S^{1/2}))^2)\mathrm{tr}(S^{-1}S^{1/2}\Sigma^{-1}S^{1/2})$$
$$=2E(|S|/(\mathrm{tr}(S^{1/2}))^2)(\mathrm{tr}(\Sigma^{-1})+\mathrm{tr}(S^{1/2}\Sigma^{-1})/\mathrm{tr}(S^{1/2}))$$
$$+2E\mathrm{tr}[(-|S|/(\mathrm{tr}(S^{1/2}))^3)S^{-1/2}_{(2)}+(|S|/(\mathrm{tr}(S^{1/2}))^2)S^{-1}_{(2)}](S^{1/2}\Sigma^{-1}S^{1/2})_{(1/2)}$$
$$+(n-p-1)E(|S|/(\mathrm{tr}(S^{1/2}))^2)\mathrm{tr}(S^{-1}S^{1/2}\Sigma^{-1}S^{1/2}))$$

(By using (3.4) (ii) and (iv)).

Again by using the relation $tr\,M_{(t)}N_{(1/t)}=trMN$ and after some simplification, we have

$$E(|S|/(\mathrm{tr}(S^{1/2}))^2)\mathrm{tr}(S^{1/2}\Sigma^{-1}S^{1/2}\Sigma^{-1})$$
$$=(n-p+3)E(|S|/\mathrm{tr}(S^{1/2}))^2)\mathrm{tr}(\Sigma^{-1}). \quad \text{... (3.5)}$$

Again by applying the identity (3.1) with $T=I$, $h(S)=|S|/(\mathrm{tr}(S^{1/2}))^2$,

$$E(|S|/(\mathrm{tr}(S^{1/2}))^2)tr(\Sigma^{-1})=-2E(|S|^{1/2}/(\mathrm{tr}(S^{1/2}))^2)$$
$$+(n-p+1)E(\mathrm{tr}(S)/(\mathrm{tr}(S^{1/2}))^2)$$
$$=(n-p+2)E(\mathrm{tr}(S)/(\mathrm{tr}(S^{1/2}))^2)-1 \quad \text{... (3.6)}$$

(Since, $(\mathrm{tr}S+2\,|S|^{1/2})/(\mathrm{tr}(S^{1/2}))^2=(\mathrm{tr}(S^{1/2}))^2/(\mathrm{tr}(S^{1/2}))^2=1$, for $p=2$). By substituting (3.6) in (3.5), we get

$$E(|S|/(\mathrm{tr}(S^{1/2}))^2)\mathrm{tr}(S^{1/2}\Sigma^{-1}S^{1/2}\Sigma^{-1})$$
$$=(n-p+3)((n-p+2)E(\mathrm{tr}(S)/(\mathrm{tr}S^{1/2}))^2)-1). \quad \ldots \quad (3.7)$$

To find $E(|S|^{1/2}/\mathrm{tr}(S^{1/2}))(\mathrm{tr}(S^{1/2}\Sigma^{-1}))$ the third term in the square bracket of (3.3), take $h(S)=|S|^{1/2}/\mathrm{tr}(S^{1/2})$ and $T=S^{1/2}$ and apply the identity (3.1). Making use of (3.4) (i) and (iv), we get

$$E(|S|^{1/2}/\mathrm{tr}(S^{1/2}))\mathrm{tr}(S^{1/2}\Sigma^{-1})=(n-p+1). \quad \ldots \quad (3.8)$$

To evaluate $E(|S|^{1/2}/\mathrm{tr}(S^{1/2}))\mathrm{tr}(S\Sigma^{-1}S^{1/2}\Sigma^{-1})$, take $h(S)=|S|^{1/2}/\mathrm{tr}(S^{1/2})$ and $T=S\Sigma^{-1}S^{1/2}$, then from the identity (3.1) and (3.4) (iii) and (iv) we obtain,

$$E(|S|^{1/2}/\mathrm{tr}(S^{1/2}))\mathrm{tr}(S\,\Sigma^{-1}S^{1/2}\Sigma^{-1})$$
$$-(n-p+4)E((|S|^{1/2}/\mathrm{tr}(S^{1/2}))\ \mathrm{tr}(S^{1/2}\Sigma^{-1})$$
$$=(n-p+4)\ (n-p+1)\ (\text{from (3.8)}). \quad \ldots \quad (3.9)$$

By substituting (3.7), (3.8) and (3.9) in (3.3), we get,

$$R_1(\Sigma,\psi(S))=R_1(\Sigma,aS)+b^2(n-p+3)((n-p+2)E(\mathrm{tr}(S)/(\mathrm{tr}(S^{1/2}))^2)-1)$$
$$+2b(n-p+1)(a(n-p+4)-1). \quad \ldots \quad (3.10)$$

Since, $R_1(\Sigma, aS)$ is constant, the expression in the right hand side of (3.10) without the expectation is an unbiased estimator of $R_1(\Sigma, \psi(S))$. Replacing a by a_0 and b by b_0 in (3.10), we get an unbiased estimator of $R_1(\Sigma, \psi_1(S))$.

For $S_{p\times p}>0$ and $m=1,2,\ldots$, we know that $\mathrm{tr}(S^m)/(\mathrm{tr}(S))^m\leq 1$. Making use of this inequality and after substituting $p=2$ in (3.10) we obtain

$$R_1(\Sigma,\psi(S))\leqslant R_1(\Sigma,aS)|_{p=2}+2\,b(n-1)\,(a(n+2)-1)\,+b^2(n^2-1) \quad \ldots \quad (3.11)$$

where $R_1(\Sigma, aS)|_{p=2}=2n(n+3)\,a^2-4na+2$.

Since $\psi(S))$ is a scale and orthogonal equivariant estimator, it is enough to consider $R_1(\Sigma, \psi(S))$ at $\Sigma=\begin{pmatrix}1&0\\0&\theta\end{pmatrix}$.

It can be easily seen that $E(\mathrm{tr}(S)/(\mathrm{tr}(S^{1/2}))^2) \to 1$ as $c \to 0$, (see (8.4)), and so from (3.10) $R_1(\Sigma, \psi(S))$ attains its upper bound for all $a>0$ and $b \geq 0$, when $c \to 0$. Also, we can show that $\psi_1(S)$ given in (2.8) is the unique estimator which minimize the upper bound of $R_1(\Sigma, \psi(S))$. Differentiating the upper bound of $R_1(\Sigma, \psi(S))$ with respect to a and b and equating to zero, we get

$$(n+2)a+(n+1)b=1 \text{ and } 2n(n+3)a+(n-1)(n+2)b=2n.$$

Solving for a and b, we obtain $a=a_0$ and $b=b_0$, where the values of a_0 and b_0 are given in (2.8). The matrix of the second order derivatives is

$$\begin{pmatrix} 2n(n+3) & (n-1)(n+2) \\ (n-1)(n+2) & (n-1)(n+1) \end{pmatrix}$$

and is positive definite. Thus, we proved that $\psi_1(S)$ is the unique minimax estimator among the class of estimators $\psi(S)$.

4. The scale and orthogonal equivariant estimators of Σ under the loss L_2

For the loss L_2 Stein (1961) has given a lower triangular equivariant minimax estimator, which is of the form $TD_\eta T'$, where $TT'=S$ and $D_\eta = \mathrm{diag}\,(\eta_1,\ldots,\eta_p)$. The value of $\eta_j\,(j=1,2,\ldots,p)$ which minimize $R_2(\Sigma, TD_\eta T')$ are $\eta_j^* = \frac{1}{(n+p+1-2j)}\ (j=1,2,\ldots,p)$. Since the loss function is non-negative, non-singular invariant and also a strictly convex function of $\hat{\Sigma}$, it satisfies the conditions of Lemma 2.1 and Lemma 2.2. and so

$$\psi_2(S) = \int_{O(P)} \varphi_{2\Gamma'}(S)\, d\nu(\Gamma) \qquad \ldots \quad (4.1)$$

where $\varphi_{2\Gamma}(S) = \Gamma(T_\Gamma D_{\eta*} T_\Gamma')\Gamma'$ and $T_\Gamma T_\Gamma' = \Gamma' S \Gamma$, is better than $TD_{\eta*}T'$. Since $\psi_2(S)$ is exactly similar to $\psi_1(S)$, it is also scale and orthogonal equivariant. For $p=2$, the equation similar to (2.7) is

$$\psi_2\left(\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}\right) = \frac{1}{1+c^{1/2}} \begin{pmatrix} \eta_1^* + \eta_2^* c^{1/2} & 0 \\ 0 & c(\eta_1^* c^{1/2} + \eta_2^*) \end{pmatrix}.$$

Substituting the values of η_1^* and η_2^* and making use of the relation

$$\psi_2(S) = \lambda_1\, \alpha'\, \psi_2\left(\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}\right) \alpha \text{ where } c = \lambda_2/\lambda_1,$$

(α, λ_i are the same as in (2.4)) $\psi_2(S)$ is seen to be

$$aS+\frac{b\,|\,S\,|^{1/2}\,S^{1/2}}{\text{tr}\,S^{1/2}},\ \text{with}\ a=\frac{1}{n+1}\ \text{and}\ b=\frac{2}{n^2-1}. \quad \dots \quad (4.2)$$

5. THE LOWER TRIANGULAR EQUIVARIANT MINIMAX ESTIMATOR OF Σ^{-1} UNDER $L^{(1)}$

In this section, we first derive a minimax estimator of Σ^{-1} for the loss function $L^{(1)}$, from which we obtain a scale and orthogonal equivariant minimax estimator of Σ^{-1} which is better than the minimax estimator.

It is easy to verify that the lower triangular equivariant estimator is of the form $T'^{-1}\,D_{\delta'}T^{-1}$ where T is a lower triangular matrix such that $TT'=S$ and $D_{\delta'}=\text{diag}\,(\delta'_1 \dots \delta'_p)$.

For the special case $p=2$, let $T=\begin{pmatrix} t_{11} & 0 \\ t_{21} & t_{22} \end{pmatrix}$. Since, $L^{(1)}$ is invariant under non-singular transformations, without loss of generality we can consider the risk of the estimator $T'^{-1}D_{\delta'}T^{-1}$ when $\Sigma=I$. At $\Sigma=I$,

$$\begin{aligned} R^{(1)}(I,\,T'^{-1}D_{\delta'}T^{-1}) &= E\,\text{tr}\,(T'^{-1}D_{\delta'}T^{-1}-I)^2 \\ &= E\,\text{tr}\,(D_{\delta'}\,(T^{-1}T'^{-1})D_{\delta'}(T^{-1}T'^{-1})) \\ &\quad -2E\,\text{tr}\,(D_{\delta'}T^{-1}T'^{-1})+p. \end{aligned} \quad \dots \quad (5.1)$$

From Kshirsagar (1978) we know that t_{ii}^2, $(i=1,2,\dots,p)$ are independent and follow $\chi^2_{(n-i+1)}$ distribution and t_{ij} $(i>j)$ independently follow standard normal distribution. Therefore, for $p=2$,

$$\begin{aligned} E\,\text{tr}\,(D_{\delta'}(T^{-1}T'^{-1})D_{\delta'}(T^{-1}T'^{-1})) &= \frac{(\delta'_1)^2}{(n-2)\,(n-4)}+\frac{2\delta'_1\delta'_2}{(n-2)(n-3)(n-4)} \\ &\quad +\left[\frac{(n-1)\,(n-3)}{(n-2)(n-3)(n-4)(n-5)}\right](\delta'_2)^2. \end{aligned} \quad \dots \quad (5.2)$$

$$E\,\text{tr}\,(D_{\delta'}(T^{-1}T'^{-1}))=\frac{\delta'_1}{(n-2)}+\frac{(n-1)\,\delta'_2}{(n-2)(n-3)}. \quad \dots \quad (5.3)$$

After substituting (5.2) and (5.3) in (5.1), for $p=2$, we have

$$\begin{aligned} R^{(1)}(I,\,T'^{-1}D_{\delta'}T^{-1}) &= \frac{(\delta'_1)^2}{(n-2)(n-4)}+\frac{2(\delta'_1)\,(\delta'_2)}{(n-2)(n-3)(n-4)} \\ &\quad +\left(\frac{(n-1)\,(n-3)}{(n-2)\,(n-3)\,(n-4)\,(n-5)}\right)(\delta'_2)^2 \\ &\quad -2\left(\frac{\delta'_1}{(n-2)}+\frac{(n-1)\,\delta'_2}{(n-2)(n-3)}-2\right). \end{aligned} \quad \dots \quad (5.4)$$

The minimizing $\overset{*}{\delta'_1}$, $\overset{*}{\delta'_2}$ satisfy

$$\delta'_1(n-3)+\delta'_2=(n-3)(n-4), \qquad \ldots (5.5)$$

$$\delta'_1(n-5)+(n-1)(n-3)\delta'_2=(n-1)(n-4)(n-5). \qquad \ldots (5.6)$$

Solving (5.5) and (5.6) for δ'_1 and δ'_2, we have the minimizing

$$\overset{*}{\delta'_1}=\frac{n^4-12n^3+53n^2-98n+56}{n^3-7n^2+14n-4},$$

$$\overset{*}{\delta'_2}=\frac{n^4-14n^3+71n^2-154n+120}{n^3-7n^2+14n-4}.$$

Again from Kiefer's (1957) theorem $T'^{-1}D_{\delta'}{}^*T^{-1}$ is minimax for Σ^{-1}. Since, $\delta_1'^* \neq \delta_2'^*$, no estimator of the form aS^{-1} is minimax for the loss function $L^{(1)}$.

To find the minimax risk, observe that if $\triangle$ denotes $T'^{-1}D_{\delta'}T^{-1}$,

$$E\,\mathrm{tr}\,(c\triangle-I)^2=c^2E\,\mathrm{tr}\,\triangle^2-2c\,E\,\mathrm{tr}\,\triangle+p$$

has minimum at $c=1$ and so we must have $E\,\mathrm{tr}\,\triangle^2=E\,\mathrm{tr}\,\triangle$.
Thus, from (5.1), the minimax risk is

$$p-E\,(\mathrm{tr}\,T'^{-1}D_{\delta'}T^{-1})\,|_{\delta_1'^*,\,\delta_2'^*}$$

$$=2-\frac{\delta_1'^*}{(n-2)}-\frac{(n-1)\delta_2'^*}{(n-2)(n-3)}$$

$$=2-\frac{(n-3)(n-4)+\delta_2'^*(n-2)}{(n-2(n-3)}\text{ from (5.5).} \qquad \ldots (5.7)$$

Remark 5.1. Haff (1979) has given the best multiple of S^{-1} as bS^{-1} with $b=\frac{(n-p-3)(n-p)}{(n-1)}$ under the loss $L^{(1)}$ and its risk is $\frac{2(3n-7)}{(n-1)(n-3)}$ (for $p=2$). Notice that from (5.2) $R^{(1)}(I, T'^{-1}D_{\delta'}T^{-1})$ does not exist for $n=5$. For $n=6$, the minimax estimator is approximately 10% better than the best multiple of S^{-1}, i.e.

$$\frac{R^{(1)}(\Sigma^{-1}, bS^{-1})-R^{(1)}(\Sigma^{-1}, T'^{-1}D_{\delta'}{}^*T^{-1})}{R^{(1)}(\Sigma^{-1}, bS^{-1})}\times 100=10.124.$$

6. The scale and orthogonal equivariant estimator of Σ^{-1} under the loss $L^{(1)}$

Proceeding as in Section 2, we obtain a scale and orthogonal equivariant estimator of Σ^{-1} which is better than the lower triangular equivariant estimator given in Section 5. Lemmas 6.1 and 6.2 needed for this purpose are essentially Lemmas 2.1 and 2.2.

Lemma 6.1. *Let L be a loss function satisfying $L(\Sigma, \hat{\Sigma}) \geqslant 0$, $L(\Sigma, \hat{\Sigma}) = L(B\Sigma B', B'^{-1} \hat{\Sigma} B^{-1})$ for all n. s. B. Then, if $\varphi^{(1)}(S)$ is a minimax estimator with constant risk so is $\varphi_B^{(1)}(S) = B'\varphi^{(1)}(BSB')B$.*

Lemma 6.1 can be applied for the loss function $L^{(1)}$ and so

$$\varphi_{\Gamma'}^{(1)}(S) = \Gamma\varphi^{(1)}(\Gamma' S\Gamma)\Gamma' = \Gamma(U_\Gamma D_{\delta'} {}^* U_{\Gamma'})\Gamma', \qquad \dots (6.1)$$

where $U_{\Gamma'}$ is an upper triangular matrix satisfying $U_\Gamma U_{\Gamma'} = \Gamma' S^{-1} \Gamma$, is a constant risk minimax estimator of Σ^{-1} for any orthogonal Γ. We can construct from the estimators $\varphi_{\Gamma'}^{(1)}(S)$ an orthogonal and scale equivariant minimax estimator $\psi^{(1)}(S)$ which is better than any $\varphi_{\Gamma'}^{(1)}(S)$, making use of the following lemma.

Lemma 6.2. *Let ν be the invariant probability measure over the orthogonal group O(2) and define*

$$\psi^{(1)}(S) = \int_{O(2)} \varphi_{\Gamma'}^{(1)}(S) d\nu(\Gamma')$$

Then $\psi^{(1)}(S)$ is an orthogonal and scale equivariant estimator, is minimax and is better than any $\varphi^{(1)}(S)$. In particular, it is better than $T'^{-1} D_{\delta'} {}^* T^{-1}$ for $\varphi^{(1)}(S) = T'^{-1} D_{\delta'} T^{-1}$.

Next, let S have the spectral decomposition

$$S = \alpha' D_\lambda \alpha, \text{ where } \alpha \perp \text{ and } S^{-1} = \alpha' D_{1/\lambda} \alpha.$$

As $\psi^{(1)}(S)$ is scale and orthogonal equivariant

$$\psi^{(1)}(S) = 1/\lambda_1 \, \alpha' \psi^{(1)}(\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix})\alpha, \text{ where } c = \lambda_2/\lambda_1, \qquad \dots (6.2)$$

it suffices to find $\psi^{(1)} = (\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix})$.

Write

$$\begin{pmatrix} \delta_1^* & 0 \\ 0 & \delta_2^* \end{pmatrix} = \delta_2^{'*} I + (\delta_1^{'*} - \delta_2^{'*})\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \qquad \dots (6.3)$$

Using (6.3), it can be easily seen that

$$\varphi^{(1)}_{\Gamma'}\left(\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}\right)=\delta_2'^*\begin{pmatrix} 1 & 0 \\ 0 & c' \end{pmatrix}+(\delta_1'^*-\delta_2'^*)u_{11}^2\begin{pmatrix} \gamma_{11}^2 & \gamma_{11}\gamma_{21} \\ \gamma_{11}\gamma_{21} & \gamma_{21}^2 \end{pmatrix} \quad \ldots (6.4)$$

where γ_{ij} and u_{ij} are the (i, j) elements of Γ and U respectively. Since, $U_\Gamma U_{\Gamma'}=\Gamma' S^{-1}\Gamma$, $u_{11}^2=c'/(1+(c'-1)\gamma_{22}^2)$ where $c'=1/c$.

In the evaluation of $\psi^{(1)}(S)$, the values of the integrals

$$\int_{O(2)} \frac{\gamma_{11}\gamma_{21}}{1+(c'-1)\gamma_{22}^2}\, d\nu(\Gamma') \text{ and } \int_{O(2)} \frac{\nu_{i1}^2}{1+(c'-1)\gamma_{22}^2}\, d\nu(\Gamma'),\ i=1,2 \quad \ldots (6.5)$$

are required, which we already know from Section 2 as

$$0,\ (1-(c')^{-1/2})\,(c'-1)^{-1} \text{ and } (c')^{-1/2}-(1-(c')^{-1/2})\,(c'-1)^{-1}$$

respectively.

After some simplification we get,

$$\psi^{(1)}\left(\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}\right)=\delta_2'^*\begin{pmatrix} 1 & 0 \\ 0 & c' \end{pmatrix}+(\delta_1'^*-\delta_2'^*)\begin{pmatrix} \dfrac{(c')^{1/2}}{(c')^{1/2}+1} & 0 \\ 0 & (c')^{1/2}-\dfrac{c^{1/2}}{(c^{1/2}+1)} \end{pmatrix}. \quad \ldots (6.6)$$

From (6.2), after substituting the values of $\delta_1'^*$, $\delta_2'^*$, we get

$$\psi^{(1)}=\frac{n^4-14n^3+71n^2-154n+120}{n^3-7n^2+14n-4}\, S^{-1}+\frac{2n^3-18n^2+56n-64}{n^3-7n^2+14n-4}\,\frac{S^{-1/2}}{\operatorname{tr}(S^{1/2})}. \quad \ldots (6.7)$$

7. THE LOWER TRIANGULAR EQUIVARIANT MINIMAX ESTIMATOR AND THE ORTHOGONAL EQUIVARIANT ESTIMATOR OF Σ^{-1} UNDER $L^{(2)}$ LOSS

The lower triangular equivariant minimax estimator of Σ^{-1} is $T'^{-1}D_{\eta'}T^{-1}$, where $TT'=S$, as in the previous section. Since the loss function is non-singular invariant it is enough to consider $R^{(2)}(\Sigma, T'^{-1}D_{\eta'}T^{-1})$ at $\Sigma=I$.

$$R^{(2)}(I, T'^{-1}D_{\eta'}T^{-1})=E\,\mathrm{tr}(D_{\eta'}(T'^{-1}T^{-1}))+E\log|S|$$
$$-\sum_{i=1}^{p}\log\eta'_i-p$$
$$=\frac{\eta'_1}{(n-2)}+\frac{(n-1)\eta'_2}{(n-2)(n-3)}+E\log|S|$$
$$-\log\eta'_1-\log\eta'_2-2\ (\text{when } p=2). \qquad \dots\ (7.1)$$

Differentiating with respect to η'_1 and η'_2 and equating to zero, we obtain minimizing η'^*_1, η'^*_2 as

$$\eta'^*_1=(n-2) \text{ and } \eta'^*_2=\frac{(n-2)(n-3)}{(n-1)}.$$

Now $|S|=|TT'|=\prod_1^p t^2_{ii}$, where $t^2_{ii}\sim\chi^2_{(n-i+1)}$ independently. (Kshirsagar (1978)). So $E\log|S|=\sum_1^p E(\log\chi^2_{n-i+1})$. As is Sugiura and Fujimoto (1982) using digamma function $\Psi(x)=\frac{d\log\Gamma x}{dx}$, we can rewrite it as

$$\sum_1^p E\ (\log(\chi^2_{(n-i+1)})=p\log 2+\sum_1^p\Psi\left(\frac{(n-i+1)}{2}\right).$$

If n is an integer larger than one,

$$\Psi(n)=1+\frac{1}{2}+\dots+\frac{1}{n-1}-\gamma$$

and for half integer argument ($n\geqslant1$),

$$\Psi\left(n+\frac{1}{2}\right)=-\gamma-2\log 2+2\left(1+\frac{1}{3}+\dots+\frac{1}{2n-1}\right),$$ where γ is Euler's constant=0.57721 56649 01532 9...Thus, the minimax risk is

$$R^{(2)}(\Sigma^{-1}, T'^{-1}D_{\eta'}T^{-1})\Big|_{\eta'^*_1,\eta'^*_2}=\Psi\left(\frac{n}{2}\right)+\Psi\left(\frac{n-1}{2}\right)$$
$$-\log\frac{(n-2)^2(n-3)}{(n-1)}+2\log 2. \qquad \dots\ (7.2)$$

Orthogonal equivariant estimator of Σ^{-1} under $L^{(2)}$.

Since $L^{(2)} \geqslant 0$, satisfies the conditions of Lemmas 6.1 and 6.2, the scale and orthogonal equivariant estimator is of the form

$$\psi^{(2)}(S) = \int_{O(2)} \varphi_\Gamma^{(2)}(S) d\nu(\Gamma') \quad \ldots \quad (7.3)$$

where $\varphi_\Gamma^{(2)}(S) = \Gamma\varphi^{(2)}(\Gamma' S \Gamma)\Gamma' = \Gamma(U_\Gamma D_{\eta'^*} U'_\Gamma)\Gamma'$ and U_Γ is an upper triangular matrix which satisfies $U_\Gamma U'_\Gamma = \Gamma' S^{-1}\Gamma$.

The equations similar to (6.2) and (6.5) are

$$\psi^{(2)}(S) = \frac{1}{\lambda_1}\, \alpha' \psi^{(2)}\left(\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}\right) \alpha \text{ where } \alpha \perp \text{ such that } \alpha D_\lambda \alpha' = S, \quad \ldots \quad (7.4)$$

$c = \lambda_2/\lambda_1$ and for $c' = \dfrac{1}{c}$. ... (7.5)

$$\psi^{(2)}\left(\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}\right) = \eta_2'^* \begin{pmatrix} 1 & 0 \\ 0 & c' \end{pmatrix} + (\eta_1'^* - \eta_2'^*) \begin{pmatrix} \dfrac{(c')^{1/2}}{(c')^{1/2}+1} & 0 \\ 0 & (c')^{1/2} - \dfrac{c^{1/2}}{c^{1/2}+1} \end{pmatrix}.$$

After substituting the values of $\eta_1'^*$ and $\eta_2'^*$ in (7.5) and making use of (7.4), we get

$$\psi^{(2)}(S) = (n-2)\, S^{-1} + \frac{2(n-2)}{(n-1)} \, \frac{S^{-1/2}}{\operatorname{tr} S^{1/2}}. \quad \ldots \quad (7.6)$$

Remark 7.1. Haff (1979, 1980) considered the estimators of the form $aS + bg(S)I$ (for Σ) and $a\, S^{-1} + bh(S)I$ (for Σ^{-1}) under Selliah's and Stein's loss functions. Note that these estimators 'correct' only the diagonal elements of aS (or) aS^{-1}, whereas our estimators 'correct' all the elements of aS or aS^{-1}.

8. Comparison of the new estimators with Haff's estimators

Haff (1980) has considered the estimator $\hat{\Sigma}_{1H} = AS + \dfrac{B}{\operatorname{tr} S^{-1}}\, I$ of Σ for the loss function L_1. The optimal choices for A and B are $1/n+3$ and $1/(n+3)(n+1)$ respectively, when $p=2$. We shall show that Haff's estimator $\hat{\Sigma}_{1H}$ with these A and B is not minimax for the loss function L_1.

Theorem 8.1. *When $p=2$, $\hat{\Sigma}_{1H} = AS + BI/trS^{-1}$, with $A = 1/n+3$, $B = 1/(n+3)(n+1)$ is not a minimax estimator under the loss function L_1.*

Proof. The risk of $\hat{\Sigma}_{1H}$ w.r.t. L_1 is

$$\begin{aligned} R_1(\Sigma, \hat{\Sigma}_{1H}) &= E \operatorname{tr} (\hat{\Sigma}_{1H}\Sigma^{-1} - I)^2 \\ &= R_1(\Sigma, AS) + E(B^2 \operatorname{tr} \Sigma^{-2}/(\operatorname{tr} S^{-1})^2 \\ &\quad + 2AB \operatorname{tr} S\Sigma^{-2}/\operatorname{tr} S^{-1} - 2B \operatorname{tr} \Sigma^{-1}/\operatorname{tr} S^{-1}). \quad \ldots (8.1) \end{aligned}$$

Since $\hat{\Sigma}_{1H}$ is a scale and orthogonal equivariant estimator, without loss of generality, we can consider $R_1(\Sigma, \hat{\Sigma}_{1H})$ at $\Sigma = \left(\begin{smallmatrix}1 & 0\\ 0 & c\end{smallmatrix}\right)$ $(c>0)$. Let $1/c=\beta$. The risk of AS is $p(p+1)/(n+p+1)$ when $A=1/(n+p+1)$; using this fact and applying the transformation, $S \to \Sigma^{-1/2}\ S\ \Sigma^{-1/2}$ in (8.1), we have

$$\begin{aligned} R_1(\Sigma, \hat{\Sigma}_{1H}) &= 6/(n+3) + B^2\ E_I\ (1+\beta^2)/(v^{11}+v^{22}\beta) \\ &\quad + 2ABE_I\ (v_{11}+v_{22}\beta)/(v^{11}+v^{22}\beta) \\ &\quad - 2BE_I\ (1+\beta)/(v^{11}+v^{22}\beta) \quad \ldots \quad (8.2) \end{aligned}$$

where v^{ij} and v_{ij} are the elements of V^{-1} and V respectively and $V \sim W_2\ (I, n)$.

Thus,

$$\lim_{\beta\to\infty} R_1(\Sigma, \hat{\Sigma}_{1H}) = 6/(n+3) + B^2E_I(1/(v^{22})^2) + 2ABE_I(v_{22}/v^{22}) - 2BE_I(1/v^{22})$$

$$\begin{aligned} &= 6/(n+3)B^2E_I\ (\ |\ V\ |^2/v_{11}^2) + 2ABE_I\ (v_{22}\ |\ V\ |/v_{11}) - 2BE_I\ (\ |\ V\ |/v_{11}) \\ &= 6/(n+3) + B^2\ (n+1)\ (n-1) + 2AB(n-1)(n+2) - 2B(n-1). \quad \ldots \quad (8.3) \end{aligned}$$

Substituting the optimal values of A and B in (8.3) and after some simplification, we obtain

$$\lim_{\beta\to\infty} R_1(\Sigma, \hat{\Sigma}_{1H}) = 6/(n+3) - (n-1)/((n+3)^2(n+1)) = \lim_{\beta\to 0} R_1(\Sigma, \hat{\Sigma}_{1H})$$

which is greater than $2(3n^2+5n+4)/(n^3+5n^2+6n+4)$ — the minimax risk for the loss function L_1.

Let $\Sigma = \begin{pmatrix}1 & 0\\ 0 & c\end{pmatrix}$. Consider $E\,(\operatorname{tr}S/(\operatorname{tr}S^{1/2})^2)$.

$$E\frac{\operatorname{tr}S}{(\operatorname{tr}S^{1/2})^2} = E_I\frac{\operatorname{tr}(\Sigma^{1/2}V\Sigma^{1/2})}{\operatorname{tr}[(\Sigma^{1/2}V\Sigma^{1/2})^{1/2}]^2} = E\frac{v_{11}+v_{22}c}{v_{11}+v_{22}c+2\ V^{1/2}\sqrt{c}}.$$

Therefore,

$$\lim_{c\to 0} E\frac{\operatorname{tr} S}{(\operatorname{tr} S^{1/2})^2} = 1 = \lim_{c\to\infty} E\ \frac{\operatorname{tr} S}{(\operatorname{tr}S^{1/2})^2}. \quad \ldots \quad (8.4)$$

For the values of a and b given in (2.8), it follows from (8.4) and (3.11) that

$$\lim_{c\to 0} R_1(\Sigma, \psi_1(S)) = \lim_{c\to\infty} R_1(\Sigma, \psi_1(S)) = \frac{2(3n^2+5n+4)}{n^3+5n^2+6n+4}.$$

Thus, $R_1(\Sigma, \psi_1(S))$ attains minimax risk when $c\to 0$ (or) $c\to\infty$ while $R_1(\Sigma, \hat{\Sigma}_{1H})$ tends to a value greater than the minimax risk.

For the loss function L_2, Haff (1980) has considered the estimator $\hat{\Sigma}_{2H} = AS + \frac{BI}{\operatorname{tr}(S^{-1})}$ with the optimal values of $A = \frac{1}{n}$ and $B = \frac{1}{n^2}$. By similar arguments, given in Theorem 8.1, we can prove that $R_2(\Sigma, \psi_2(S))$ attains the minimax risk when $c\to 0$ (or) $c\to\infty$ while $R_2(\Sigma, \hat{\Sigma}_{2H})$ tends to a value greater than the minimax risk.

Theorem 8.2. *For $p=2$, $\hat{\Sigma}_{2H} = AS + \frac{BI}{trS^{-1}}$ with $A = 1/n$ and $B = 1/n^2$ is not a minimax estimator under the loss function L_2.*

Proof.

$$R_2(\Sigma, \hat{\Sigma}_{2H}) = R_2(\Sigma, AS) + BE\frac{\operatorname{tr}\Sigma^{-1}}{\operatorname{tr}S^{-1}} - E\log\left| I + \frac{B}{A}\frac{S^{-1}}{\operatorname{tr}S^{-1}}\right|. \tag{8.5}$$

Let $S \to \Sigma^{-1/2} S \Sigma^{-1/2} = V$. At $\Sigma^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \beta \end{pmatrix}$, we have

$$R_2(\Sigma, \hat{\Sigma}_{2H}) = R_2(\Sigma, AS) + BE_I\frac{(1+\beta)}{(v^{11}+v^{22}\beta)} - E_I\log\left| I + \frac{B}{A}\frac{\Sigma^{-1/2}V^{-1}\Sigma^{-1/2}}{tr\Sigma^{-1/2}V^{-1}\Sigma^{-1/2}}\right|$$

$$= R_2(\Sigma, AS) + BE_I\frac{(1+\beta)}{(v^{11}+v^{22}\beta)} - E_I\log\left| I + (B/A(v^{11}+v^{22}\beta))\begin{pmatrix} v_{11} & v^{12}\cdot\sqrt{\beta} \\ v^{12}\cdot\sqrt{\beta} & v^{22}\beta \end{pmatrix}\right|. \quad \ldots \tag{8.6}$$

Taking limit on both sides of (8.6), we obtain

$$\lim_{\beta\to 0} R_2(\Sigma, \hat{\Sigma}_{2H}) = R_2(\Sigma, AS) + BE_I\frac{|V|}{v_{22}} - \log\left(\frac{B}{A}+1\right)$$

$$= 2\log\frac{n}{2} - \sum_{j=1}^{2}\Psi\left(\frac{(n-j+1)}{2}\right) + \frac{(n-1)}{n^2} - \log\left(\frac{n+1}{2}\right)$$

which can be seen to be greater than the minimax risk

$$\sum_{j=1}^{2} \log (n+3-2j) - \sum_{j=1}^{2} \Psi\left(\frac{n-j+1}{2}\right) - 2 \log 2.$$

Since, the estimator is scale equivariant,

$$\lim_{\beta \to 0} R_2 (\Sigma, \hat{\Sigma}_{2H}) = \lim_{\beta \to \infty} R_2 (\Sigma, \hat{\Sigma}_{2H}) > \text{minimax risk.}$$

Now, we can show that the risk of $\psi_2(S)$ attains minimax risk as $\beta \to 0$ (or) ∞ with respect to the loss L_2.

$$R_2 (\Sigma, \psi_2 (S)) = a\, E \operatorname{tr} S\, \Sigma^{-1} + bE \frac{|S|^{1/2}}{\operatorname{tr} S^{1/2}} \operatorname{tr} S^{1/2} \Sigma^{-1}$$

$$-E \log \left|\left(aS + \frac{b\,|S|^{1/2}}{\operatorname{tr} S^{1/2}}\right)\Sigma^{-1}\right| - 2. \quad \ldots \quad (8.7)$$

After substituting $a = \frac{1}{n+1}$ and $b = \frac{2}{n^2-1}$ in (8.7) and using (3.9) and the relation $\frac{|S|^{1/2}}{\operatorname{tr} S^{-1/2}} = \frac{1}{\operatorname{tr} S^{1/2}}$, we get

$$R_2(\Sigma, \psi_2(S)) = \frac{2n}{n+1} + \frac{2}{n+1} - E \log |S\Sigma^{-1}| + 2 \log (n+1)$$

$$- E \log \left| I + \frac{2}{n-1} \frac{S^{-1/2}}{\operatorname{tr} S^{-1/2}} \right| - 2.$$

$$\ldots \quad (8.8)$$

Consider the transformation $S \to \Sigma^{-1/2} S \Sigma^{-1/2} = V$. This implies that $S^{-1} \to \Sigma^{1/2} S^{-1} \Sigma^{1/2} = V^{-1}$. From (8.6), we have $\Sigma^{-1/2} V^{-1} \Sigma^{-1/2}$ tends to $\begin{pmatrix} v^{11} & 0 \\ 0 & 0 \end{pmatrix}$ as β tends to zero. Hence $(\Sigma^{-1/2} V^{-1} \Sigma^{-1/2})^{1/2}$ tends to $\begin{pmatrix} \sqrt{v^{11}} & 0 \\ 0 & 0 \end{pmatrix}$. Therefore after taking limit on both sides of (8.8), we get

$$\lim_{\beta \to 0} R_2(\Sigma, \psi_2(S)) = 2 \log (n+1) - \sum_{j=1}^{2} E \log \chi^2_{(n-j+1)} - \log \left(\frac{n+1}{n-1}\right)$$

$$=\log(n+1)+\log(n-1)-\sum_{j=1}^{2}\Psi\left(\frac{n-j+1}{2}\right)-2\log 2$$

$$=\text{minimax risk.}$$

As the estimator is scale equivariant $R_2(\Sigma,\psi_2(S))$ also tends to minimax risk as $\beta\to\infty$.

9. Simulation results

Since the actual comparison of the risks of these estimators with Haff's estimators is difficult, we did some simulation study. Ten thousand random matrices from Wishart population were generated with $\Sigma=\left(\begin{smallmatrix}1 & 0\\ 0 & c\end{smallmatrix}\right)$, $n=10$ and $p=2$—using the Fortran program given by Smith and Hocking (1972). For each random matrix, we calculated the value of the loss and took the average of these 10,000 values. To find the risk of $\psi_1(S)$, we used the unbiased estimator given in (3.10).

Haff's estimator $\hat{\Sigma}_{1H}$ (given in Section 8) is compared with $\psi_1(S)$ given by (2.8) under the loss L_1, their risks are given in Table 1. It shows that $\psi_1(S)$ is uniformly better than $\hat{\Sigma}_{1H}$. Also, we see that for large values of c $R_1(\Sigma,\psi_1(S))$ is very near the minimax risk whereas $R_1(\Sigma,\hat{\Sigma}_{1H})$ is greater than the minimax risk, as we have shown theoretically.

Table 2 shows the values of the risks of $\psi_2(S)$ and Haff's estimator $\hat{\Sigma}_{2H}$ given in section 8 for the loss L_2. For $n=10$, $p=2$, Table 2 shows that $\psi_2(S)$ is uniformly better than $\hat{\Sigma}_{2H}$ for all values of c with respect to the loss L_2. Also, we see that $R_2(\Sigma,\psi_2(S))$ approaches the minimax risk as $c\to 0$, whereas as $R_2(\Sigma,\hat{\Sigma}_{2H})$ is greater than the minimax risk for $c\geqslant 10$.

Table 3 shows the values of the risks of $\psi^{(1)}(S)$ and $\hat{\Sigma}_H^{(1)}$, for given values of c under the loss $L^{(1)}$. $\hat{\Sigma}_H^{(1)}$ given by Haff (1979) is

$$\hat{\Sigma}_H^{(1)}=AS^{-1}+\frac{B}{\operatorname{tr}S}I \text{ with } A=(n-5)(n-2)/(n-1) \text{ and } B=(n-5)/(n-1).$$

From Table 3 it seems that $\psi^{(1)}(S)$ is better than $\hat{\Sigma}_H^{(1)}$ for all values of c. Here also, $R^{(1)}(\Sigma,\psi^{(1)}(S))$ approaches the minimax risk for the large values of c, whereas $R^{(1)}(\Sigma,\hat{\Sigma}_H^{(1)})$ is greater than the minimax risk.

Table—1

Loss function L_1 (n=10)

c	Risk of $\psi_1(S)$	Risk of $\hat{\Sigma}_{1H}$
1	.44442	.44576
2	.44444	.44678
3	.44478	.44807
5	.44539	.45002
10	.44640	.45254
20	.44751	.45442
30	.44814	.45519
40	.44857	.45561
50	.44889	.45588
100	.44978	.45644
150	.45022	.45663
200	.45051	.45673
300	.45086	.45684
400	.45108	.45689
10^6	.45265	.45704

For n=10, the minimax risk is =.45269

Table—2

Loss function L_2 (n=10)

c	Risk of $\psi_2(S)$	Risk of $\hat{\Sigma}_{2H}$
1	.30435	.30834
2	.30461	.30920
3	.30500	.31035
5	.30567	.31209
10	.30681	.31434
20	.30804	.31601
30	.30875	.31670
40	.30923	.31707
50	.30958	.31730
100	.31057	.31780
150	.31108	.31800
200	.31139	.31806
300	.31178	.31815
400	.31203	.31819
10^6	.31380	.31833

For n=10, the minimax risk is .31384.

Table—3

Loss function $L^{(1)}$ $(n=10)$

c	Risk of $\psi^{(1)}(S)$	Risk of $\hat{\Sigma}_H^{(1)}$
1	.68195	.69832
2	.68182	.69951
3	.68233	.70128
5	.68353	.70417
10	.68595	.70827
20	.68886	.71168
30	.69063	.71318
40	.69186	.71403
50	.69279	.71458
100	.69545	.71578
150	.69682	.71621
200	.69769	.71644
300	.69880	.71666
400	.69949	.71678
10^6	.70005	.71714

For $n=10$, the minimax risk is .69954.

APPENDIX

For $p=2$, write $S^{1/2}=\begin{pmatrix} t_{11} & t_{12} \\ t_{12} & t_{22}\end{pmatrix}=T.$

We nave to find $D^*(S^{1/2}_{(1/2)})$, $D^*((S\Sigma^{-1}S^{1/2}{}_{(1/2)})$ and $D^*((S^{1/2}\Sigma^{-1}S^{1/2})_{(1/2)})$. From $S^{1/2}S^{1/2}=S$ we have

$$t_{11}^2+t_{12}^2=s_{11},\ t_{12}(t_{11}+t_{22})=s_{12},\ t_{12}^2+t_{22}^2=s_{22}.$$

Differentiate both sides of these with respect to s_{ij} and solve the resulting equations for $\partial t_{11}/\partial s_{ij}$, $\partial t_{12}/\partial s_{ij}$, $\partial t_{22}/\partial s_{ij}$.

For $T_{p\times p}=(T_{ij}(S))$, $(T_{ij}=T_{ji})$

$$D^*(T_{(1/2)})=\text{tr}\left[\left(\frac{\partial}{\partial s_{ij}}\right)_{p\times p} T_{(1/2)}\right]=\sum_{i\leq j}\frac{\partial t_{ij}}{\partial s_{ij}}.$$

For $p=2$, after some simplification we get

(i) $D^*(S^{1/2}_{(1/2)})=\partial t_{11}/\partial s_{11}+\partial t_{12}/\partial s_{12}+\partial t_{22}/\partial s_{22}$

$$=\frac{1}{2\,|T|\,\text{tr}T}\,[(t_{11}+t_{22})^2+2(t_{11}t_{22}-t_{12}^2)]$$

$$=\frac{\text{tr}(S^{1/2})}{2\,|S|^{1/2}}+\frac{1}{\text{tr}(S^{1/2})}.$$

(ii) $D^*((S^{1/2}\Sigma^{-1/2}S^{1/2})_{(1/2)})$

$$=\partial a_{11}/\partial s_{11}+\partial a_{12}/\partial s_{12}+\partial a_{22}/\partial s_{22}$$

where a_{ij} is (i,j) element of $(S^{1/2}\Sigma^{-1}S^{1/2})$. Let σ^{ij} be (i,j) element of Σ^{-1}. After finding $\partial a_{ij}/\partial s_{ij}$, $(i\leqslant j)$ and simplifying we get

$$D^*((S^{1/2}\Sigma^{-1}S^{1/2})_{(1/2)})=\left(1+\frac{t_{11}}{\mathrm{tr}T}\right)\sigma^{11}+\left(1+\frac{t_{22}}{\mathrm{tr}T}\right)\sigma^{22}+\frac{2t_{12}}{\mathrm{tr}T}\sigma^{12}$$

$$=\mathrm{tr}\Sigma^{-1}+\frac{\mathrm{tr}(S^{1/2}\Sigma^{-1})}{\mathrm{tr}(S^{1/2})}.$$

(iii) Similarly, we can find

$$D^*((S\Sigma^{-1}S^{1/2})_{(1/2)})=2\mathrm{tr}(S^{1/2}\Sigma^{-1})+\frac{\mathrm{tr}(S\Sigma^{-1})}{2\mathrm{tr}(S^{1/2})}.$$

(iv) Let $S=\Gamma D_\lambda\Gamma'$, where Γ is orthogonal and $D_\lambda=\mathrm{diag}\ (\lambda_1,\ldots,\lambda_p)$, and $f(\lambda_1,\ldots,\lambda_p)$ be a function of eigen values.

Then,

$$\frac{\partial f}{\partial s_{ij}}=\begin{cases}\sum_{k=1}^{p}\gamma_{ik}^2\frac{\partial f}{\partial\lambda_k}, & i=j\\ 2\sum_{k=1}^{p}\gamma_{ik}\gamma_{kj}\frac{\partial f}{\partial\lambda_k}, & i\neq j\end{cases}$$

where γ_{ik} is (i,j) element of Γ.

Hence

$$\frac{\partial|S|}{\partial s_{ij}}=\frac{\partial\Pi\lambda_i}{\partial s_{ij}}=\begin{cases}\sum_{k=1}^{p}\gamma_{ik}^2(\lambda_1\times\ldots,\lambda_{k-1}\times\lambda_{k+1}\cdots\times\lambda_p), & i=j\\ 2\sum_{k=1}^{p}\gamma_{ik}\gamma_{kj}(\lambda_1\times\ldots,\lambda_{k-1}\times\lambda_{k+1}\cdots\times\lambda_p), & i\neq j\end{cases}$$

$$=\begin{cases}\sum_{k=1}^{p}(\gamma_{ik}^2/\lambda_k)\,|S|\\ 2\sum_{k=1}^{p}(\gamma_{ik}\gamma_{kj}/\lambda_k)\,|S|\end{cases}$$

$$=|S|S_{(2)}^{-1}.$$

$$\frac{\partial\mathrm{tr}S^{1/2}}{\partial s_{ij}}=\begin{cases}\sum_{k=1}^{p}\gamma_{ik}^2/2\sqrt{\lambda_k} & i=j\\ 2\sum_{k=1}^{p}\gamma_{ik}\gamma_{kj}/2\sqrt{\lambda_k} & i\neq j\end{cases}$$

$$=\frac{1}{2}S_{(2)}^{-1/2}.$$

REFERENCES

Eaton, M. L. (1970): Some problems in covariance estimation. Tech. Rep. No. 49, Dept. of Statist. Stanford Univ.

Haff, L. R. (1977): Minimax estimator for multivariate normal precision matrix. *J. Mult. Anal.* 7, 374-385.

Haff, L. R. (1979): An Identity for the Wishart Distribution with Applications. *J. Mult. Anal.* 9, 531-544.

Haff, L. R. (1979): Estimation of the inverse covariance matrix: Random mixtures of the inverse Wishart matrix and the Identity. *Ann. Statist.* 7, 1264-1276.

Haff, L. R. (1980): Empirical Bayes estimation of the multivariate normal covariance matrix. *Ann. Statist.* 8, 586-597.

James, A. T. (1954): Normal multivariate Analysis and orthogonal group. *Ann. Math. Statist.* 25, 40-75.

James, W. and Stein, C. (1961): Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* Univ. California Press.

Kiefer, J. (1957): Invariance minimax sequential estimation and continuous time processes. *Ann. Math. Statist.* 25, 40-75.

Kshirsagar, A. M. (1978): *Multivariate Analysis*. Marcell Dekker, New York.

Lehmann, E. L. (1959): *Testing Statistical hypotheses.* John Wiley, New York.

Selliah, J. B. (1964): Estimation and testing problems in Wishart distribution. *Tech.* Rep. No. 10, Dep. of Statist., Stanford Univ.

Sharma, Divakar, (1980): An estimator of normal covariance matrix. *Calcutta Statist. Assoc. Bull.* 29, 161-167.

Smith, W. B. and Hocking, R. R. (1972): Wishart variate generator, (Algorithm AS 53). *Appl. Statist.* 21, 341-345.

Sugiura, N. and Fujimoto, M. (1982): Asymptotic risk comparison of improved estimators for normal covariance matrix. *Tsukuba J. Math.* 6, 103-126.