Reprinted from :
CALCUTTA STATISTICAL ASSOCIATION BULLETIN
Vol. 34, March & June, 1985, Nos. 133—134

IMPROVED MINIMAX ESTIMATORS OF NORMAL COVARIANCE AND PRECISION MATRICES FROM INCOMPLETE SAMPLES

DIVAKAR SHARMA and K. KRISHNAMOORTHY

Indian Institute of Technology, Kanpur

ABSTRACT : Let X be an N_p $(\mathbf{o}, \Sigma)$ random vector. Suppose besides n observations on X, m observations on the first $q(q < p)$ coordinates are available. Eaton (1970), for this set up, has given a minimax estimator of Σ, which is better than the MLE. We, in this paper, obtain a class of constant risk minimax estimators (Eaton's estimator is its member), and hence estimators better than any member of this class. Similar results are derived also for the estimation of Σ^{-1}. The loss functions considered are those of Selliah (1964) and James and Stein (1961) for the estimation of Σ and an analogue of Stein's loss function for the estimation of Σ^{-1}.

1. Introduction

Consider a p-varitate normal population with a known mean vector and an unknown covariance matrix Σ. The mean vector, without loss of generality, can be assumed to be null. Let $(X_1, \cdots, X_n)$ be a random sample from this population. Also, suppose that we have m observations $Y_1, \cdots, Y_m$ available from the $N_q(\mathbf{o}, \Sigma_{11})$ population where $q < p$ and

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \ \Sigma_{11} \text{ of order } q \times q.$$

This amounts to having $m+n$ observations from the $N_p(\mathbf{o}, \Sigma)$ population in which m observations have information only on the first q coordinates. The problem, we are interested in, is one of estimating Σ or Σ^{-1} from this incomplete sample. Anderson (1957) and Sylvan (1969) derived the MLE of Σ in such a set up. Eaton (1970) obtained a minimax estimator of Σ, when the loss function is

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

The loss function L_1 was first considered by James and Stein (1961) and Eaton's estimator is, in fact, best lower triangular equivariant. For its derivation, he has used a theorem due to Stein (see, for example, Zidek (1969), which reprssents the best equivariant estimator formal Bayes estimator with respect to a right invariant prior. as a Kiefer's (1957) theorem ensures the minimaxity of the estimator. The MLE also being lower triangular equivariant is worse than Eaton's estimator.

We, in Section 2, following Eaton's approach, obtain minimax estimator of Σ under the loss function

$$L_2(\Sigma, \hat{\Sigma}) = \text{tr } (\hat{\Sigma}\Sigma^{-1} - I)^2.$$

As the computation for general p and q is difficult, we restrict to the cases $p=2$, $q=1$; and $p=3$, $q=1$.

Let $\phi_i(S, V)$ $(i=1, 2)$, where $S=\sum_i X_i X'_i$ and $V=\sum_i Y_i Y'_i$, denote the minimax estimators, mentioned above, when the loss function is L_i. In Section 3, we obtain a class of minimax estimators under $L_i (i=1, 2)$ which contains $\phi_i(S, V)$. Then we find an estimator $\hat{\Sigma}_i^*$ which is better than any member of this class ; $\hat{\Sigma}_i^*$ is, in fact, the best average of two members of the class. It is shown that the best average is the simple average for any p and q when the loss function is L_2 and for $p-q=2$ when the loss function is L_1.

In Section 4, we derive an estimator $\psi_i(S, V)$ $(i=1, 2)$ which is equivariant under a group of transformations G_Γ for the case $p=3$, $q=1$, a member of G_Γ is of the form $\begin{pmatrix} 1 & O \\ O & \Gamma_{22} \end{pmatrix}$ with Γ_{22} orthogonal.

Section 5 is concerned with the estimation of Σ^{-1} under the loss

$$L^{(1)}(\Sigma^{-1}, \hat{\Sigma}^{-1}) = \text{tr } (\hat{\Sigma}^{-1} \Sigma) - \log | \hat{\Sigma}^{-1} \Sigma | - p$$

We first obtain a minimax estimator $\phi^{(1)}(S, V)$ for the cases $p=3$, $q=1$ and $p=2$, $q=1$. As in Section 3, we then find a class of minimax estimators containing $\phi^{(1)}(S, V)$ and derive estimators $\hat{\Sigma}_*^{(1)}$ and $\psi^{(1)}(S, V)$ similar to $\hat{\Sigma}_1^*$ and $\psi^{(1)}(S, V)$ respectively. Both $\hat{\Sigma}_*^{(1)}$ and $\psi^{(1)}(S, V)$ are better than $\phi^{(1)}(S, V)$.

In Section 6, we find expressions for the risks of the MLE under the losses L_1 and L_2 when $p=3$ and $q=1$. We also obtain in risks of the estimators $\phi_1(S, V)$ and $\phi^{(1)}(S, V)$ through numerical integration. Comparison of the risks shows us the amount by which $\phi_i(S, V)$ dominates the MLE. The exact risk difference between

$\phi_1(S, V)$ and $\hat{\Sigma}_1^*$ and between $\phi^{(1)}(S, V)$ and $\hat{\Sigma}_*^{(1)}$ are obtained in Section 7 and presented in Tables 3(a) and 3(b).

The evaluation of the risks of $\phi_2(S, V)$, $\hat{\Sigma}_2^*$ and $\psi_1(S, V)$ is difficult and we estimate them using Monte-Carlo method. This is the topic of Section 8. Both $\psi_i(S, V)$ and $\hat{\Sigma}_i^*$ seem to be considerable improvements over the MLE under the loss $L_i (i=1, 2)$. The estimatoss $\hat{\Sigma}_i^*$ and $\psi_i(S, V)$ are not comprable, though $\psi_i(S, V)$ is an average oaer a larger set of constant risk minimax estimators.

2. Best lower triangular equivariant estimators of Σ

Let the $p \times p$ positive definite matrix Σ be partitioned as $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, where Σ_{11} is $q \times q$ $(q < q)$. Suppose $X_1, \cdots X_n$ are i.i.d. $N_p(O, \Sigma)$ and $Y_1, \cdots, Y_m$ are i. i.d. $N_q(\mathbf{o}, \Sigma_{11})$ and the X's are statistically independent of the Y's. Consider the estimation of Σ under a loss function L_0 which is fully invariant, that is,

$$L_0(\Sigma, \hat{\Sigma}) = L_0(A\Sigma' A, \; A\hat{\Sigma}A') \text{ for any non-singular } A.$$

Since (S, V), where $S = \sum_i X_i X'_i$ and $V = \sum_i Y_i Y'_i$, is a sufficient statistic for $\{ \Sigma : \Sigma > O \}$, we consider estimators depending on S and V only. Notice that S and V are independently $W_p(\Sigma, n)$ and $W_q(\Sigma_{11}, m)$ respectively.

Let G_1 denote the class of lower triangular matrices of order $p \times p$ and any matrix A in G_1 be partitioned as

$$A = \begin{pmatrix} A_{11} & O \\ A_{21} & A_{22} \end{pmatrix}, \qquad \cdots \qquad (2.1)$$

where A_{11} is of order $q \times q$. Then, the estimation problem is invariant under the transformations $S \to ASA'$, $V \to A_{11} V A'_{11}$, $\Sigma \to A\Sigma A'$, $\Sigma_{11} \to A_{11} \Sigma A'_{11}$ and $\hat{\Sigma} \to A\hat{\Sigma}A'$, that is, it is lower triangular invariant. According to a theorem due to Stein (Zidek (1969)), the best lower tringular equivariant estimator is formal Bayes with respect to a right invariant prior. It is minimax since G_1 is solvable (Kiefer (1957)). Eaton (1970), used these results to obtain minimax estimator. His estimator is described below.

Let S be factorized as $S = TT'$, $T \varepsilon G_1$. Denote $T_{11}^{-1} V T'^{-1}_{11}$ (T_{11} is as in (2.1)) by U and $\xi \varepsilon G_1$ be such that

$$\xi\xi' = \begin{pmatrix} I+U & O \\ O & I \end{pmatrix}$$

Also, let $L_0(\Sigma, \hat{\Sigma})$ be denoted by $L(B\hat{\Sigma}B')$, $B \in G_1$ and $BB' = \Sigma^{-1}$, and d^* minimize

$$H(d) = \int_{G_1} L(ZdZ') \mid Z \mid^n \mid Z_{11} \mid^m$$

$$\exp\left[-\tfrac{1}{2} \operatorname{tr} ZZ' + \tfrac{1}{2} \operatorname{tr} Z_{21}\xi_{11}^{-1}U\xi'^{-1}_{11}Z'_{21}\right] \nu(dZ) \quad \cdots \quad (2.2)$$

with respect to d, where $\nu(dZ)$ is $(\prod_1^p Z_{ii}^{-i})dZ$ and corresponds to the left invariant measure on G_1. Then the best lower tringular equivariant estimator is given by $T\xi d^*\xi' T'$. When the loss function is L_1, Eaton has shown that the estimator is

$$T\begin{pmatrix} \xi'^{-1}_{11} D_1 \xi_{11}^{-1} + (p-q)I)^{-1} & O \\ O & D_2^{-1} \end{pmatrix} T', \quad \cdots \quad (2.3)$$

where $D_1 = \operatorname{diag}(d_1^{(1)}, \cdots, d_q^{(2)})$; $d_i^{(1)} = m + r + q - 2i + 1$ $(i = 1, \cdots, q)$ and $D_2 = \operatorname{diag}(d_1^{(2)}, \cdots, d_{p-q}^{(2)})$: $d_i^{(2)} = n + p + 2q - 2i + 1$ $(l = 1, \cdots, p-q)$.

We derive below the best lower triangular equivariant estimator of Σ for the loss function L_2. The minimization of (2.2) is equivalent to the minimization of

$$\int_{G_1} L(ZdZ') \mid Q \mid^{\frac{p-q}{2}} \mid Z_{11} \mid^{m+n} \mid Z_{22} \mid^n .$$

$$\exp\left[-\tfrac{1}{2}\operatorname{tr} Z_{11}Z'_{11} - \tfrac{1}{2}\operatorname{tr} Z_{22}Z'_{22} - \tfrac{1}{2}\operatorname{tr} Z_{21}QZ'_{21}\right]\nu(dZ) \quad \cdots \quad (2.4)$$

where $L(ZdZ') = \operatorname{tr}\ (ZdZ')^2 - 2\ \operatorname{tr}\ (ZdZ') - p$ and $Q = I - \xi_{11}^{-1}U\xi'^{-1}_{11}$. Since evaluating (2.4) for general p and q is difficult, we consider the special cases of $p = 2$, $q = 1$; and $p = 3$, $q = 1$.

Case I. $p = 2, q = 1$. The quantity (2.4) is

$$\int_{G_1} L(LdZ') \mid Q \mid^{\frac{1}{2}} (z_{11}^2)^{\frac{m+n}{2}-1} (z_{22}^2)^{\frac{n-1}{2}-1}$$

$$\exp\left[-\frac{z_{11}^2}{2} - \frac{z_{22}^2}{2} - \frac{z_{21}^2}{2\beta}\right] dz_{11}^2\, dz_{22}^2\, dz_{21}, \quad \cdots \quad (2.5)$$

where Q is a scalar and $\pi = Q^{-1}$. Except for a normalizing constants (2.5) is the expection of $L(ZdZ')$ under the joint distribution of z_{11}^2, z_{22}^2 and z_{21} which are independently χ^2_{m+n}, χ^2_{n-1} and $N(0, \beta)$ respectively. It can be easily verified that

$$\begin{aligned} EL(ZdZ') = {} & [M(M + 2 + 2\beta) + 2\beta^2]\ d_{11}^2 + (n^2 - 1)d_{22}^2 \\ & + 2(n-1)(M + 2\beta)d_{12}^2 + 2\beta(n-1)d_{11}d_{22} - (2M + \beta)d_{11} \\ & - 2(n-1)d_{22} + 2, \end{aligned} \quad \cdots \quad (2.6)$$

where d_{ij} are the elements of d (d is a symmetric matrix) and $M = m+n$, The minimizing d is then given by

$$\begin{pmatrix} d_{H}^{*} & 0 \\ 0 & d_{22}^{*} \end{pmatrix} = \begin{pmatrix} \frac{M(n+1)+2\beta}{(n+1)x-(n-1)\beta^2} & 0 \\ 0 & \frac{x-(M+\beta)\beta}{(n+1)x-(n-1)\pi^2} \end{pmatrix},$$

where $x = M(M+2+2\beta)+3\beta^2$ and $\beta = 1+v/s_{11}(s_{11}$ is the (1.1) element of S). Thus the best lower triangular equivariant estimator is

$$\phi_2(S, V) = T\,\xi\, d^*\, \xi'\, T' = T \begin{pmatrix} d_{11}^{*}\beta & 0 \\ 0 & d_{22}^{*} \end{pmatrix} T'.$$

Case II. $p=3$, $q=1$. Noting that (2.4) is proportional to $EL(ZdZ')$ where $z_{11}^2, z_{22}^2, z_{33}^2, z_{32}, z_{21}$ and z_{31} are innependently $\chi^2_{m+n}, \chi^2_{n-1}, \chi^2_{n-2}$, $N(0, 1)$, $N(0, \beta)$ and $N(0, \beta)$ respectively and proceeding as in the earlier cases, we find

$$\phi_2(S, V) = T \begin{pmatrix} d_{11}^{*}(1+V/s_{11}) & 0 & 0 \\ 0 & d_{22}^{*} & 0 \\ 0 & 0 & d_{33}^{*} \end{pmatrix} T',$$

with

$$d_{11}^{*} = \frac{1}{\beta} - \frac{(n^3+2n^2-n+2)d_{33}^{*}}{n(n+1)\beta},\quad d_{22}^{*} = \frac{n^2+2-n}{n(n+1)} d_{33}^{*}$$

$$d_{33}^{*} = \frac{n\{y-\beta(M+2\beta)\}(n+1)}{2n(n-1)(y-n\beta^2)+y(n^2-n+2)(n+1)}, \quad \cdots \quad (2.7)$$

$$y = M(M+2)+4\beta(M+2\beta),\ M = m+n$$

and $\beta = 1+V/s_{11}$.

3. Derivation of $\hat{\Sigma}_i^*$ under the loss $L_i (i=1, 2)$.

The estimator $\phi_i(S, V)$ of Section 2 is a constant risk minimax estimator and one can use the following lemma to generate a class of such estimators.

Lemma 3.1. *Suppose L is a fully invariant loss function, then, if $\phi(S, V)$ is a minimax estimotor with constant risk so is $\phi_{\Gamma'}(S, V) = \Gamma\phi(\Gamma' S\Gamma, \Gamma'_{11} V\Gamma_{11})\Gamma'$ for any orthogonal matrix Γ of the form* $\begin{pmatrix} \Gamma_{11} & 0 \\ 0 & \Gamma_{22} \end{pmatrix}$.

Proof: Let $E_{\Sigma, \Sigma_{11}}$ denote the expectation under the joint dis-

tribution of $S \sim W_p(\Sigma, n)$ and $V \sim W_q(\Sigma_{11}, m)$. Then $E_{\Sigma, \Sigma_{11}} L(\Sigma, \Gamma\varphi(\Gamma' S\Gamma, \Gamma'_{11} V\Gamma_{11})\Gamma') = E_{\Gamma'\Sigma\Gamma, \Gamma'_{11}\Sigma_{11}\Gamma_{11}} L(\Sigma, \Gamma\varphi(S, V)\Gamma')$
$= E_{\Gamma'\Sigma\Gamma, \Gamma'_{11}\Sigma_{11}\Gamma_{11}} L(\Gamma'\Sigma\Gamma, \varphi(S, V)) = E_{\Sigma, \Sigma_{11}} L(\Sigma, \varphi(S, V))$,
which proves the assertion.

If $L(\Sigma, \hat{\Sigma})$ is also a strictly convex function of $\hat{\Sigma}$, any average $\alpha\varphi_\Gamma(S, V) + (1-\alpha)\varphi_\Gamma(S, V)$, $0 < \alpha < 1$, $\Gamma \neq \eta$, is better than $\varphi(S, V)$ and is minimax. The loss functions L_1 and L_2 can be easily seen to be strictly convex. We show that $\alpha = \frac{1}{2}$ is the best choice for any $p - q = 2$ when the loss function is L_1 and for any p and q when the loss function is L_2. Recall that the best average under L_1 is being denoted by $\hat{\Sigma}_1^*$. One could also consider the average of estimators $\varphi_{\Gamma'}(S, V)$ with respect to an invariant probability measure over the group of orthogonal matrices of the form $\begin{bmatrix} I & O \\ O & \Gamma_{22} \end{bmatrix}$ to obtain an orthogonal and scale equivariant estimator $\psi(S, V)$. Jensen's inequality ensures the dominance of $\psi(S, V)$ over $\varphi(S, V)$. As the derivation of such estimators is difficult for general p and q, we take the special case of $p - q = 2$ for L_1 and $p = 3$, $q = 1$ for L_2. Such estimators denoted by $\psi_i(S, V)$ to indicate the dependence on the corresponding loss function L_i $(i = 1, 2)$ are considered in the next section.

We first take up the loss function L_1 and obtain an expression for $\hat{\Sigma}_1^*$. Let the estimator $\varphi_1(S, V)$ given by (2.3) be written as $T\begin{bmatrix} D_1^* & O \\ O & D_2^* \end{bmatrix} T'$, then it is equal to

$$\begin{bmatrix} T_{11} D_1^* T'_{11} & T_{11} D_1^* T'_{21} \\ T_{21} D_1^* T'_{11} & T_{21} D_1^* T'_{21} + T_{22} D_2^* T'_{22} \end{bmatrix}.$$

Notice that $T_{22}T'_{22} = S_{22} - S_{21}S_{11}^{-1}S_{12} = S_{22.1}$ (say) has a $W_{p-q}(\Sigma_{22.1}, n-q)$ distribution and

$$D_2^* = \text{diag}\ (d_1^{(2)^{-1}}, \ldots d_{p-q}^{(2)^{-1}}),$$

so that $T_{22} D_2^* T'_{22}$ is Stein's [4] estimator of $\Sigma_{22.1}$ based on $S_{22.1}$. If Γ is taken to be

$$\begin{bmatrix} I & O \\ O & J \end{bmatrix}, \text{ where } J = \begin{bmatrix} 0 \cdots 1 \\ \cdots\cdots \\ 1 \cdots 0 \end{bmatrix},$$

one can easily check that the simple average of $\varphi_1(S, V)$ and $\varphi_{1\Gamma'}(S, V) = \Gamma\varphi_1(\Gamma' S\Gamma, \Gamma'_{11} V\Gamma_{11})\Gamma'$ is

$$\begin{bmatrix} T_{11} D_1^* T'_{11} & T_{11} D_1^* T'_{21} \\ T_{21} D_1^* T'_{11} & T_{21} D_1^* T_{21} + \tfrac{1}{2}(T_{22} D_2^* T'_{22} + U_{22} D_2^{0*} U'_{22}) \end{bmatrix}$$

$$= \varphi_1^0(S, V) \text{ (say)}, \qquad \dots (3.1)$$

where U_{22} is an upper triangular matrix satisfying $U_{22}U'_{22} = S_{22.1}$ and $D_2^{0*} = \text{diag}\,(d_{p-q}^{(2)^{-1}}, \dots, d_1^{(2)^{-1}})$. The estimator $U_{22}D_2^{0*}U_{22}$ can be described as the best upper triangular equivariant estimator of $\Sigma_{22.1}$ based on $S_{22.1}$.

For the special case of $p=3$ and $q=1$, (3.1) is

$$\begin{bmatrix} A_{11} & A_{12} \\ A'_{12} & A'_{12} A_{11}^{-1} A_{12} + \tfrac{1}{2}(T_{22} \; D_2^* \; T'_{22} + U_{22} \; D_2^{0*} \; U'_{22}) \end{bmatrix}, \dots \quad (3.2)$$

where $A_{11} = d_1^*\,(S, V)\, s_{11}$, $A_{12} = d_1^*\,(S, V)\,(s_{12}\, s_{13})$,

$$d_1^*\,(S, V) = \left(\frac{M}{1 + V/s_{11}} + 2\right)^{-1} \text{and } D_2^* = \text{diag}\left(\frac{1}{n}, \frac{1}{n-2}\right).$$

Because of the following theorem, (3.2) is $\hat{\Sigma}_1^*$, the best average of $\varphi_1(S, V)$ and $\varphi_1^0(S, V)$. Incidentally, $\frac{1}{2}(T_{22}D_2^* \; T'_{22} + U_{22}D_2^{0*}U'_{22})$ is also the best average of $T_{22}D_2^* \; T'_{22}$ and $U_{22}D_2^{0*}U'_{22}$ for estimating $\Sigma_{22.1}$ from $S_{22.1}$.

Theorem 3.2. *For the loss function L_1 and any p and q with $p-q=2$, the best average of the estimators $\varphi_{1\Gamma'}(S, V)$ and $\varphi_{1\eta'}(S, T)$, where Γ and η are orthogonal matrices of the form $\begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix}$ and B_{11} is $q \times q$, is the simple average.*

Proof: The risk difference

$$R_1(\Sigma, \varphi_{1\Gamma'}(S, V)) - R_1(\Sigma, \alpha\varphi_{1\Gamma'}(S, V) + (1-\alpha)\varphi_{1\eta'}(S, V))$$

$$= -E \log \begin{vmatrix} A_{11} & A_{12} \\ A'_{12} & A'_{12}A_{11}^{-1}A_{12} + P_1 \end{vmatrix}$$

$$+ E \log \begin{vmatrix} A_{11} & A_{12} \\ A'_{12} & A'_{12}A_{12}^{-1}A_{12} + (\alpha P_1 + (1-\alpha) P_2) \end{vmatrix}$$

$$= E(\log |\alpha P_1 + (1-\alpha) P_2| - \log |P_1|),$$

where $P_1 = \Gamma_{22} T_{22\Gamma} D_2^* \; T'_{22\Gamma}\Gamma'_{22}$, $P_2 = \eta_{22} T_{22\eta} D_2^* \; T'_{22\eta}\eta_{22}$, $T_{22\Gamma}$ and $T_{22\eta}$ are lower triangular matrices of order $p-q$ such that $T_{22\Gamma} T'_{22\Gamma} = \Gamma'_{22} S_{22.1} \Gamma_{22}$ and $T_{22\eta} T'_{22\eta} = \eta'_{22} S_{22.1} \eta_{22}$. As shown in Fujimoto (1982), using the fact that $|P_1| = |P_2|$, it can be easily verified that $|\alpha P_1 + (1-\alpha) P_2|$ is maximized by $\alpha = \frac{1}{2}$.

Remark 3.2.1. We do not know what the best choice of α, if any, for $p-q \geqslant 3$ is. A necessary and sufficient condition for $\alpha = \frac{1}{2}$ to be the best value is

$$E \operatorname{tr} P_1 (P_1 + P_2)^{-1} = (p-q)/2. \quad \cdots \quad (3.4)$$

When $\Sigma_{22.1} = J\Sigma_{22.1} J$, (3.4) can be easily seen to hold.

Consider next the loss function L_2. Denote the estimator $\Gamma\varphi_2 (\Gamma' S\Gamma, \Gamma'_{11} V\Gamma_{11})\Gamma'$ by $\varphi_{2\Gamma'}(S, V)$. We state and prove below a theorem similar to Theorem 3.2.

Theorem 3.3. *For the loss function L_2 and any p and q, the best average of the estimators $\varphi_{2\Gamma'}(S, V)$ and $\varphi_{2\eta'}(S, V)$, where Γ and η are as in Theorem 3.2, is the simple average.*

Proof: Since $\varphi_{2\Gamma'}(S, V)$ and $\varphi_{2\eta'}(S, V)$ have the same risk,

$$R_2(\Sigma, \varphi_{2\Gamma'}(S, V) - R_2 (\Sigma, \alpha\varphi_{2\Gamma'}(S, V) + (1-\alpha)\varphi_{2\eta'}(S, V))$$
$$= \alpha R_2 (\Sigma, \varphi_{3\Gamma'}(S, V) + (1-\alpha) R_2 (\Sigma, \varphi_{2\eta'}(S, V))$$
$$- R_2 (\Sigma, \alpha\varphi_{2\Gamma'}(S, V) + (1-\alpha)\varphi_{2\eta'}(S, V)$$
$$= 2\alpha(1-\alpha) E_{\Sigma, \Sigma_{11}} [\operatorname{tr} (\varphi_{2\Gamma'}(S, V)\Sigma^{-1} - I)^2$$
$$- \operatorname{tr} (\varphi_{2\Gamma'}(S, V)\Sigma^{-1} - I)(\varphi_{2\eta'}(S, V)\Sigma^{-1} - I)],$$

which is maximised by $\alpha = \frac{1}{2}$.

If, in particular, $p=3$, $q=1$, $\Gamma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

the best average $\hat{\Sigma}_2^*$ is given by

$$\hat{\Sigma}_2^* = \begin{bmatrix} \delta_1^*(S, V)s_{11} & \delta_1^*(S, V)S_{12} \\ \delta_1^*(S, V)S_{21} & \delta_1^*(S, V)S_{21}s_{11}^{-1}S_{12} + \frac{1}{2}(T_{22}\Delta^* T'_{22} \\ & + U_{22}\Lambda^{o*}U'_{22}) \end{bmatrix},$$

where $\delta_1^*(S, V) = d_{11}^*(1 + V/s_{11})$, $\Delta^* = \operatorname{diag}(d_{22}^*, d_{33}^*)$, $\Delta^{o*} = \operatorname{diag}(d_{33}^*, d_{22}^*)$, d_{ii}^* $(i = 1, 2, 3)$ are given by (2.7), and T_{22} and U_{22} are as before lower triangular and upper triangular matrices with $T_{22}T'_{22} = U_{22}U'_{22} = S_{22.1}$. Observe that we cannot describe $T_{22}\Delta^* T'_{22}$ and $U_{22}\Delta^{o*}U'_{22}$ as the best lower triangular equivariant and the best upper triangular equivariant estimators of $\Sigma_{22.1}$ based on $S_{22.1}$, as we could for similar expressions in $\hat{\Sigma}_1^*$.

4. Derivation $\psi_i(S, V)$ under the loss L_i $(i=1, 2)$.

As mentioned in Section 3, since $\varphi_i(S, V)$ is a constant risk minimax estimator, so is $\varphi_{i\Gamma'}(S, V)$ for any orthogonal matrix Γ of of the form $\begin{bmatrix}\Gamma_{11} & 0 \\ 0 & \Gamma_{22}\end{bmatrix}$. Let G_Γ be the group of such orthogonal, matrices. Define

$$\psi_i(S, V) = \int_{G_\Gamma} \varphi_{i\Gamma'}(S, V)\, d\nu(\Gamma) \qquad \text{... (4.1)}$$

where ν is an invariant probability measure on G_Γ. The estimator $\psi_i(S, V)$ can be easily verified to be scale and G_Γ-equivarient. Strict convexity of the loss function and Jensen's inequality imply that $\psi_i(S, V)$ is better than $\varphi_i(S, V)$. As the derivation of $\psi_i(S, V)$ for general Γ, p and q is difficult, we shall take $\Gamma_{11}=1$ and $p-q=2$. For the derivation of $\psi_2(S, V)$ we shall consider only $p=3$ and $q=1$.

As in Section 3, write $\varphi_1(S, V) = T\begin{bmatrix}D_1^* & 0 \\ 0 & D_2^*\end{bmatrix}T'$. Then since

$$D_i^*\ (S, V) = D_i^*\ (\Gamma S \Gamma', \Gamma_{11})\ ;\ i=1, 2, \qquad \text{... (4.2)}$$

it can be easily seen that

$$\varphi_{1\Gamma'}(S, V) = \begin{bmatrix} T_{11} D_1^* T'_{11} & T_{11} D_1^* T'_{21} \\ T_{21} D_1^* T'_{11} & T_{21} D_1^* T'_{21} + \Gamma_{22} T_{22\Gamma} D_2^* T'_{22\Gamma} \Gamma'_{22}\end{bmatrix}.$$

where $T_{22\Gamma}$ is a lower triangular matrix satisfying $T_{22\Gamma}T'_{22\Gamma}=\gamma'_{22} S_{22.1}\Gamma_{22}$. Taking ν now to be the invariant probability measure over the group of orthogonal matrices of order 2×2, $\psi_1(S, V)$, from Sharma and Krishnamoorthy (1983), is

$$\begin{bmatrix} T_{11} D_1^* T'_{11} & T_{11} D_1^* T'_{21} \\ T_{21} D_1^* T'_{11} & T_{21} D_1^* T'_{21} + \psi_1^0(S_{22.1})\end{bmatrix}$$

with

$$\psi_1^0(S_{22.1}) = \frac{1}{n-q+1}\, S_{22.1} = \frac{2}{(n-q)^2-1}\, \frac{S_{22.1}^{1/2}}{\operatorname{tr} S_{22.1}^{-1/2}}.$$

To obtain $\psi_2(S, V)$, we consider $p=3$ and $q=1$. This sort of restriction becomes necessary since we do not know if a relation similar to (4.2) exists for general p and q. Proceeding as for $\psi_1(S, V)$, it can be easily verified that

$$\psi_2(S, V) = \begin{bmatrix} \delta_1^*\ (S, V)\, s_{11} & \delta_1^*\ (S, V)\, S_{12} \\ \delta_1^*\ (S, V)\, S_{21} & \delta_1^*\ (S, V)\, S_{21}\, s_{11}^{-1}\, S_{12} + \psi_2^0(S_{22.1})\end{bmatrix}.$$

where $\psi_2^0(S_{22.1}) = \delta_2^*\ (S, V) S_{22.1} + (\delta_3^*\ (S, V) - \delta_2^*\ (S, V))\, \dfrac{S_{22.1}^{1/2}}{\operatorname{tr} S_{22.1}^{-1/2}}$,
$\delta_1^*\ (S, V) = d_{11}^*(1+V/s_{11})$, $\delta_2^*\ (S, V) = d_{22}^*$, $\delta_3^*\ (S, V) = d_{33}^*$, and

d^*_{ii} $(i=1, 2, 3)$ are given by (2.7). However, unlike $\psi^0_1(S_{22.1})$ the estimator $\psi^0_2(S_{22.1})$ is not a function of $S_{22.1}$ alone.

5. DERIVATION OF THE ESTIMATORS $\varphi^{(1)}(S, V)$, $\hat{\Sigma}^{(1)}_*$ AND $\psi^{(1)}(S, V)$ OF Σ^{-1} UNDER THE LOSS $L^{(1)}$

In this section, we consider the estimation of Σ^{-1} under the loss function $L^{(1)}(\Sigma^{-1}, \hat{\Sigma}^{-1}) - \text{tr}\,(\hat{\Sigma}^{-1}\Sigma) - \log|\hat{\Sigma}^{-1}\Sigma| - p$, which we can denote by $L((B\hat{\Sigma}^{-1}B')^{-1})$ where $B \in G_1$ is such that $BB' = \Sigma^{-1}$. Proceeding as in Section 2, it is seen that the best lower triangular equivariant and hence minimax estimator

$$\varphi^{(1)}(S, V) = T'^{-1}\xi'^{-1}d^{*-1}\xi^{-1}T^{-1},$$

where T and ξ are lower triangular matrices satisfying

$$TT' = S \text{ and } \xi\xi' = \begin{bmatrix} I+U & O \\ O & I \end{bmatrix},$$

$U = T_{11}^{-1}VT'^{-1}_{11}$ and $d^* = E(Z'Z)^{-1}$ with the density of Z proportional to

$$|Z_{11}|^{m+n}|Z_{22}|^n \exp[-\tfrac{1}{2}\text{tr}\,Z_{11}Z'_{11} - \tfrac{1}{2}\text{tr}\,Z_{22}Z'_{22} - \tfrac{1}{2}\text{tr}\,Z_{21}(I - \xi'^{-1}_{11}U\xi'^{-1}_{11})Z'_{21}]\,\nu(dZ)$$

As evaluating d^* general p and q is difficult, take $p=2$, $q=1$; and $p=3$, $q=1$, For $p=2$, $q=1$ we know from Section 2 that Z^2_{11}, Z^2_{22} and Z_{21} are independently χ^2_{m+n}, χ^2_{n-1} and $N(0, \beta)$ respectively so that

$$d^{*-1} = \begin{pmatrix} M-2 & 0 \\ 0 & (M-2)(n-3)/(M-2+\beta) \end{pmatrix}$$

where $M = m+n$ and $\beta = 1 + V/s_{11}$. The estimator $\varphi^{(1)}(S, V)$ is

$$T'^{-1}\begin{pmatrix} (M-2)/\beta & 0 \\ 0 & (M-2)(n-3)/(M-2+\beta) \end{pmatrix}T^{-1}.$$

For $p=3$, $q=1$, d^{*-1} is seen to be

$$\begin{bmatrix} (M-2)/2 & 0 & 0 \\ 0 & (M-2)(n-3)/(M-2+\beta) & 0 \\ 0 & 0 & \dfrac{(M-2)(n-3)(n-4)}{(n-2)(M-2+\beta)} \end{bmatrix}$$

so that $\varphi^{(1)}(S, V)$ is

$$T'^{-1}\begin{pmatrix} d^{(1)} & 0 \\ 0 & D^{(2)}h(S, V) \end{pmatrix}T^{-1},$$

where $d^{(1)} = (M-2)/\beta$, $D^{(2)} = \text{diag}\,(n-3, (n-3)(n-4)/(n-2))$, and $h(S, V) = (M-2)/(M-2+\beta)$.

We next obtain, for $p=3$, $q=1$, estimators $\hat{\Sigma}_*^{(1)}$ and $\psi^{(1)}(S, V)$ similar to $\hat{\Sigma}_1^*$ and $\psi_1(S, V)$. The estimator $\varphi^{(1)}(S, V)$ can be written as

$$\begin{bmatrix} s_{11}/d^{(1)} & S_{12}/d^{(1)} \\ S_{21}/d^{(1)} & S_{21}s_{11}^{-1}S_{12}/d^{(1)} + T_{22}(D^{(2)})^{-1}T'_{22}/h(S, V) \end{bmatrix}^{-1}$$

in which $T'^{-1}_{22}D^{(2)}T^{-1}_{22}$ is the best lower triangular equivariant estimator of $\Sigma^{-1}_{22.1}$ based on $S_{22.1}$ (Sharma and Krishnamoorthy[8]). Since $L^{(1)}$ is fully invariant, $\varphi^{(1)}_{\Gamma}(S, V) = \Gamma\varphi^{(1)}(\Gamma' S\Gamma, \Gamma'_{11}V\Gamma_{11})\Gamma'$ for any orthogonal Γ of the form $\begin{pmatrix} \Gamma_{11} & 0 \\ 0 & \Gamma_{22} \end{pmatrix}$. Then

$$\varphi^{(1)}_{\Gamma}(S, V) = \begin{bmatrix} s_{11}/d^{(1)} & S_{12}/d^{(1)} \\ S_{21}/d^{(1)} & S_{21}s_{11}^{-1}S_{12}/d^{(1)} + \\ & \Gamma_{22}T_{22\Gamma}(D^{(2)})^{-1}T'_{22\Gamma}\Gamma'_{22}/h(S, V) \end{bmatrix}^{-1} \quad \cdots(5.1)$$

where $T_{22\Gamma}T'_{22\Gamma} = \Gamma'_{22}S_{22.1}\Gamma_{22}$ with $T_{22\Gamma}$ lower triangular.

If $\Gamma_{22} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,

$$\varphi^{(1)}_{\Gamma}(S, V) = \begin{bmatrix} s_{11}/d^{(1)} & S_{12}/d^{(1)} \\ S_{21}/d^{(1)} & S_{21}s_{11}^{-1}S_{12}/d^{(1)} + U_{22}(D_0^{(2)})^{-1}U'_{22}/h(S, V) \end{bmatrix}^{-1}$$

$$= \varphi_0^{(1)}(S, V) \text{ (say)},$$

where U_{22} is an upper triangular matrix such that $U_{22}U'_{22} = S_{22.1}$, $D_0^{(2)} = \text{diag}\,((n-3)(n-4)/(n-2),\ n-3)$ and $U'^{-1}_{22}D_0^{(2)}U^{-1}_{22}$ can be described as the best upper triangular equivariant minimax estimator of $\Sigma^{-1}_{22.1}$ based on $S_{22.1}$.

Strict convexity of the loss function implies that, for $0 < \alpha < 1$, $\alpha\varphi^{(1)}(S, V) + (1-\alpha)\varphi_0^{(1)}(S, V)$ is better than $\varphi^{(1)}(S, V)$. We show that the best choice of α is $\frac{1}{2}$ so that $\hat{\Sigma}_*^{(1)}(S, V) = \frac{1}{2}(\varphi^{(1)}(S, V) + \varphi_0^{(1)}(S, V))$: Using the relation

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{21}A_{11}^{-1}A_{12} + B_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11}^{-1}(I + A_{12}(B_{22})^{-1}A_{21}A_{11}^{-1}) & -A_{11}^{-1}A_{12}B_{22}^{-1} \\ -B_{22}^{-1}A_{21}A_{11}^{-1} & B_{22}^{-1} \end{pmatrix}, \quad \cdots (5.2)$$

we have

$$\alpha\varphi^{(1)}(S, V) + (1-\alpha)\varphi_0^{(1)}(S, V)$$

$$= \begin{bmatrix} s_{11}/d^{(1)} & S_{12}/d^{(1)} \\ S_{21}/d^{(1)} & S_{12}s_{11}^{-1}S_{21}/d^{(1)} + \{\alpha T'^{-1}_{22}D^{(2)}T^{-1}_{22} + (1-\alpha)U'^{-1}_{22}D_0^{(2)} \\ & U'^{-1}_{22}\}^{-1}/h(S, V) \end{bmatrix}^{-1} \quad \cdots (5.3)$$

Hence, the risk of $\alpha\varphi^{(1)}(S, V)+(1-\alpha)\varphi_0^{(1)}(S, V)$ is

$$E \log (s_{11}/d^{(1)}) - E \log | \alpha T'^{-1}_{22} D^{(2)} T^{-1}_{22} + (1-\alpha) U'^{-1}_{22} D_0^{(2)} U^{-1}_{22}| - \log|\Sigma| - 2E \log h(S, V),$$

which again can be sent to be minimixed for $\alpha=\frac{1}{2}$, as we have proved in Theorem 3.2.

A convenient expression for $\hat{\Sigma}_*^{(1)}(S, V)$ is (5.3) with $\alpha=\frac{1}{2}$; the term $\frac{1}{2}(T'^{-1}_{22}D^{(2)}T^{-1}_{22}+U'^{-1}_{22}D_0^{(2)}U^{-1}_{22})$ in it can be described as the best average of the best lower triangular equivariant and the best upper triangular equivariant estimators of $\Sigma^{-1}_{22.1}$ based on $S_{22.1}$.

Like $\psi_1(S, V)$ we also define

$$\psi^{(1)}(S, V)=\int_{O(2)} \varphi^{(1)}_{\Gamma'}(S, V)\, dv(\Gamma_{22}), \qquad \dots (5.4)$$

where v is the invariant probability measure on the group $O(2)$ of orthogonal matrices of order 2×2 and $\varphi^{(1)}_{\Gamma'}$ is the estimator (5.1). Making use of (5.2), we get

$$\varphi^{(1)}_{\Gamma'}(S, V)=$$

$$\begin{bmatrix} \frac{d^{(1)}}{s_{11}}+\frac{S_{12}}{s_{11}^2}\eta^{(22.1)}_{\Gamma'_{22}}(S_{22.1})\,(S_{21}h(S, V) & -\frac{S_{12}}{s_{11}}\eta^{(22.1)}_{\Gamma'_{22}}(S_{22.1})h(S, V) \\ -\eta^{(22.1)}_{\Gamma'_{22}}(S_{22.1})\,\frac{S_{21}}{s_{11}}\,h(S. V) & \eta^{(22.1)}_{\Gamma'_{22}}(S_{22.2})h(S, V) \end{bmatrix}$$

where $\eta^{(22.1)}(S_{22.1})$ is the best lower triangular equivariant estimator of $\Sigma^{-1}_{22.1}$ based on $S_{22.1}$ (Sharma and Krishnamoorthy [ibid.]) and $\eta^{(22.1)}_{\Gamma'_{22}}(S_{22.1})=\Gamma_{22}\,\eta^{(22.1)}(\Gamma'_{22}S_{22.1}\,\Gamma_{22})\Gamma'_{22}$. Once again using the result in Sharma and Krishnamoorthy [ibid.] and (5.2),

$$\psi^{(1)}(S, V)=\begin{bmatrix} s_{11}/d^{(1)} & S_{12}/d^{(1)} \\ S_{21}/d^{(1)} & S_{22}s^{-1}_{11}S_{12}/d^{(1)}+(h(S, V)\psi^{(22.1)}(S_{22.1}))^{-1} \end{bmatrix}^{-1}$$

where $\psi^{(22.1)}(S_{22.1})=\frac{(n-3)(n-4)}{(n-2)}\,S^{-1}_{22.1}+\frac{2(n-3)}{(n-2)}\cdot\frac{S^{-\frac{1}{2}}_{22.1}}{\operatorname{tr} S^{\frac{1}{2}}_{22.1}}.$

Remarks 6.1. The admissibility of $\psi_i(S, V)$ $(i=1, 2)$ and $\psi^{(1)}(S, V)$ is an open question. We feel that they are inadmissible as they "correct" the best lower triangular equivariant estimator only in (i, j) elements, $i, j > 2$. In view of the fact that we are considering the group of orthogonal matrices, which is compact, if ψ's are inadmissible, one must be able to find a better orthogonal equivariant estimator,

6. RISKS OF THE MLE AND THE BEST LOWER TRIANGULAR EQUIVARIANT ESTIMATORS φ_1 AND $\varphi^{(1)}$

Anderson (1957) has derived the MLE of Σ for general p and q. For $p=3$ and $q=1$, it is

$$T\begin{bmatrix} (1+V/s_{11})/M & 0 & 0 \\ 0 & \frac{1}{n} & 0 \\ 0 & 0 & \frac{1}{n} \end{bmatrix} T' = TDT' \text{ (say)},$$

where $T \in G_1$, $TT'=S$ and $M=m+n$.

Its risk is

$$R_1(I, \hat{\Sigma}_{\text{mle}}) = E_I \text{ tr } (TDT') - E_I \log |S| - E_I \log |D| - 3$$

where E_I denotes the expectation when $\Sigma = I$. Let $T=(t_{ij})$, then t_{ij} $(i \geqslant j)$ are independent and t_{ii}^2 and t_{ij} $(i>j)$ have χ^2_{n-i+1} and $N(0, 1)$ distributions resspectively. Denoting the digamma function $\Gamma'(x)/\Gamma(x)$ by $\zeta(x)$, we can obtain

$$E_I \log |S| = 3 \log 2 + \sum_{j=1}^{3} \zeta(\tfrac{n-j+1}{2}),$$

so that

$$R_1(I, \hat{\Sigma}_{\text{mle}}) = E_I \text{ tr } TDS' - 3 \log 2 - \sum_{j=1}^{3} \zeta(\tfrac{n-j+1}{2}) + \log (n^2 M) - E_I \log (1+V/s_{11})$$
$$= 2\left(\frac{M-2}{M(n-2)} - \frac{1}{n}\right) - 3 \log 2 - \sum_{j=1}^{3} \zeta(\tfrac{n-j+1}{2}) - \zeta(M/2) + \zeta(n/2) + \log (n^2 M).$$

Similarly, under the loss L_2, the risk of $\hat{\Sigma}_{\text{mle}}$ can be seen to be

$$R_2(I, \hat{\Sigma}_{\text{mle}}) = \frac{M-2}{M(n+2)}\left[\frac{7n(M-4)+3M(n-1)(n-4)}{Mn(n-4)}\right] + \frac{n(M+2)-M(3n-4)}{Mn} + \frac{(n-1)(2n+3)}{n^2}.$$

Risks of φ_1 *and* $\varphi^{(1)}$ *for* $p=3$, $q=1$.

Consider the estimator $\varphi_1(S, V)$. Since $E_{\Sigma, \Sigma_{11}} \text{ tr } (\varphi_1(S, V)\, \Sigma^{-1}) = p$, the risk of φ_1 is

$$R_1(I, \varphi_1(S, V)) = -E_I \log |\varphi_1(S, V)|$$
$$= -E_I \log d_1^*(S, V) s_{11} - E_I \log |T_{22} D_2^* T'_{22}|,$$

where $d_1^*(S, V) = \beta/(M+2\beta)$, $\beta = 1+V/s_{11}$, T_{22} is lower triangular such that $T_{22}T'_{22}=S_{22.1}$ and $D_2^* = \text{diag}\ (\frac{1}{n}, \frac{1}{n-2})$. Since

$E_I \log |T_{22}T'_{22}| = E \log |S_{22.1}| = \sum_1^2 E \log \chi^2_{n-j}$.

$$R_1(I, \varphi_1(S, V)) = -3 \log 2 - \zeta(M/2) - \zeta((n-1)/2) - \zeta((n-2)/2) + \log (n(n-2)) + E_I \log (M+2\beta)$$

$E_I \log (M+2\beta)$ can be evaluated through numerical integration, we have done it for $n=10$ and $m=4$; its value is found to be 2.831472 so that $R_1(I, \varphi_1(S, V)) = {\cdot}61559$,

Similarly, the risk of $\varphi^{(1)}(S, V)$ is

$$R^{(1)}(I, \varphi^{(1)}(S, V)) = -E_I \log |\varphi^{(1)}(S, V)|$$
$$= -E_I \log d^{(1)} - 2E_I \log h(S, V) + E_I \log |T_{22}D^{(2)-1}T'_{22}|,$$

where $d^{(1)} = (M-2)/\beta$, $h(S, V) = (M-2)/(M-2-\beta)$ and $D^{(2)} = \text{diag}$ $(n-3, (n-3)(n-4)/(n-2))$, and is seen to equal $-3\log(M-2) + 2E_I \log (M-2+\beta) + \zeta(M/2) + \zeta(\frac{n-1}{2}) + \zeta((n-2)/2) + 3 \log 2 - \log ((n-3)^2 (n-4)/(n-2))$.

Once again, one can find $E \log (M-2+\beta)$ through numerical integration. Its value, for $n=10$ and $m=4$, is 2.60188 so that $R^{(1)}(I, \varphi^{(1)}(S, V)) = .74212$.

Incidentally, using concavity of $\log x$, one can obtain upper bounds of the risks of φ_1 and $\varphi^{(1)}$:

Since $E_1 \log (M+2\beta) \leqslant \log [M+2+2m/(n-2)]$
and $E_1 \log (M-2+\beta) \leqslant \log [M-1+m/(n-2)]$.

$$R_1(I, \varphi_1(S, V)) \leqslant -3 \log 2 - \zeta(M/2) - \zeta((n-1)/2) - \zeta((n-2)/2) + \log (n(n-2)) + \log [M+2+2m/(n-2)]$$

and

$$R^{(1)}(I, \varphi^{(1)}(S, V)) \leqslant -3 \log ((M-2)/2) + \zeta(M/2) + \zeta((n-1)/2) + \zeta((n-2)/2) - \log ((n-3)^2(n-4)/(n-2)) + 2 \log (M-1+m/(n-2)).$$

7. Exact comparison of $\hat{\Sigma}_1^*$ with $\varphi_1(S, V)$ & $\hat{\Sigma}_*^{(1)}$ with $\varphi^{(1)}(S,V)$

From (3.3), when $\Gamma = I$ and $\eta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1\end{pmatrix}$, we have

$$R_1(\Sigma, \hat{\Sigma}_1^*) - R_1(\Sigma, \varphi_1(S, V))$$
$$= E \log |T_{22}D_2^*T'_{22}| - E \log |\tfrac{1}{2}(T_{22}D_2^* T'_{22} + U_{21}D_2^{o*}U'_{22})|$$

Notice that $T_{22}D_2^* T'_{22}$ and $U_{22}D_2^{o*}U'_{22}$ are the best lower tringular equivariant and the best upper triangular equivariant estimators of $\Sigma_{22.1}$ based on $S_{22.1}$. Thus, the problem of comparing the estimators $\hat{\Sigma}_1^*$ and $\varphi_1(S, V)$ under L_1 is equivalent to the problem of comparing the estimators $\hat{\Sigma}_1$ and $\frac{1}{2}(\hat{\Sigma}_1 + \hat{\Sigma}_u)$, where $\hat{\Sigma}_1$ and Σ_u are the best lower triangular and the best upper triangular equivariant estimators of Σ, based on $S \sim W_{p-q}(\Sigma, n-q)$. We take $p=3$ and $q=1$, in which case the risk difference, after some simplification, is found to be

$$\log\left(1 - \frac{1}{n_1^2}\left(-E\left(1 - \frac{r^2}{n_1^2}\right)\right.\right.$$

$$= \log(1 - 1/n_1^2) + \sum_{k=0}^{\infty} \frac{1}{(k+1)n_1^{2(k+1)}} E(r^2)^{k+1}, \qquad \ldots \quad (7.1)$$

where $n_1 = n-1$ and $r = s_{12}/(s_{11}s_{22})^{1/2}$ is the sample corelation coefficient. Fujimoto (1982) has calculated the exact values of the risk difference between $\frac{1}{2}(\hat{\Sigma}_1 + \hat{\Sigma}_u)$ and $\hat{\Sigma}_1$ using the relation

$$E(r^{2j}) = \frac{(1-\rho^2)^{n/2}}{\Gamma(n/2)} \sum_{i=0}^{\infty} \frac{\Gamma(\frac{8}{2}+i+j)\,\{\Gamma(\frac{n}{2}+i)\}^2 \rho^{2i}}{i!\,\Gamma(\frac{1}{2}+i)\,\Gamma(\frac{n}{2}+i+j)}, \qquad \ldots \quad (7.2)$$

where $S \sim W_2(\Sigma, n)$, j is an integer $\geqslant 1$ and ρ is the correlation coefficient calculated from Σ. For our problem, $R_1(\Sigma, \hat{\Sigma}_1^*) - R_1(\Sigma, \varphi_1(S, V))$ is given by (7.1) where $n_1 = n-1$ and $E(r^{2j})$ is equal to the expression on the right hand side of (7.2) with n replaced by $n-1$.

We next find an expression of the risk difference of $\hat{\Sigma}_*^{(1)}$ and $\varphi^{(1)}(S, V)$. Clearly,

$$R^{(1)}(\Sigma, \hat{\Sigma}_*^{(1)}) - R^{(1)}(\Sigma, \varphi^{(1)}(S, V))$$

$$= E \log |T'^{-1}_{22} D^{(2)} T^{-1}_{21}| - E \log |\tfrac{1}{2}(T'^{-1}_{22} D^{(2)} T^{-1}_{22} + U'^{-1}_{22} D_0^{(2)} U^{-1}_{22})|,$$

where $T'^{-1}_{22} D^{(2)} T^{-1}_{22}$ and $U'^{-1}_{22} D_0^{(2)} U^{-1}_{22}$ are the best tower triangular and the best upper triangular equivariant estimators of $\Sigma^{-1}_{22.1}$ based on $S_{22.1} \sim W_{p-q}(\Sigma_{22.1}, n-q)$. Thus, the comparison between $\hat{\Sigma}_*^{(1)}$ and $\varphi^{(1)}(S, V)$ is the same as the comparison of $\frac{1}{2}(\hat{\Sigma}_1^{-1} + \hat{\Sigma}_u^{-1})$ and $\hat{\Sigma}_1^{-1}$ where $\hat{\Sigma}_1^{-1}$ and $\hat{\Sigma}_u^{-1}$ are the best lower triangular equivariant and the best upper triangular equivariant estimators of Σ^{-1} respectively based on $S \sim W_{p-q}(\Sigma, n-q)$. From Sharma and Krishnamoorthy (ibid), for $p=3$, $q=1$,

$$\hat{\Sigma}_1^{-1} = \begin{bmatrix} (n-3)s^{11} - \frac{2(n-3)}{(n-2)}\frac{(s^{12})^2}{s^{22}} & \frac{(n-3)(n-4)s^{12}}{(n-2)} \\ \frac{(n-3)(n-4)s^{12}}{(n-2)} & \frac{(n-3)(n-4)s^{22}}{(n-2)} \end{bmatrix}$$

$$\text{and } \hat{\Sigma}_u^{-1} = \begin{bmatrix} \frac{(n-3)(n-4)}{(n-2)}s^{11} & \frac{(n-3)(n-4)}{(n-2)}s^{12} \\ \frac{(n-3)(n-4)}{n-2)}s^{12} & (n-3)s^{22} - \frac{2(n-3)}{(n-2)}\frac{(s^{12})^2}{s^{12}} \end{bmatrix}$$

where $S^{-1} = (s^{ij})$.

After some calculation, we find the difference between the risk $\frac{1}{2}(\hat{\Sigma}_1^{-1} + \hat{\Sigma}_u^{-1})$ and the risk of $\hat{\Sigma}_1^{-1}$ and hence $R^{(1)}(\Sigma, \hat{\Sigma}_*^{(1)}) - R^{(1)}(\Sigma, \varphi^{(1)}(S, V))$ equal to

$$\log \frac{(n-2)(n-4)}{(n-3)^2} + \sum_{k=0}^{\infty} \frac{1}{(k+1)(n-3)^{2(k+1)}} E(r^2)^{k+1}, \quad \ldots \quad (7.3)$$

We have calculated the risk differences (7.1) and (7.3) for $n=$ 5, 10, 15, 20 and for different values of ρ : ρ in the original comparison problem is, in fact, the partial correlation coefficient $\rho_{23.1}$. The calculated values are shown in Tables 3(a) and 3(b). We dotice that $R_1(\Sigma, \hat{\Sigma}_1^*) \to R_1(\Sigma, \varphi_1(S, V))$ and $R^{(1)}(\Sigma, \hat{\Sigma}_*^{(1)}) \to R^{(1)}(\Sigma, \varphi^{(1)}(S, V))$ as $\rho_{23.1} \to 1$.

8. Monte-carlo study ($p=3$, $q=1$)

Since evaluating the risks of $\varphi_2(S, V)$, $\hat{\Sigma}_2^*$ and $\psi_2(S, V)$ under L_2, the risk of $\psi_1(S, V)$ under L_1 and the risk of $\psi^{(1)}(S, V)$ under $L^{(1)}$ is difficult, we do a Monte-Carlo study on the basis of samples of size 10,000 for different Σ, $n=10$ and $m=4$. Note that $\varphi_2(S, V)$ is lower triangular equivariant and $\varphi_2^0(S, V)$ is equivariant under G_A, the group of matrices A of the form

$$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & 0 & a_{33} \end{bmatrix}.$$

Hence $\hat{\Sigma}_2^*$ is equivariant under the group G_B of matrices B of the form

$$\begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & 0 & b_{33} \end{bmatrix} = \begin{bmatrix} b_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix} \text{(say)}.$$

Choosing $B_{21} = -\frac{\Sigma_{12}B'_{22}}{\sigma}$, it can be seen that it is enough to evaluate $R_i(\Sigma, \hat{\Sigma}_2^*)$ at

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \rho_{23.1} \\ 0 & \rho_{23.1} & 1 \end{bmatrix},$$

whereas for $\psi_i(S, V)$, Σ can be taken to be

$$\begin{bmatrix} 1 & c_1 & c_1 \\ c_1 & c_3 & 0 \\ c_2 & 0 & c_4 \end{bmatrix}$$

without loss of generality.

Tables 1.1—1.3 present the values of $R_1(\Sigma, \hat{\Sigma}_1^*)$ and $R_1(\Sigma, \psi_1(S, V))$ while Tables 2.1—2.3 present the values of $R_2(\Sigma, \hat{\Sigma}_2^*)$ add $R_2(\Sigma, \psi_2(S, F))$.

From these tables $\hat{\Sigma}_i^*$ and $\psi_i(S, V)$ $(i=1, 2)$ do not seem to be comparable.

Tables 1.1—1.3

$m=4$, $n=10$, $p=3$, $q=1$, $R_1(\Sigma, \hat{\Sigma}_{\text{mle}}) = .66130$, $R_1(\Sigma, \varphi_1(S, V)) = .61559$

Column (a) gives the values of $R_1(\Sigma, \hat{\Sigma}_1^*)$. Column (b) gives the values of $R_1(\Sigma, \psi_1(S, V))$.

Table 1.1. $\Sigma = \begin{bmatrix} 1 & c_1 & c_2 \\ c_1 & 1 & 0 \\ c_2 & 0 & 1 \end{bmatrix}$: $1 - c_1^2 - c_2^2 > 0$.

c_2 \ c_1	.1 (a)	.1 (b)	.2 (a)	.2 (b)	.3 (a)	.3 (b)	.5 (a)	.5 (b)
.2	.60410	.60335	.60411	.60334	.60413	.60334	.60419	.60338
.4	.60412	.60338	.60415	.60338	.60423	.60339	.60458	.60348
.6	.60414	.60353	.60427	.60355	.60452	.60360	.60567	.60392
.8	.60423	.60418	.60467	.60429	.60556	.60454	.60714	.60629
.9	.60442	.60534	.60558	.60570	.60799	.60656	—	—

Table 1.2. $\Sigma=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & c\\ 0 & c & 1\end{bmatrix}$.

c	(a)	(b)
0	.60410	.60334
.1	.60420	.60350
.2	.60447	.60362
.3	.60490	.60368
.4	.63553	.60393
.5	.60639	.60429
.6	.60756	.60478
.7	.60907	.60547
.8	.61095	.60648
.9	.61316	.60821
.99999	.61558	.61563

Table 1.3. $\Sigma=\begin{bmatrix}1 & 0 & 0\\ 0 & \sigma_{22} & 0\\ 0 & 0 & \sigma_{33}\end{bmatrix}$.

$(\sigma_{22}, \sigma_{33})$	(a)	(b)
(1, 1)	.60410	.60334
(2, 3)	.60410	.60341
(10, 100)	.60410	.60615
(20, 300)	.60410	.60702

Tables 2.1-2.3

$m=4$, $n=10$, $p=3$, $q=1$, $R_2(\Sigma, \hat{\Sigma}_{mle})=99143$,

$R_2(\Sigma, \varphi_2(S, V))=.81091$ (simulated). Column (a) gives the values of $R_2(\Sigma, \hat{\Sigma}_2^*)$ and column (b) gives the values of $R_2(\Sigma, \psi_2(S, V))$

Table 2.1. $\Sigma=\begin{bmatrix}1 & c_1 & c_2\\ c_1 & 1 & 0\\ c_2 & 0 & 1\end{bmatrix}$, $1-c_1^2-c_2^2>0$

c_2 \ c_1	.1 (a)	.1 (b)	.2 (a)	.2 (b)	.3 (a)	.3 (b)	.5 (a)	.5 (b)
.2	.80295	.80264	.80295	.80264	.80296	.80265	.80301	.80268
.4	.80295	.80266	.80298	.80267	.80304	.80268	.80329	.80276
.6	.80297	.80276	.80306	.80278	.80324	.80283	.80411	.80307
.8	.80303	.80320	.80336	.80329	.80403	.80347	.80732	.80471
.9	.80317	.80399	.80404	.80425	.80587	.80485	—	—

Table 2.2. $\Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 1 & c & 1 \end{bmatrix}$

c	a	b
0	.80295	.80265
.1	.80302	.80267
.2	.80323	.80274
.3	.80356	.80286
.4	.80403	.80303
.5	.80467	.80327
.6	.80551	.80360
.7	.80659	.80406
.8	.80789	.80474
.9	.80933	.80589
.99999	.81135	.81054

Table 2.3. $\Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{bmatrix}$.

$(\sigma_{22}, \sigma_{33}$	a	b
(1; 1)	.80295	.80265
(2, 3)	.80295	.80270
(10, 100)	.80295	.80469
(20, 300)	.80295	.80531

Table 3.1. Exact Risk Difference $R(\Sigma, \varphi_1(S, V)) - R(\Sigma, \hat{\Sigma}_1^*)$

$\rho_{23.1}$ \ n	5	10	15	20
0	.04866	.01103	.00475	.00262
.05	.04858	.01099	.00474	.00261
.1	.04834	.01088	.00471	.00258
.3	.04571	.00975	.00437	.00232
.5	.04009	.00738	.00368	.00179
.7	.02403	.00283	.00241	.00086
.9	.01365	.00087	.00039	.00007

Table 3.2. Exact Risk Difference $R(\Sigma, \varphi^{(1)}(S, V)) - R(\Sigma, \Sigma_*^{(1)})$ for $p=3$, $q=1$

$\rho_{28.1}$ \ n	5	10	15	20
0	.22081	.01831	.00640	.00327
.05	.22046	.01825	.00639	.00326
.1	.21942	.01807	.00633	.00323
,3	.20802	.01619	.00573	.00289
.5	.18351	.01227	.00469	.00224
.7	.11140	.00483	.00329	.00107
.9	.07056	.00096	.00072	.00008

ACKNOWLEDGEMENTS

The authors are thankful to Vishnu Kant and R. Venkatesh for their help in the numerical computation.

REFERENCES

Anderson, T. W. (1957) : Maximum likelihood estimates for a multivariate normal distribution when some observation are missing. *J. Amer. Statist. Assoc.* **52**, 200-203.

Eaton, M. L. (1970) : Some problems in covariance estimation. *Technical Report* No. 49. Department of Statistics, Stanford University.

Fujimoto, M. (1982) : Series expressions of exact risks of improved estimators for bivariate normal covariance matrix. To appear in *Ann. Inst. Statist. Math.*

James, W. and Stein, C. (1961) : Estimation with quadratic loss. *Proc. Fourth Berkeley Symposium on Math. Statist. and Prob.* University of California Press, Berkeley. **1**, 361-380.

Kiefer, J. (1957) : Invariance, minimax sequential estimation and continuous time processes. *Ann. Math. Statist.* **28**, 573-601.

Selliah, J. B. (1964) : Estimation and testing problems in Wishart distribution. *Technical Report No.* 10. Department of Statistics, Stanford University.

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

Sharma, D. and Krishnamoorthy, K. (1983) : Orthogonal equivariant minimax estimators of bivariate normal covariance matrix and precision matrix. *Calcutta Statist. Assoc. Bull.* **32**, 23-45.

Sylvan, M. (1969) : Estimation and hypothesis testing for Wishart matrices when part of covariance matrix is known. *Technical Report No.* 38. Department of Statistics, Stanford University.

Zidek, J. V. (1969) : A representation of Bayes invariant procedures in terms of Haar measure. *Ann. Inst. Statist. Math.* **21**, 291-308,