COMMUN. STATIST.-THEORY METH., 20(3), 757-770 (1991)

ESTIMATION OF NORMAL COVARIANCE AND PRECISION MATRICES WITH INCOMPLETE DATA

K. KRISHNAMOORTHY

Department of Statistics
Temple University
Philadelphia, PA 19122

Key Words and Phrases: Wishart distribution, loss function, minimax estimator.

ABSTRACT

Suppose that we have n independent observations from $N_p(0, \Sigma)$ and, in addition, we have m independent observations available on the first $q(q < p)$ coordinates. Assuming that X_i's and Y_i's are independent , we consider the problem of estimation of Σ and Σ^{-1} respectively under the loss functions $L(\Sigma, \hat{\Sigma}) = \text{tr}(\hat{\Sigma}\Sigma^{-1}) - \log|\hat{\Sigma}\Sigma^{-1}| - p$ and $L^1(\Sigma^{-1}, \hat{\Sigma}^{-1}) = \text{tr}(\hat{\Sigma}^{-1}\Sigma) - \log|\hat{\Sigma}^{-1}\Sigma| - p$. We propose some new estimators that dominate the best lower triangular invariant minimax estimator under the loss L. We also derive the best lower triangular invariant minimax estimator of Σ^{-1} under L^1 and suggest some estimators that dominate it.

1. INTRODUCTION

Suppose that $X_1, ..., X_n$ are independent observations from a p-variate normal population with a known mean vector μ and an unknown covariance matrix Σ, $N_p(\mu, \Sigma)$. The mean vector, without loss of generality, can be assumed to be null. Besides these n observations, we also have m independent observations $Y_1, ..., Y_m$ available from the $N_q(0, \Sigma_{11})$ population, where $q < p$, $m \geq q$ and

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

Σ_{11} is of order $q \times q$. We assume that X_i's and Y_i's are independent and Σ is positive definite. We are interested in estimation of Σ and Σ^{-1} from this incomplete data. In this setup, Anderson (1957) derived the maximum likelihood estimator $\hat{\Sigma}_{mle}$ of Σ. Eaton (1970) derived the best lower triangular invariant estimator of Σ under the entropy loss function

$$L(\Sigma, \hat{\Sigma}) - \text{tr}(\hat{\Sigma}\Sigma^{-1}) - \log|\hat{\Sigma}\Sigma^{-1}| - p \tag{1.1}$$

where $|.|$ denote the determinant. As the $\hat{\Sigma}_{mle}$ is also a lower triangular invariant and being different from the best one, it is inadmissible under the loss (1.1). Sharma and Krishnamoorthy (1985) proposed some estimators that dominate Eaton's (1970) estimator (under the loss (1.1)) for the special cases $p = 2, q = 1$ and $p = 3, q = 1$.

Although the best lower triangular invariant estimator is minimax, it has an unappealing feature that it changes under the permutations of the coordinates. Since we have only partial information on the last $p - q$ coordinates , we like an estimator to be invariant under the permutations of the first q coordinates and the last $p - q$ coordinates. A way out is to look for an estimator that is invariant under the group of transformations G_Γ, where a member of G_Γ is of the form $\begin{pmatrix} \Gamma_{11} & 0 \\ 0 & \Gamma_{22} \end{pmatrix}$, $\Gamma_{11} : q \times q$ and $\Gamma_{22} : (p - q) \times (p - q)$ are orthogonal matrices.

In Section 2, we first briefly outline the Eaton's method of deriving the best lower triangular inavariant minimax estimator. Then following the approach by Krishnamoorthy and Gupta (1989), we find two estimators, namely, ψ_1 and ψ_2 of Σ under the loss (1.1). The estimator ψ_1 is better than the Eaton's minimax estimator (hence minimax) and invariant under $G_\Gamma^o = \{\Gamma : \begin{pmatrix} I & 0 \\ 0 & \Gamma_{22} \end{pmatrix}, \Gamma\Gamma' = I\}$, for any p and q such that $p - q \geq 2$.

The estimator ψ_2 is invariant under G_Γ, but we do not know whether it is minimax or not. However, Monte-Carlo simulation study (in Section 4) indicates that ψ_2 is better than both Eaton's minimax estimator and ψ_1.

In Section 3, we derive the best lower triangular invariant minimax estimator of Σ^{-1} under the loss function

$$L^1(\Sigma^{-1}, \hat{\Sigma}^{-1}) = \mathrm{tr}(\hat{\Sigma}^{-1}\Sigma) - \log|\hat{\Sigma}^{-1}\Sigma| - p. \tag{1.2}$$

Also, estimators of Σ^{-1} that are analogous to ψ_1 and ψ_2 are obtained.

To indicate the nature of the risk improvement of the proposed estimators, we carry out a Monte-Carlo simulation study in Section 4. The study shows that the estimators that are invariant under G_Γ dominate others.

2. ESTIMATION OF Σ

Let $X_1, ..., X_n$ be independent observations from $N_p(0, \Sigma)$ and $Y_1, ..., Y_m$ be independent observations from $N_q(0, \Sigma_{11})$ where $\Sigma_{11} : q \times q$ is (1,1) partitioned matrix of Σ. Assume that X_i's and Y_i's are independent. Define $S = \sum_{i=1}^n X_i X_i'$ and $V = \sum_{i=1}^m Y_i Y_i' (n > p, m > q)$. Then $S \sim W_p(n, \Sigma)$ independently of $V \sim W_q(m, \Sigma_{11})$ where $W_p(n, \Sigma)$ denotes the Wishart distribution with n degrees of freedom and parameter matrix Σ. Consider the problem of estimation of Σ under a nonsingular invariant loss function $L^*(\Sigma, \hat{\Sigma})$.

Let G_l denote the group of lower triangular matrices. Consider the transformation

$$\begin{gathered} S \to ASA', \Sigma \to A\Sigma A' \\ V \to A_{11} V A_{11}', \Sigma_{11} \to A_{11}\Sigma_{11}A_{11}' \\ \hat{\Sigma} \to A\hat{\Sigma}A' \end{gathered} \tag{2.1}$$

where $A \in G_l$ and $A_{11} : q \times q$ is (1,1) partitioned matrix of A. It can be easily checked that the estimation problem is invaraint under the transformation of (2.1). According to a theorem due to Stein (see, Zidek (1969)), the best lower triangular invariant estimator is formal Bayes with respect to a right invariant prior on G_l. Since G_l is solvable, it is also minimax. Using these results, Eaton(1970) obtained minimax estimator and is described below.

Write $\Sigma = \Theta\Theta', \Theta \in G_l$ and partition Θ as $\Theta = \begin{pmatrix} \Theta_{11} & 0 \\ \Theta_{21} & \Theta_{22} \end{pmatrix}$ where Θ_{11} is of order $q \times q$. Factorize S as $S = TT'$, where T is the lower triangular matrix with positive diagonal elements, and partition T as Θ. Since the loss L^* is nonsingular invaraint, we can write $L^*(\Sigma, \hat{\Sigma}) = L^*(I, \Theta^{-1}\hat{\Sigma}\Theta'^{-1})$, and the posterior loss of $\hat{\Sigma}$ is given by

$$\begin{aligned} H(\hat{\Sigma}) &= c(S,V)\int_{G_l} L^*(I, \Theta^{-1}\hat{\Sigma}\Theta'^{-1})|\Theta\Theta'|^{-n/2}|\Theta_{11}\Theta'_{11}|^{-m/2} \\ &\times \exp\{-\frac{1}{2}[\mathrm{tr}\Theta^{-1}S\Theta'^{-1} + \mathrm{tr}\Theta_{11}^{-1}V\Theta_{11}'^{-1}]\}\nu_r(d\Theta) \end{aligned} \tag{2.2}$$

where ν_r is right invariant measure on G_l and $c(S,V)$ is a constant that depends on S and V. We want to find $\hat{\Sigma}$ that minimizes (2.2). Let $\eta = \Theta^{-1}T$ so that $\eta_{11} = \Theta_{11}^{-1}T_{11}$. Then (2.2) becomes

$$\begin{aligned} H_1(b) &= c_1(S,V)\int_{G_l} L^*(I, \eta b\eta')|\eta|^n|\eta_{11}|^m \\ &\times \exp\{-\frac{1}{2}[\mathrm{tr}\eta\eta' + \mathrm{tr}\eta_{11}T_{11}^{-1}VT_{11}'^{-1}\eta_{11}']\}\nu_l(d\eta), \end{aligned} \tag{2.3}$$

where ν_l is left invariant measure on G_l and $b = T^{-1}\hat{\Sigma}T'^{-1}$. Now, let $\xi = \begin{pmatrix} \xi_{11} & 0 \\ 0 & I \end{pmatrix}$ be the positive lower triangular square root of $\begin{pmatrix} I + T_{11}^{-1}VT_{11}'^{-1} & 0 \\ 0 & I \end{pmatrix}$ such that $\xi_{11}\xi_{11}' = I + T_{11}^{-1}VT_{11}'^{-1}$. Then (2.3) can be written as

$$\begin{aligned} H_1(b) &= c_1(S,V)\int_{G_l} L^*(I, \eta b\eta')|\eta|^n|\eta_{11}|^m \\ &\times \exp\{-\frac{1}{2}[\mathrm{tr}\eta\xi\xi'\eta' - \mathrm{tr}\eta_{21}T_{11}^{-1}VT_{11}'^{-1}\eta_{21}']\}\nu_l(d\eta). \end{aligned} \tag{2.4}$$

Again setting $Z = \eta\xi$ (so that $Z_{21} = \eta_{21}\xi_{11}$) and $\delta = \xi^{-1}b\xi'^{-1}$, it then suffices to minimize

$$\begin{aligned} H_2(\delta) &= c_o\int_{G_l} L^*(I, Z\delta Z')|Z|^n|Z_{11}|^m \\ &\times \exp\{-\frac{1}{2}\mathrm{tr}Z_{11}Z_{11}' - \frac{1}{2}\mathrm{tr}Z_{22}Z_{22}' - \frac{1}{2}\mathrm{tr}Z_{21}QZ_{21}'\}\nu_l(dZ), \end{aligned} \tag{2.5}$$

where $Q = I - \xi_{11}^{-1}T_{11}^{-1}VT_{11}'^{-1}\xi_{11}'^{-1} = \xi_{11}^{-1}\xi_{11}'^{-1}$. The constant c_o is chosen to make (2.5) the expectation of $L^*(I, Z\delta Z')$. Let δ^* be the matrix for δ that minimizes (2.5). Since $\delta^* = \xi^{-1}b\xi'^{-1} = \xi^{-1}T^{-1}\hat{\Sigma}T'^{-1}\xi'^{-1}$, the $\hat{\Sigma}$ that minimizes (2.2) is given by

$$\hat{\Sigma} = T \begin{pmatrix} \xi_{11} & 0 \\ 0 & I \end{pmatrix} \delta^{\star} \begin{pmatrix} \xi_{11}' & 0 \\ 0 & I \end{pmatrix} T'. \tag{2.6}$$

Noticing that

$$\nu_l(dZ) = (\prod_{i=1}^{p} z_{ii}^{i})^{-1} dZ = (\prod_{i=1}^{q} z_{ii}^{i})^{-1} dZ_{11} (\prod_{i=q+1}^{p} z_{ii}^{i})^{-1} dZ_{22} dZ_{21}, \tag{2.7}$$

where z_{ii} denotes the i th diagonal elements of Z(i=1,...,p), it follows from (2.5) that Z_{11}, Z_{22} and Z_{21} are all independent with $Z_{11}Z_{11}' \sim W_q(n+m, I)$, $Z_{22}Z_{22}' \sim W_{p-q}(n-q, I)$ and $Z_{21}' \sim N(0, I_{p-q} \otimes Q^{-1})$. Using these results, for any given nonsingular inavriant loss function $L^{\star}$, the minimizing δ for (2.5) may be evaluated.

In particular, for the loss in (1.1), the best lower triangular invaraiant estimator of Σ is given by

$$\phi(S,V) = T \begin{pmatrix} \xi_{11} & 0 \\ 0 & I \end{pmatrix} D^{-1} \begin{pmatrix} \xi_{11}' & 0 \\ 0 & I \end{pmatrix} T' \tag{2.8}$$

where

$$D = (E(Z'Z)) = \begin{pmatrix} D_1 + (p-q)Q^{-1} & 0 \\ 0 & D_2 \end{pmatrix}. \tag{2.9}$$

In (2.9), $D_1 : q \times q$ and $D_2 : (p-q) \times (p-q)$ are diagonal matrices whose i th diagonal elements, respectively, are $d_{1i} = m+n+q+1-2i (i = 1, 2, .., q)$ and $d_{2i} = n+p-2q+1-2i (i = 1, 2, ..p-q)$. More explicitly, using some matrix algebra, the estimator $\phi(S,V)$ in (2.8) can be written as

$$\phi(S,V) = \begin{pmatrix} \phi_{11}(S,V) & \phi_{11}(S,V)S_{11}^{-1}S_{12} \\ S_{21}S_{11}^{-1}\phi_{11}(S,V) & S_{21}S_{11}^{-1}\phi_{11}(S,V)S_{11}^{-1}S_{12} + T_{22}D_2^{-1}T_{22}' \end{pmatrix} \tag{2.10}$$

where

$$\phi_{11}(S,V) = [(T_{1v}D_1^{-1}T_{1v}')^{-1} + (p-q)S_{11}^{-1}]^{-1} \tag{2.11}$$

and $T_{1v}$is the unique lower triangular square root of $(S_{11}+V)$.

In the following we derive some estimators that dominate $\phi(S,V)$.

Improved estimators of Σ

Let $G_{\Gamma} = \{\Gamma : \begin{pmatrix} \Gamma_{11} & 0 \\ 0 & \Gamma_{22} \end{pmatrix}, \Gamma\Gamma' = I\}$. Then for any Γ, independent of

(S, V), in G_Γ, it can be shown that

$$\phi_\Gamma(S, V) = \Gamma\phi(\Gamma' S\Gamma, \Gamma'_{11} V \Gamma_{11})\Gamma' \tag{2.12}$$

is also minimax with constant risk equal to the risk of $\phi(S, V)$ (for a proof, see Sharma and Krishnamoorthy (1985)).

Write $S_{22} - S_{21}S_{11}^{-1}S_{12} = S_{22.1} = R_2 L_2 R_2'$, where $R_2 R_2' = I$ and $L_2 = \text{diag}(l_{21}, ..., l_{2(p-q)})$ with $l_{21} > ... > l_{2(p-q)} > 0$. Let $R_o = \begin{pmatrix} I & 0 \\ 0 & R_2 \end{pmatrix}$. As in Krishnamoorthy and Gupta (1989), if we let $\Gamma = R_o$ in (2.12), then the estimator

$$\psi_1(S, V) = \phi_{R_o}(S, V) = R_o\phi(R_o' S R_o, V)R_o' \tag{2.13}$$

can be seen to be invariant under $G_\Gamma^o = \{\Gamma : \begin{pmatrix} I & 0 \\ 0 & \Gamma_{22} \end{pmatrix}, \Gamma\Gamma' = I\}$ and we prove in the following theorem that it dominates $\phi(S, V)$.

Theorem 2.1. For $p - q \geq 2$, the estimator $\psi_1(S, V)$ dominates $\phi(S, V)$ under the loss (1.1).

Proof. Let $T_{R_o} = \begin{pmatrix} T_{11} & 0 \\ R_2' T_{21} & L_2^{\frac{1}{2}} \end{pmatrix}$ so that

$$T_{R_o} T_{R_o}' = R_o' S R_o = \begin{pmatrix} S_{11} & S_{12}R_2 \\ R_2' S_{21} & L_2 + R_2' S_{21} S_{11}^{-1} S_{12} R_2 \end{pmatrix}.$$

Using these relations, it follows from (2.13) and (2.8) that

$$\psi_1(S, V) = R_o T_{R_o} \begin{pmatrix} \xi_{11} & 0 \\ 0 & I \end{pmatrix} D^{-1} \begin{pmatrix} \xi_{11}' & 0 \\ 0 & I \end{pmatrix} T_{R_o}' R_o'$$

and after some calculations, we can write

$$\psi_1(S, V) = \begin{pmatrix} \phi_{11}(S, V) & \phi_{11}(S, V)S_{11}^{-1}S_{12} \\ S_{21}S_{11}^{-1}\phi_{11}(S, V) & S_{21}S_{11}^{-1}\phi_{11}(S, V)S_{11}^{-1}S_{12} + R_2 D_2^{-1} L_2 R_2' \end{pmatrix}, \tag{2.14}$$

where $\phi_{11}(S, V)$ is given in (2.11). Thus, from (2.10), (2.14) and using the fact that $|T_{22}D_2^{-1}T_{22}'| = |R_2 L_2 D_2^{-1} R_2'|$, the risk difference can be written as

$$\begin{aligned} R(\Sigma, \psi_1(S, V)) - R(\Sigma, \phi(S, V)) &= EL(\Sigma, \psi_1(S, V)) - EL(\Sigma, \phi(S, V)) \\ &= E\text{tr}[(R_2 D_2^{-1} L_2 R_2' - T_{22} D_2^{-1} T_{22}')\Sigma_{22.1}^{-1}]. \end{aligned} \tag{2.15}$$

The *rhs* is the risk difference between Dey and Srinivasan's (1985) orthogonally invariant minimax estimator and Stein's (1961) minimax estimator of $\Sigma_{22.1}$ based on $S_{22.1} \sim W_{p-q}(n-q, \Sigma_{22.1})$ under the loss (1.1). Thus, for $n > p+1$ and $p-q \geq 2$, the *rhs* of (2.15) is less than zero by Theorem 3.1 of Dey and Srinivasan (1985).

Although the estimator $\psi_1(S,V)$ dominates the best lower triangular invariant estimator $\phi(S,V)$, it is not invariant under the permutations of the first q components. So we look for a minimax estimator that is invariant under G_Γ.

Let $(S_{11}+V) = R_1 L_1 R_1'$ where $R_1 R_1' = I$ and $L_1 = \text{diag}(l_{11}, ..., l_{1q})$ with $l_{11} > ... > l_{1q} > 0$. Define $R = \begin{pmatrix} R_1 & 0 \\ 0 & R_2 \end{pmatrix}$. Then it can be easily checked that the estimator

$$\psi_2(S,V) = \phi_R(S,V) = R\phi(R'SR, R_1'VR_1)R' \tag{2.16}$$

is invariant under the group G_Γ. Some calculations show that the estimator $\psi_2(S,V)$ can be written more explicitly as

$$\psi_2(S,V) = \begin{pmatrix} \psi_{2(11)}(S,V) & \psi_{2(11)}(S,V)S_{11}^{-1}S_{12} \\ S_{21}S_{11}^{-1}\psi_{2(11)}(S,V) & S_{21}S_{11}^{-1}\psi_{2(11)}(S,V)S_{11}^{-1}S_{12} + R_2 D_2^{-1} L_2 R_2' \end{pmatrix} \tag{2.17}$$

where $\psi_{2(11)}(S,V) = [R_1 D_1 L_1^{-1} R_1' + (p-q)S_{11}^{-1}]^{-1}$. Monte-Carlo simulation study(in Section 4) indicates that $\psi_2(S,V)$ dominates both $\phi(S,V)$ and $\psi_1(S,V)$. However, proving the dominance results theoretically seems to be difficult.

Remark 2.1.

For $p-q \geq 3$, the estimator $\psi_i(S,V)$ can further be improved as follows: Recall that $R_2 D_2^{-1} L_2 R_2'$ is a minimax estimator (of $\Sigma_{22.1}$ under the loss (1.1)) based on $S_{22.1} \sim W_{p-q}(n-q, \Sigma_{22.1})$. Let $\psi_i^o(S,V)$ be equal to $\psi_i(S,V)$(i=1,2) with $R_2 D_2^{-1} L_2 R_2'$ replaced by its superior (for $p-q \geq 3$) given in Theorem 3.2 of Dey and Srinivasan (1985). Then, it can be easily checked that $\psi_i^o(S,V)$ dominates $\psi_i(S,V)$, $i = 1, 2$, under the loss (1.1).

3. ESTIMATION OF Σ^{-1}

In this section, we first derive the best lower triangular invariant estimator of Σ^{-1} under the loss function (1.2).

It can be easily seen that the estimation problem is invariant under the transformations (2.1) and the Bayes estimator with respect to ν_r on G_l is given by

$$\phi^1(S,V) = T'^{-1}\begin{pmatrix} \xi_{11}'^{-1} & 0 \\ 0 & I \end{pmatrix} \Delta \begin{pmatrix} \xi_{11}^{-1} & 0 \\ 0 & I \end{pmatrix} T^{-1} \tag{3.1}$$

where Δ is the value of δ which minimizes

$$\begin{aligned} H^1(\delta) &= c_o \int_{G_l} L^1(I, Z'^{-1}\delta Z^{-1})|Q|^{\frac{p-q}{2}}|Z_{11}|^{m+n}|Z_{22}|^n \\ &\times \exp\{-\frac{1}{2}tr[Z_{11}Z_{11}' + Z_{22}Z_{22}' + Z_{21}QZ_{21}']\}\nu_l(dZ) \end{aligned} \tag{3.2}$$

($\nu_l(dZ)$ is given in (2.7)). The minimizing Δ can be shown to be $[E(Z^{-1}Z'^{-1})]^{-1}$ and is computed as (see appendix)

$$\Delta = [E(Z^{-1}Z'^{-1})]^{-1} = \begin{pmatrix} \Delta_1 & 0 \\ 0 & g\Delta_2 \end{pmatrix}, \tag{3.3}$$

where $\Delta_1 = \text{diag}(\delta_{11}, ..., \delta_{1q}), \delta_{1i} = \frac{(m+n-i)(m+n-i-1)}{m+n-1}, (i = 1, ..., q)$, $\Delta_2 = \text{diag}(\delta_{21}, ..., \delta_{2(p-q)}), \delta_{2i} = \frac{(n-q-i)(n-q-i-1)}{n-q-1}, (i = 1, ..., p-q)$ and $g = 1/(\text{tr}(\Delta_1^{-1}Q^{-1}) + 1)$.

Since $\phi^1(S,V)$ is the best lower triangular invariant estimator, it is minimax with constant risk. As the maximum likelihood estimator (see, Anderson (1957))

$$\hat{\Sigma}_{mle}^{-1} = T'^{-1}\begin{pmatrix} \xi_{11}'^{-1} & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} (m+n)I & 0 \\ 0 & nI \end{pmatrix}\begin{pmatrix} \xi_{11}^{-1} & 0 \\ 0 & I \end{pmatrix} T^{-1} \tag{3.4}$$

of Σ^{-1} is also a lower triangular invariant and being different from $\phi^1(S,V)$ it is inadmissible under the loss (1.2).

Explicitly, the estimator (3.1) can be written as $\phi^1(S,V)$

$$= \begin{pmatrix} T_{1v}'^{-1}\Delta_1 T_{1v}^{-1} + gS_{11}^{-1}S_{12}T_{22}'^{-1}\Delta_2 T_{22}^{-1}S_{21}S_{11}^{-1} & -gS_{11}^{-1}S_{12}T_{22}'^{-1}\Delta_2 T_{22}^{-1} \\ -gT_{22}'^{-1}\Delta_2 T_{22}^{-1}S_{21}S_{11}^{-1} & gT_{22}'^{-1}\Delta_2 T_{22}^{-1} \end{pmatrix} \tag{3.5}$$

where $T_{1v}T_{1v}' = (S_{11} + V)$, $T_{22}T_{22}' = S_{22.1}$ and $g = [tr(\Delta_1^{-1}Q^{-1}) + 1]^{-1}$.

We next develop estimators of Σ^{-1} that are analogous to ψ_1 and ψ_2.

For any Γ, independent of (S,V), in G_Γ it can be seen that the estimator

$$\phi_\Gamma^1(S,V) = \Gamma\phi^1(\Gamma' S\Gamma, \Gamma_{11}' V\Gamma_{11})\Gamma' \tag{3.6}$$

is minimax with constant risk equal to that of $\phi^1(S,V)$.

Let, as in Section 2, $(S_{11}+V)$ and $S_{22.1}$ have spectral decompositions $R_1L_1R_1'$ and $R_2L_2R_2'$ respectively. For $R_o = \begin{pmatrix} I & 0 \\ 0 & R_2 \end{pmatrix}$, $\psi^{(1)}(S,V)$

$$\begin{aligned} &= R_o\phi^1(R_o'SR_o, V)R_o' \\ &= \begin{pmatrix} T_{1v}'^{-1}\Delta_1 T_{1v}^{-1} + gS_{11}^{-1}S_{12}\psi_{22}^{(2)}(S_{22.1})S_{21}S_{11}^{-1} & -gS_{11}^{-1}S_{12}\psi_{22}^{(2)}(S_{22.1}) \\ -g\psi_{22}^{(2)}(S_{22.1})S_{21}S_{11}^{-1} & g\psi_{22}^{(2)}(S_{22.1}) \end{pmatrix} \end{aligned} \tag{3.7}$$

where $\psi_{22}^{(2)}(S_{22.1}) = R_2\Delta_2 L_2^{-1}R_2'$, is invariant under G_Γ^o. For $R = \begin{pmatrix} R_1 & 0 \\ 0 & R_2 \end{pmatrix}$, the estimator

$$\psi^{(2)}(S,V) = R\phi(R'SR, R_1'VR_1)R'$$

can be shown to be invaraint under G_Γ. Explicitly, the estimator $\psi^{(2)}(S,V)$

$$= \begin{pmatrix} R_1\Delta_1 L_1^{-1}R_1' + hS_{11}^{-1}S_{12}\psi_{22}^{(2)}(S_{22.1})S_{21}S_{11}^{-1} & -hS_{11}^{-1}S_{12}\psi_{22}^{(2)}(S_{22.1}) \\ -h\psi_{22}^{(2)}(S_{22.1})S_{21}S_{11}^{-1} & h\psi_{22}^{(2)}(S_{22.1}) \end{pmatrix} \tag{3.8}$$

where $h = [\mathrm{tr}(R_1\Delta_1^{-1}L_1R_1'S_{11}^{-1}) + 1]^{-1}$. Although Monte-Carlo simulation study in the next section indicates that the estimator $\psi^{(1)}(S,V)$ *dominates* $\phi^1(S,V)$ and $\psi^{(2)}(S,V)$ dominates both $\psi^{(1)}(S,V)$ and $\phi^1(S,V)$, we do not know whether $\psi^{(2)}(S,V)$ and $\psi^{(1)}(S,V)$ are minimax or not.

4. MONTE-CARLO SIMULATION STUDY

Since evaluating the exact risk expressions of the proposed estimators seems to be difficult, we estimate them based on 2000 S's generated from $W_p(n,\Sigma)$ and 2000 V's generated from $W_q(n-q,\Sigma_{11})$.

We use the following expressions to estimate the risks.

As $\phi(S,V)$ and $\phi^1(S,V)$ are lower triangular invariant, the risks of them are independent of Σ and also as they are the best ones, it can be checked that $E\text{tr}(\phi(S,V)\Sigma^{-1}) = E\text{tr}(\phi^1(S,V)\Sigma) = p$. Further, using the result that for $W \sim W_p(n,\Sigma)$, $E\log|W\Sigma^{-1}| = p\log(2) + \sum_{i=1}^{p}\varsigma(\frac{n-i+1}{2})$, where $\varsigma(.)$ denotes the digamma function, we get

$$\begin{aligned}
R(I,\phi(S,V)) &= -E\log|\phi(S,V)| \\
&= -E\log|T_{22}D_2^{-1}T_{22}'| - E\log|\phi_{11}(S,V)| \\
&= \sum_{i=1}^{p-q}\log(d_{2i}) - (p-q)\log(2) - \sum_{i=1}^{p-q}\varsigma(\frac{n-q-i+1}{2}) \\
&- E\log|\phi_{11}(S,V)|, \\
R(\Sigma,\psi_1(S,V)) &= E\text{tr}(R_2D_2^{-1}L_2R_2'\Sigma_{22.1}^{-1}) - (p-q) + R(I,\phi(S,V)), \\
R(\Sigma,\psi_2(S,V)) &= E\text{tr}(\psi_2(S,V)\Sigma^{-1}) - (p-q)\log(2) - \sum_{i=1}^{p-q}\varsigma(\frac{n-q-i+1}{2}) \\
&+ \sum_{i=1}^{p-q}\log(d_{2i}) - E\log|\psi_{2(11)}(S,V)\Sigma_{11}^{-1}| - p.
\end{aligned}$$

Similarly, we can write

$$\begin{aligned}
R^1(I,\phi^1(S,V)) &= -\sum_{i=1}^{q}\log(\delta_{1i}) - \sum_{i=1}^{p-q}\log(\delta_{2i}) + p\log(2) \\
&+ \sum_{i=1}^{q}\varsigma(\frac{n+m-i+1}{2}) + \sum_{i=1}^{p-q}\varsigma(\frac{n-q-i+1}{2}) \\
&+ (p-q)E\log(\text{tr}(\Delta_1^{-1}Q^{-1})+1), \\
R^1(I,\psi^{(1)}(S,V)) &= E\text{tr}(\psi^{(1)}(S,V)\Sigma^{-1}) - p + R^1(I,\phi^1(S,V)), \\
R^1(I,\psi^{(2)}(S,V)) &= E\text{tr}(\psi^{(2)}(S,V)\Sigma^{-1}) - \sum_{i=1}^{q}\log(\delta_{1i}) - \sum_{i=1}^{p-q}\log(\delta_{2i}) \\
&+ p\log(2) + \sum_{i=1}^{q}\varsigma(\frac{n+m-i+1}{2}) + \sum_{i=1}^{p-q}\varsigma(\frac{n-q-i+1}{2}) \\
&+ (p-q)E\log(\text{tr}(R_1\Delta_1^{-1}L_1R_1'S_{11}^{-1})+1) - p.
\end{aligned}$$

Finally, the exact risks of $\hat{\Sigma}_{mle}$ (under the loss (1.1)) and $\hat{\Sigma}_{mle}^{-1}$ (under the loss (1.2)) can be evaluated as

$$\begin{aligned}
R(I,\hat{\Sigma}_{mle}) &= (p-q)[\frac{q(m+n-q-1)}{(m+n)(n-q-1)} - \frac{q}{n}] + q\log(\frac{m+n}{n}) + p\log(\frac{n}{2}) \\
&- \sum_{i=1}^{q}\varsigma(\frac{m+n-i+1}{2}) - \sum_{i=1}^{p-q}\varsigma(\frac{n-q-i+1}{2}),
\end{aligned}$$

Table 1: Risks of $\hat{\Sigma}_{mle}$, $\phi(S,V)$, $\hat{\Sigma}_{mle}^{-1}$ and $\phi^1(S,V)$.

n=20, m=5, p=5, q=2 and $\Sigma = I$

$R(I,\hat{\Sigma}_{mle})$	$R(I,\phi)$	$R^1(I,\hat{\Sigma}_{mle}^{-1})$	$R^1(I,\phi^1)$
0.8056(exact)	0.7484(.004)	1.2676(exact)	0.8478(.005)
0.8105(.006)*		1.2734(.006)*	
n=20, m=5, p=5, q=1			
0.8212(exact)	0.7721(.006)	1.3019(exact)	0.8705(.007)
0.8205(.005)*		1.3028(.004)*	

* estimated.

[The numbers in parantheses represent the estimated values of the standard error. The matrix Σ_i(i=1,...,4) is of the form $\begin{pmatrix} (J-I)\rho+I & 0 \\ 0 & \text{diag}(\sigma_{33},\sigma_{44},\sigma_{55}) \end{pmatrix}$, where $\Sigma_1 : \rho = .7, \Sigma_{22} = \text{diag}(3,3,3); \Sigma_2 : \rho = .9, \Sigma_{22} = I; \Sigma_3 : \rho = .7, \Sigma_{22} = \text{diag}(2,3,10); \Sigma_4 : \rho = .8, \Sigma_{22} = \text{diag}(10,5,1)$. The matrix Σ_i(i=5,6,7) is of the form $\begin{pmatrix} \text{diag}(\sigma_{11},\sigma_{22}) & 0 \\ 0 & (J-I)\rho+I \end{pmatrix}$, where $\Sigma_5 : \Sigma_{11} = I, \rho = .9; \Sigma_6 : \Sigma_{11} = \text{diag}(2,10), \rho = .7$. The matrix $\Sigma_7 : \Sigma_{11} = I, \Sigma_{12} = \begin{pmatrix} .8 & .9 & .8 \\ .7 & .6 & .9 \end{pmatrix}$, $\Sigma_{22} = \text{diag}(10,10,10)$.]

$$
\begin{aligned}
R^1(I,\hat{\Sigma}_{mle}^{-1}) &= \frac{n(n-1)(p-q)}{(n-p-1)(n-q-1)} + \frac{(m+n)q}{m+n-q-1} - q\log(\frac{m+n}{n}) \\
&- p\log(\frac{n}{2}) + \sum_{i=1}^{q}\varsigma(\frac{m+n-i+1}{2}) + \sum_{i=1}^{p-q}\varsigma(\frac{n-q-i+1}{2}) - p.
\end{aligned}
$$

Table 1 gives the risks of $\hat{\Sigma}_{mle}$, ϕ (under the loss (1.1)),$\hat{\Sigma}_{mle}^{-1}$ and ϕ^1 (under the loss (1.2)) for $n = 20, m = 5; p = 5, q = 2$ and $p = 5, q = 1$. We give both the exact and the estimated risks of MLEs in order to understand the 'credibility' of the simulation estimates.

For the values of $n = 20, m = 5, p = 5$ and $q = 2$ in Table 2(a) we present the estimated values of $R(\Sigma,\psi_1)$, $R(\Sigma,\psi_2)$, $R(\Sigma,\psi^{(1)})$ and $R(\Sigma,\psi^{(2)})$ for $\Sigma = \text{diag}(c_1,...,c_p)$, $\Sigma = (J-I)\rho + I$, where $0 < \rho < 1$ and J is the $p \times p$ matrix of ones, and for some patterned Σ's. These values indicate that the improvement of these new estimators over the MLEs and the minimax estimators is maximum when $\Sigma \doteq kI$, where k is a positive constatnt. This also seems to be true when $\Sigma_{11} \doteq k_1 I$, $\Sigma_{22} \doteq k_2 I$ and $\Sigma_{12} = 0$, where

Table 2(a): Risks of $\psi_1(S,V)$, $\psi_2(S,V)$, $\psi^{(1)}(S,V)$ and $\psi^{(2)}(S,V)$.

n=20, m=5, p=5 and q=2

$\Sigma = diag(c_1, ..., c_p)$	$R(\Sigma,\psi_1)$	$R(\Sigma,\psi_2)$	$R^1(\Sigma,\psi^{(1)})$	$R^1(\Sigma,\psi^{(2)})$
(1,1,1,1,1)	0.69(.008)	0.68(.007)	0.75(.008)	0.73(.008)
(.5,.4,.3,.2,.1)	0.72(.009)	0.71(.008)	0.79(.007)	0.77(.009)
(11,12,10,11,10)	0.69(.008)	0.68(.009)	0.75(.008)	0.73(.008)
(20,15,10,2,1)	0.74(.008)	0.74(.008)	0.82(.006)	0.80(.009)
(30,20,10,5,1)	0.75(.008)	0.74(.006)	0.82(.010)	0.81(.009)
(2,1.8,15,15,14)	0.69(.006)	0.68(.004)	0.75(.008)	0.73(.007)
(1,1,10,10,10)	0.69(.005)	0.68(.006)	0.75(.007)	0.73(.008)
(10,10,1,1,1)	0.69(.006)	0.68(.006)	0.75(.006)	0.73(.007)
(10,10,10,1,1)	0.74(.007)	0.73(.006)	0.81(.010)	0.79(.007)
(10,1,1,1,1)	0.69(.008)	0.69(.008)	0.75(.004)	0.75(.005)
(10,1,10,1,1)	0.74(.004)	0.74(.005)	0.81(.006)	0.80(.007)
(10,10,10,10,10)	0.69(.005)	0.68(.006)	0.75(.005)	0.73(.007)
$\Sigma = (J - I)\rho + I$				
$\rho = .1$	0.70(.007)	0.68(.007)	0.75(.008)	0.73(.008)
.3	0.70(.009)	0.70(.007)	0.76(.009)	0.75(.006)
.8	0.72(.006)	0.71(.007)	0.80(.007)	0.78(.006)
.9	0.74(.006)	0.73(.010)	0.82(.007)	0.82(.007)
Σ_1	0.69(.008)	0.69(.006)	0.75(.007)	0.74(.008)
Σ_2	0.69(.007)	0.68(.006)	0.75(.007)	0.74(.006)
Σ_3	0.74(.007)	0.73(.006)	0.82(.006)	0.81(.008)
Σ_4	0.74(.006)	0.74(.005)	0.81(.008)	0.80(.005)
Σ_5	0.74(.005)	0.73(.005)	0.81(.006)	0.79(.007)
Σ_6	0.74(.009)	0.74(.005)	0.81(.010)	0.81(.008)
Σ_7	0.70(.008)	0.69(.008)	0.76(.009)	0.73(.008)

Table 2(b): Risks of $\psi_1(S,V)$ and $\psi^{(1)}(S,V)$.

n=20, m=5, p=5, q=1

$(\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{44}, \sigma_{55}, \sigma_{12}, \sigma_{13}, \sigma_{14}, \sigma_{15})$	$R(\Sigma,\psi_1)$	$R^1(\Sigma,\psi^{(1)})$
(1,1,1,1,1,0,0,0,0)	0.60(.008)	0.67(.008)
(1,10,10,10,10,0,0,0,0)	0.62(.009)	0.60(.006)
(10,1,1,1,1,0,0,0,0)	0.60(.008)	0.68(.008)
(3,8,7,8,8,0,0,0,0)	0.61(.007)	0.70(.006)
(8,3,2,3,3,0,0,0,0)	0.62(.007)	0.73(.006)
(10,1,10,1,10,0,0,0,0)	0.70(.005)	0.80(.007)
(30,20,10,5,1,0,0,0,0)	0.75(.005)	0.82(.008)
(5,5,5,5,5,2,1,2,3)	0.67(.008)	0.76(.007)
(10,10,10,10,10,2,1,2,3)	0.61(.007)	0.70(.007)
(30,20,10,5,1,2,1,2,3)	0.76(.009)	0.82(.008)

k_1 and k_2 are some positive constatnts. Also, we can observe that the risks of the estimators ψ_i and $\psi^{(i)}$(i=1,2) approach the minimax risks when $\Sigma = \text{diag}(\sigma_{11},...,\sigma_{pp})$ and σ_{ii}'s are much spread out or $\Sigma = (J-I)\rho + I$ and $\rho \to 1$.

For $q = 1$ case, we note that the estimator $\psi_2(\psi^{(2)})$ reduces to $\psi_1(\psi^{(1)})$. Further, as ψ_1 and $\psi^{(1)}$ are invariant under G_T^o, without loss of generality, we estimate their risks at Σ's of the form $\begin{pmatrix} \sigma_{11} & \sigma'_{21} \\ \sigma_{21} & \text{diag}(\sigma_{22},...,\sigma_{pp}) \end{pmatrix}$, where $\sigma'_{21} = (\sigma_{12},...,\sigma_{1p})$. For $p = 5$ and $q = 1$, the estimated values of $R(\Sigma, \psi_1)$ and $R^1(\Sigma, \psi^{(1)})$ are presented in Table 2(b). We here observe that the maximum relative improvement of $\psi_1(\psi^{(1)})$ over $\phi(\phi^1)$ is higher than that of $p = 5$ and $q = 2$ case. Apart from this, both the cases are almost similar.

Over all the study indicates that, as an estimator of Σ, ψ_2 is preferable to others under the loss (1.1), and $\psi^{(2)}$ is preferable to others as an estimator of Σ^{-1} under the loss (1.2).

APPENDIX

To compute $E(Z^{-1}Z'^{-1})$, write $Z^{-1} = \begin{pmatrix} Z_{11}^{-1} & 0 \\ -Z_{22}^{-1}Z_{21}Z_{11}^{-1} & Z_{22}^{-1} \end{pmatrix}$ where Z_{11}^{-1} is of order $q \times q$. Then,

$$E(Z^{-1}Z'^{-1}) = E\begin{pmatrix} Z_{11}^{-1}Z_{11}'^{-1} & -Z_{11}^{-1}Z_{11}'^{-1}Z_{21}'Z_{22}'^{-1} \\ -Z_{22}^{-1}Z_{21}Z_{11}^{-1}Z_{11}'^{-1} & Z_{22}^{-1}(I + Z_{21}Z_{11}^{-1}Z_{11}'^{-1}Z_{21}')Z_{22}'^{-1} \end{pmatrix}. \tag{A.1}$$

Recall that Z_{11}, Z_{22} and Z_{21} are all independent with $Z_{11}Z_{11}' \sim W_q(n + m, I)$, $Z_{22}Z_{22}' \sim W_{p-q}(n-q, I)$ and $Z_{21}' \sim N(0, I_{p-q} \otimes Q^{-1})$, where I_{p-q} denotes the identity matrix of order $(p-q) \times (p-q)$. Therefore, it is immediate that $E(Z_{11}^{-1}Z_{11}'^{-1}Z_{21}'Z_{22}'^{-1}) = 0$ and from Krishnamoorthy and Gupta(1989), we get $E(Z_{11}^{-1}Z_{11}'^{-1}) = \Delta_1^{-1}$ where $\Delta_1 = \text{diag}(\delta_{11},...,\delta_{1q}), \delta_{1i} = \frac{(m+n-i)(m+n-i-1)}{m+n-1}(i = 1,...,q)$. We next compute

$$\begin{aligned} E(Z_{22}^{-1}(I + Z_{21}Z_{11}^{-1}Z_{11}'^{-1}Z_{21}')Z_{22}'^{-1}) &= EE[(Z_{22}^{-1}(I + Z_{21}Z_{11}^{-1}Z_{11}'^{-1}Z_{21}')Z_{22}'^{-1})|Z_{22}, Z_{21}] \\ &= EE[(Z_{22}^{-1}(I + Z_{21}\Delta_1^{-1}Z_{21}')Z_{22}'^{-1})|Z_{22}] \end{aligned}$$

$$= E[Z_{22}^{-1}(I + I\mathrm{tr}(\Delta_1^{-1}Q^{-1}))Z_{22}'^{-1})]$$
$$= \Delta_2^{-1}(1 + tr(\Delta_1^{-1}Q^{-1}))$$

where $\Delta_2 = \mathrm{diag}(\delta_{2i}, ..., \delta_{2(p-q)}), \delta_{2i} = \frac{(n-q-i)(n-q-i-1)}{n-q-1}, (i = 1, ..., p-q)$. Thus, substituting these expectations in (A.1) and after taking inverse, we get (3.3).

ACKNOWLEDGEMENT

The author is thankful to Nariaki Sugiura and the referee for their valuable comments and suggestions.

BIBLIOGRAPHY

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

DEY, D.K. and SRINIVASAN, C. (1985). Estimation of covariance matrix under Stein's loss. *Ann. Statist.* 13, 1581-1592.

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

KRISHNAMOORTHY, K. and GUPTA, A.K. (1989). Improved minimax estimation of a normal precision matrix. *Canadian J. Statist.* 17, No. 1, 91-102.

SHARMA, D. and KRISHNAMOORTHY, K. (1985). Improved minimax estimators of normal covariance and precision matrices from incomplete samples. *Calcutta Statist. Assoc. Bull.* 34, 23-42.

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

Received February 1990; Revised December 1990.

Recommended by Nariaki Sugiura, University of Tsukuba, JAPAN.

Refereed by Y. Konno, Ishinomaki Senshu University, JAPAN.