

Estimation of the Scale Matrix of a Multivariate T-Model under Entropy Loss¹

ANWAR H. JOARDER

Department of Econometrics, Monash University, Clayton, Victoria 3168, Australia

MIR M. ALI

Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario, Canada N6A 5B7

Abstract: This paper deals with the estimation of the scale matrix of a multivariate t -model with unknown location vector and scale matrix to improve upon the usual estimators based on the sample sum of product matrix. The well-known results of the estimation of the scale matrix of the multivariate normal model under the assumption of entropy loss function have been generalized to that of a multivariate t -model.

Key Words: Multivariate t -model, scale matrix, estimation of scale matrix, mixture of distributions.

1 Introduction

The estimation of the covariance matrix of the multivariate normal distribution is well-known. However, the assumption of the normality is not met in many real world situations. Many economic and business data e.g. stock return data exhibit fat tailed distributions. The multivariate t -model accomodates thin tailed as well as fat tailed distributions. Blattberg and Gonedes (1974) have assessed the suitability of multivariate t -model for stock returns. Zellner (1976) analyzed stock return data by considering a regression model under the assumption that errors have a multivariate t -model. This model has attracted considerable attention in the recent literature (Lange et. al. 1989). In this paper we consider the estimation of the scale matrix of a multivariate t -model under entropy loss function.

We consider the multivariate t -model given by

$$g(x) = \frac{|A|^{-N/2}}{K(v, Np)\pi^{Np/2}} \left(1 + \sum_{j=1}^N (x_j - \theta)'(vA)^{-1}(x_j - \theta) \right)^{-(v+Np)/2} \quad (1.1)$$

¹ The paper is based on the first author's unpublished Ph.D. dissertation 'Estimation of the Scale Matrix of a Multivariate T-model', University of Western Ontario, Canada. Present address: School of Mathematics and Statistics, The University of Sydney, NSW 2006, Australia.

where $x_j = (x_{1j}, x_{2j}, \dots, x_{pj})'$, $j = 1, 2, \dots, N$, and $x = (x_1, x_2, \dots, x_N)$ while the normalizing constant $K(v, Np) = v^{Np/2} \Gamma(v/2) / (\Gamma((v + Np)/2))$. The p -dimensional column vector θ is the location parameter and the $p \times p$ positive definite matrix A is the scale matrix while the shape parameter v is a known scalar. We remark that the above multivariate t -model can also be written as a scale mixture of multivariate normal distributions given by

$$g(x) = \int_0^\infty \frac{|\tau^2 A|^{-N/2}}{(2\pi)^{Np/2}} \exp\left(-\frac{1}{2} \sum_{j=1}^N (x_j - \theta)' (\tau^2 A)^{-1} (x_j - \theta)\right) h(\tau) d\tau \quad (1.2)$$

where τ^{-2} has a gamma distribution with mean 1 and variance $2/v$ i.e.

$$h(\tau) = \frac{2\tau^{-(v+1)}}{(2/v)^{v/2} \Gamma(v/2)} \exp\left(-\frac{v}{2\tau^2}\right). \quad (1.3)$$

This means that $X_j|\tau$ has a multivariate normal distribution $N_p(\theta, \tau^2 A)$. As $v \rightarrow \infty$, the multivariate t -model converges to the multivariate normal model given by

$$f(x) = \frac{|A|^{-N/2}}{(2\pi)^{Np/2}} \exp\left(-\frac{1}{2} \sum_{j=1}^N (x_j - \theta)' A^{-1} (x_j - \theta)\right).$$

In this paper we develop estimators for the scale matrix A , the mean vector θ being unknown, under the entropy loss function

$$L(u(A), A) = \text{tr}(A^{-1}u(A)) - \ln|A^{-1}u(A)| - p \quad (1.4)$$

where $u(A)$ is any estimator of A based on the sample sum of product matrix $A = \sum_{j=1}^N (X_j - \bar{X})(X_j - \bar{X})'$ with $\bar{X} = (\bar{X}_1, \bar{X}_2, \dots, \bar{X}_p)'$ and $\bar{X}_i = N^{-1} \sum_{j=1}^N X_{ij}$, ($i = 1, 2, \dots, p$).

Based on the form of the likelihood function, the entropy loss function has been suggested in the literature by James and Stein (1961) and is sometimes known as the Stein loss function. Some important features of the entropy loss function are that it is zero if the estimator $u(A)$ equals the parameter A , positive when $u(A) \neq A$ and invariant under translation as well as under natural group of transformations of covariance matrices. Moreover, the loss function approaches infinity as the estimator approaches a singular matrix or when one or more elements (or one or more characteristics roots) of the estimator approaches infinity. This means gross underestimation is penalized just as heavily as gross overestimation. The interested reader may be referred to Anderson (1984, pp 273–274) for a detailed discussion.

In estimating A by $u(A)$, we consider the risk function $R(u(A), A) = E[L(u(A), A)]$. An estimator $u_2(A)$ of A will be said to dominate another estimator $u_1(A)$ of A if for all A belonging to the class of positive definite matrices the inequality $R(u_2(A), A) \leq R(u_1(A), A)$ holds and the inequality $R(u_2(A), A) < R(u_1(A), A)$ holds for at least one A .

In Section 2, we consider estimators which are multiples of the sample sum of product matrix i.e. estimators of the type cA , $c > 0$, where A is the sum of product matrix based on the multivariate t -model. In Section 3, we consider estimators based on the lower triangular decomposition of the sample sum of product matrix i.e. estimators of the type $TA'T'$ where T is an upper triangular matrix such that $A = TT'$ and A is an arbitrary positive definite diagonal matrix with diagonal elements δ_i ($i = 1, 2, \dots, p$), in analogy with the work, in the context of the multivariate normal distribution, by James and Stein (1961). In Section 4, we consider estimators based on the spectral decomposition of the sample sum of product matrix i.e. estimators of the form $RAMR'$ where A has the spectral decomposition $A = RMR'$, with $M = \text{diag}(m_1, m_2, \dots, m_p)$, ($m_1 \geq m_2 \dots \geq m_p$), once again in analogy with the works of Stein (1975) and Dey and Srinivasan (1985).

Dey and Srinivasan (1985) develop improved as well as minimax estimators for the scale matrix A of the multivariate normal distribution along the line of James and Stein (1961) and Stein (1975). Some of these results have been extended in the present work to the case when the underlying observations have the multivariate t -model rather than the multivariate normal model.

In order to avoid future digressions of a trivial nature we prove some lemmas that will be needed in the sequel.

Lemma 1.1: Consider the multivariate t -model given by (1.1) with $\nu > 2$. Then (i) $A|\tau = \tau^2 W \sim \mathcal{W}_p(n, \tau^2 A)$ and (ii) $E(A) = \nu n A / (\nu - 2)$, $\nu > 2$ where τ has an inverted gamma distribution given by (1.3) and W has the usual Wishart distribution $\mathcal{W}_p(n, A)$, $n = N - 1$.

Proof: The proof of part (i) follows from Sutradhar and Ali (1989) by noting the mixture representation of the multivariate t -model given by (1.2). The result in (ii) follows from $E(A) = E[E(A)|\tau] = E(n\tau^2 A)$ since

$$E(\tau^2) = \nu n / (\nu - 2) . \quad (1.5)$$

Lemma 1.2: Consider the multivariate t -model given by (1.1) for $\nu > 2$. Then the following two identities hold:

$$E[\ln(|A|)] = E[\ln(|W|)] + 2pE(\ln \tau) \quad (1.6)$$

$$\text{and} \quad E[\ln(|A^{-1}A|)] = \sum_{i=1}^p E[\ln(\chi_{n+1-i}^2)] + 2pE(\ln \tau) \quad (1.7)$$

where τ has an inverted gamma distribution given by (1.3) and W has the usual Wishart distribution $\mathcal{W}_p(n, A)$.

Proof: Since $|A| = \tau^{2p}|W|$, the result in (1.6) is immediate. By the same argument, we have

$$E[\ln|A^{-1}A|] = E[\ln(|W|/|A|)] + 2pE(\ln \tau) .$$

The expression in (1.7) then follows by virtue of the fact that $|W|/|A| \sim \prod_{i=1}^p \chi_{n+1-i}^2$ (see e.g. Muirhead, 1982, pp 85, 100).

Lemma 1.3: Consider the triangular decomposition $A = TT'$ where A is the sum of product matrix based on the multivariate t -model with $\nu > 2$ and T is a lower triangular matrix. Then the following identities hold:

$$E[\text{tr}(A^{-1}TAT')] = E[\text{tr}(A^{-1}U\Delta U')]\nu/(\nu - 2) \quad (1.8)$$

$$\text{and} \quad E[\ln(|A^{-1}TAT'|)] = E[\ln(|A^{-1}U\Delta U'|)] + 2pE[\ln(\tau)] , \quad (1.9)$$

where U is a lower triangular matrix such that $W = UU'$ and Δ is positive definite diagonal matrix while τ has an inverted gamma distribution given by (1.3).

Proof: Under the transformation $X_j = \theta + \tau Z_j$, ($j = 1, 2, \dots, N$), the sum of product matrix $A = TT'$ becomes

$$\tau^2 \sum_{j=1}^N (Z_j - \bar{Z})(Z_j - \bar{Z})' = \tau^2 UU' = \tau^2 W$$

so that

$$\begin{aligned}
E[\text{tr}(A^{-1}TAT')] &= \int_{X \in \mathcal{H}^{Np}} \text{tr}(A^{-1}TAT') \int_0^\infty \frac{|\tau^2 A|^{-N/2}}{(2\pi)^{Np/2}} \\
&\quad \times \exp\left(-\frac{1}{2} \sum_{j=1}^N (x_j - \theta)'(\tau^2 A)^{-1}(x_j - \theta)\right) h(\tau) d\tau dX \\
&= \int_0^\infty \int_{Z \in \mathcal{H}^{Np}} E[\text{tr}(A^{-1}\tau UAU')] \frac{\tau^{-Np}|A|^{-N/2}}{(2\pi)^{Np/2}} \\
&\quad \times \exp\left(-\frac{1}{2} \sum_{j=1}^N z_j' A^{-1} z_j\right) h(\tau) \tau^{Np} dZ d\tau \\
&= E[\text{tr}(A^{-1}UAU')] v/(v-2)
\end{aligned}$$

proving the result in (1.8). The proof for the result in (1.9) runs exactly along the same lines as above.

Lemma 1.4: Consider the spectral decomposition $A = RMR'$ where A is the sum of product matrix based on the multivariate t -model with $v > 2$. Then the following results hold:

$$E[\ln(|A^{-1}RD^*MR'|)] = \sum_{i=1}^p E[\ln(\chi_{n+1-i}^2)] + \sum_{i=1}^p \ln(d_i^*) + 2pE(\ln \tau) \quad (1.10)$$

$$\text{and } E[\text{tr}(A^{-1}RD^*MR')] = 2E\left[\sum_{i=1}^p \sum_{t=i+1}^p \frac{d_i l_i - d_t l_t}{l_i - l_t}\right] + (n-p+1) \sum_{i=1}^p d_i \quad (1.11)$$

where l_i 's and m_i 's ($i = 1, 2, \dots, p$) are the characteristic roots of W and A respectively, d_i^* , ($i = 1, 2, \dots, p$) and d_i , ($i = 1, 2, \dots, p$) are the diagonal elements of the diagonal matrices D^* and D respectively with

$$d_i^* = \frac{v-2}{v}(n+p+1-2i)^{-1} = \frac{v-2}{v}d_i, \quad (i = 1, 2, \dots, p). \quad (1.12)$$

Proof: It is easily checked that

$$\begin{aligned}
E[\ln(A^{-1}RD^*MR')] &= \ln(|A^{-1}|) + E[\ln(|\phi(M)|)] + \ln(|RR'|) \\
&= \ln(|A^{-1}|) + E[\ln(|D^*M|)] \\
&= E[\ln(|A^{-1}A|)] + \sum_{i=1}^p \ln(d_i^*)
\end{aligned}$$

so that by the use of (1.7) we have (1.10).

Next under the transformation $X_j = \theta + \tau Z_j$, ($j = 1, 2, \dots, N$), the sum of product matrix $A = RMR'$ becomes

$$\tau^2 \sum_{j=1}^N (Z_j - \bar{Z})(Z_j - \bar{Z})' = R\tau^2 \text{diag}(l_1, l_2, \dots, l_p)R' = R\tau^2 LR' = \tau^2 W$$

so that $m_i = \tau^2 l_i$, ($i = 1, 2, \dots, p$) and then it is readily verified that

$$\begin{aligned} & E[\text{tr}(A^{-1}RD^*MR')] \\ &= \int_{X \in \mathbb{R}^{Np}} \text{tr}(A^{-1}D^*M) \int_0^\infty \frac{|\tau^2 A|^{-N/2}}{(2\pi)^{Np/2}} \\ & \quad \times \exp\left(\frac{-1}{2} \sum_{j=1}^N (x_j - \theta)'(\tau^2 A)^{-1}(x_j - \theta)\right) h(\tau) d\tau dX \\ &= E(\tau^2) \int_{Z \in \mathbb{R}^{Np}} \text{tr}(A^{-1}RD^*LR') \frac{|A|^{-N/2}}{(2\pi)^{Np/2}} \exp\left(\frac{-1}{2} \sum_{j=1}^N z_j' Az_j\right) dZ \\ &= \int_{Z \in \mathbb{R}^{Np}} \text{tr}(A^{-1}RDLR') \frac{|A|^{-N/2}}{(2\pi)^{Np/2}} \exp\left(\frac{-1}{2} \sum_{j=1}^N z_j' Az_j\right) dZ \end{aligned}$$

where the last step follows by the use of (1.5) and (1.12).

The result in (1.12) then follows from Dey and Srinivasan (1985, p 1583) by noting

$$\begin{aligned} & \int_{Z \in \mathbb{R}^{Np}} \text{tr}(A^{-1}R\phi(L)R') \frac{|A|^{-N/2}}{(2\pi)^{Np/2}} \exp\left(\frac{-1}{2} \sum_{j=1}^N z_j' Az_j\right) dZ \\ &= 2E \left[\sum_{i=1}^p \sum_{t=i+1}^p \frac{\phi_i(L) - \phi_t(L)}{l_i - l_t} + 2 \sum_{i=1}^p \frac{\partial \phi_i(L)}{\partial l_i} + (n - p - 1) \sum_{i=1}^p \frac{\phi_i(L)}{l_i} \right] \end{aligned}$$

where $\phi(L) = \text{diag}(\phi_1(L), \phi_2(L), \dots, \phi_p(L))$, with $\phi_i(L) = d_i l_i$, $i = 1, 2, \dots, p$.

2 Estimators Based on the Multiples of the Sample Sum of Product Matrix

The scale matrix A of the multivariate normal distribution is usually estimated by cW where $c > 0$ and W is the usual Wishart matrix. It is well known (see e.g.

Muirhead, 1982, p 129) that under the entropy loss function, the best estimator (smallest risk) of the scale matrix of the multivariate normal distribution, of the form cW , is given by W/n .

In this section we consider estimators of the form cA , where $c > 0$, for the scale matrix A of the multivariate t -model and find optimum value of c for which the risk function of the estimator under the entropy loss function is minimized. The result is summarized in the form of the following theorem.

Theorem 2.1: Consider the multivariate t -model given by (1.2) with $v > 2$. Then under the entropy loss function the unbiased estimator of A , namely $\tilde{A} = (v - 2)A/(vn)$ has the smallest risk among the class of estimators cA , for $c > 0$ and the corresponding minimum risk is given by

$$R(\tilde{A}, A) = p \ln(n) - \sum_{i=1}^p E[\ln(\chi_{n+1-i}^2)] + p \ln[v/(v - 2)] - 2pE(\ln \tau) \quad (2.1)$$

where τ has an inverted gamma distribution given by (1.3).

Proof: The risk function of the estimator cA is given by

$$\begin{aligned} R(cA, A) &= E[\text{tr}(A^{-1}cA) - \ln(|A^{-1}cA|) - p] \\ &= c \text{tr}[A^{-1}E(A)] - p \ln(c) - E[\ln(|A^{-1}A|)] - p . \end{aligned}$$

Then it follows from Lemmas 1.1 and 1.2 that

$$\begin{aligned} R(cA, A) &= c \text{tr}[A^{-1}vnA/(v - 2)] - p \ln(c) \\ &\quad - \sum_{i=1}^p E[\ln(\chi_{n+1-i}^2)] - 2pE(\ln \tau) - p \\ &= \frac{vnpc}{v - 2} - \sum_{i=1}^p E[\ln(\chi_{n+1-i}^2)] - p \ln(c) - p - 2pE(\ln \tau) . \end{aligned} \quad (2.2)$$

Now the risk function given by (2.2) minimizes at $c = (v - 2)/(vn)$ and the corresponding estimator is given by $\tilde{A} = (v - 2)A/(vn)$. It then follows from (2.2), by putting $c = (v - 2)/(vn)$, that the risk function of the estimator $\tilde{A}$ is given by (2.1).

3 Estimators Based on a Triangular Decomposition of the Sample Sum of Product Matrix

Following James and Stein (1961), we propose estimators of the form TAT' where T is a lower triangular matrix such that the sample sum of product matrix A has the decomposition $A = TT'$ and Δ an arbitrary positive definite diagonal matrix. We find the optimum value of Δ for which the risk function of the estimator TAT' under the entropy loss function is minimized and denote it by D^* . The resulting estimator $A^* = TD^*T$ dominates the unbiased estimator $\tilde{A}$. The dominance behaviour is presented in the following theorem.

Theorem 3.1: Under the entropy loss function, the estimator $A^* = TD^*T'$ where T is a lower triangular matrix such that $A = TT'$ and $D^* = \text{diag}(d_1^*, d_2^*, \dots, d_p^*)$ with d_i^* ($i = 1, 2, \dots, p$) defined by (1.12), has the smallest risk among the class of estimators TAT' where Δ belongs to the class of all positive definite diagonal matrices, and the risk function of the estimator A^* is given by

$$\begin{aligned} R(A^*, A) &= \sum_{i=1}^p \ln(n+1+p-2i) - \sum_{i=1}^p E[\ln(\chi_{n+1-i}^2)] \\ &\quad + p \ln[v/(v-2)] - 2pE(\ln \tau) \end{aligned} \quad (3.1)$$

with τ having an inverted gamma distribution given by (1.3). Furthermore, A^* dominates the unbiased estimator $\tilde{A} = (v-2)A/(vn)$.

Proof: It follows from Lemma 1.3 that the risk function of the estimator TAT' is given by

$$\begin{aligned} R(TAT', A) &= E[\text{tr}(A^{-1}TAT')] - E[\ln(|A^{-1}TAT'|)] - p \\ &= E[\text{tr}(A^{-1}U\Delta U')]v/(v-2) - E[\ln(|A^{-1}U\Delta U'|)] \\ &\quad - 2pE(\ln \tau) - p. \end{aligned} \quad (3.2)$$

Then following Muirhead (1982, pp 130–132), it can be proved that the risk function given by (3.2) does not depend on Δ and that

$$\begin{aligned} R(TAT', A) &= -v \sum_{i=1}^p \delta_i(n+1+p-2i)/(v-2) - 2pE(\ln \tau) \\ &\quad - \left[\sum_{i=1}^p \ln(\delta_i) + \sum_{i=1}^p E[\ln(\chi_{n+1-i}^2)] \right] - p. \end{aligned} \quad (3.3)$$

This attains its minimum value of

$$R(A^*, A) = -2pE(\ln \tau) - \sum_{i=1}^p \ln d_i^* - \sum_{i=1}^p E[\ln(\chi_{n+1-i}^2)] \quad (3.4)$$

when $\delta_i = d_i^*$, ($i = 1, 2, \dots, p$) defined by (1.12). Then by putting the value of d_i^* from (1.12) we get the risk function of A^* as given by (3.1). The risk function of the unbiased estimator $\tilde{A}$ has already been calculated in (2.1) and consequently the risk difference is given by

$$R(A^*, A) - R(\tilde{A}, A) = \sum_{i=1}^p \ln \left(\frac{n+1+p-2i}{n} \right).$$

The righthand side of the above equation is negative which follows by considering the arithmetic mean and geometric mean inequality of the numbers $n+1, n+1, n+3, n+3, n+5, n+5, \dots, n+p-(2p-1), n+p-(2p-1)$ and then raising each side of the resulting inequality to its p -th power, dividing both sides of the resulting expression by n^p and finally taking logarithm in both sides. Hence it follows that the estimator A^* dominates the unbiased estimator $\tilde{A}$.

4 Estimators Based on the Spectral Decomposition of the Sample Sum of Product Matrix

As mentioned earlier, the sum of product matrix A can be decomposed as $A = RMR'$. We consider, following Stein (1975) and Dey and Srinivasan (1985), estimators of the form $R\phi(M)R'$ where $\phi(M) = \text{diag}(\phi_1(M), \phi_2(M), \dots, \phi_p(M))$ and $\phi_i(M)$ ($i = 1, 2, \dots, p$) is a function of the characteristic roots $m_1, m_2, \dots, m_p$ of the Wishart matrix A satisfying $\phi_i(M) > 0$. The main result is presented in the form of the following theorem.

Theorem 4.1: Let $\hat{A} = R\phi(M)R$ be an estimator for A where A has the spectral decomposition $A = RMR'$, with $\phi(M) = D^*M$. Then under the entropy loss function given by (1.4), the estimator $\hat{A} = RD^*MR'$ dominates the estimator $A^* = TD^*T'$.

Proof: The risk function of the estimator $\hat{A}$ is given by

$$\begin{aligned}
R(\hat{A}, A) &= E[\text{tr}(A^{-1}\hat{A}) - \ln(|A^{-1}\hat{A}|) - p] \\
&= E[\text{tr}(A^{-1}R\phi(M)R')] - E[\ln(|A^{-1}R\phi(M)R'|)] - p .
\end{aligned}$$

Then it follows from Lemma 1.4 that

$$\begin{aligned}
R(\hat{A}, A) &= 2E\left[\sum_{i=1}^p \sum_{t=i+1}^p \frac{d_i l_i - d_t l_t}{l_i - l_t}\right] + (n - p + 1) \sum_{i=1}^p d_i \\
&\quad - \sum_{i=1}^p \ln(d_i^*) - \sum_{i=1}^p E[\ln(\chi_{n+1-i}^2)] - 2pE[\ln(\tau)] - p
\end{aligned} \tag{4.1}$$

where l_i 's are the characteristic roots of the Wishart matrix W and d_i 's are given by (1.12). But

$$\sum_{i=1}^p \sum_{t=i+1}^p \frac{d_i l_i - d_t l_t}{l_i - l_t} = \sum_{i=1}^p \sum_{t=i+1}^p \frac{l_i(d_i - d_t)}{l_i - l_t} + \sum_{i=1}^p \sum_{t=i+1}^p d_t ,$$

and for $t = i + 1, i + 2, \dots, p$; $i = 1, 2, \dots, p$, we have $d_i \leq d_t$, $l_i \geq l_t$ so that $l_i/(l_i - l_t) > 1$ and $l_i(d_i - d_t)/(l_i - l_t) \leq d_i - d_t$; we then have

$$\begin{aligned}
\sum_{i=1}^p \sum_{t=i+1}^p \frac{d_i l_i - d_t l_t}{l_i - l_t} &\leq \sum_{i=1}^p \sum_{t=i+1}^p (d_i - d_t) + \sum_{i=1}^p \sum_{t=i+1}^p d_t \\
&= \sum_{i=1}^p \sum_{t=i+1}^p d_i = \sum_{i=1}^p (p - i)d_i .
\end{aligned} \tag{4.2}$$

Hence, by the use of (4.2) in (4.1) we have

$$\begin{aligned}
R(\hat{A}, A) &\leq 2 \sum_{i=1}^p (p - i)d_i + (n - p + 1) \sum_{i=1}^p d_i \\
&\quad - \sum_{i=1}^p \ln(d_i^*) - \sum_{i=1}^p E[\ln(\chi_{n+1-i}^2)] - 2pE[\ln(\tau)] - p \\
&= - \sum_{i=1}^p \ln(d_i^*) - \sum_{i=1}^p E[\ln(\chi_{n+1-i}^2)] - 2pE[\ln(\tau)] .
\end{aligned}$$

It then follows from (3.4) that $\hat{A}$ dominates A^* .

5 An Example

Suppose we have the following observed sum of product matrix

$$A = \begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix}$$

with $n = 10$, $v = 5$ and $p = 3$. We calculate the estimators $\tilde{A} = (v - 2)A/(vn)$, $A^* = TD^*T'$ and $\hat{A} = RD^*MR'$ where $D^* = \text{diag}(d_1^*, d_2^*, d_3^*)$, and $M = \text{diag}(m_1, m_2, m_3)$. It is easy to verify that $D^* = \text{diag}(.05, .06, .075)$,

$$T = \begin{bmatrix} 3.60555128 & 0.00000000 & 0.00000000 \\ -1.10950039 & 3.43063125 & 0.00000000 \\ 0.55470020 & -0.40360368 & 3.08697453 \end{bmatrix},$$

$m_1 = 18, m_2 = 9, m_3 = 9, M = \text{diag}(18, 9, 9)$ and

$$R = \begin{bmatrix} 0.66666667 & 0.70710678 & 0.23570226 \\ -0.66666667 & 0.70710678 & -0.23570226 \\ 0.33333333 & 0.00000000 & -0.94280904 \end{bmatrix}$$

so that $A = TT'$, $A = RMR'$ and $D^*M = \text{diag}(0.9, 0.54, 0.675)$.

Hence the unbiased estimator $\tilde{A} = (v - 2)A/(vn)$ is given by

$$\tilde{A} = \begin{bmatrix} 0.78000000 & -0.24000000 & 0.12000000 \\ -0.24000000 & 0.78000000 & -0.12000000 \\ 0.12000000 & -0.12000000 & 0.60000000 \end{bmatrix}$$

and the estimator $A^* = TD^*T'$ is given by

$$A^* = \begin{bmatrix} 0.65000000 & -0.20000000 & 0.10000000 \\ -0.20000000 & 0.76769231 & -0.11384615 \\ 0.10000000 & -0.11384615 & 0.73986425 \end{bmatrix}$$

while the estimator $\hat{A} = RD^*MR'$ given by

$$\hat{A} = \begin{bmatrix} 0.70750000 & -0.16750000 & 0.05000000 \\ -0.16750000 & 0.70750000 & -0.05000000 \\ 0.05000000 & -0.05000000 & 0.70000000 \end{bmatrix}.$$

We note that the smallest risk estimator in the class of estimators cA is an unbiased estimator given by $\tilde{A} = (v - 2)A/(vn)$. The maximum likelihood estimator of A given by $(n + 1)^{-1}A$ (Anderson et al. 1986) also belongs to the class cA ; it, therefore, follows that $\tilde{A}$ dominates the maximum likelihood estimator. The estimator $A^* = TD^*T'$ dominates $\tilde{A}$. However, the estimator $\hat{A} = RD^*MR'$ dominates the estimator A^* . Finally, we remark that as $v > \infty$, the estimators of the scale matrix of the multivariate t -model coincide, as expected, with those of the multivariate normal model.

Acknowledgements: The authors thank Prof. M. L. King, Monash University, Australia and Prof. A. K. M. E. Saleh, Carleton University, Canada for their constructive comments on an earlier draft of this paper. The authors are also grateful to the editor Prof. Dr. W. Uhlmann and an anonymous referee for their constructive suggestions that considerably improved the presentation of this paper.

References

- Anderson TW (1984) An introduction to multivariate statistical analysis. John Wiley, New York
 Anderson TW, Fang KT, Hsu H (1986) Maximum likelihood estimates and likelihood ratio criterion for multivariate elliptically contoured distributions. *Canad J Statist* 14: 55–59
 Blattberg RC, Gonedes NJ (1974) A comparison of the stable and Student distributions as statistical models for stock prices. *J Business* 47: 224–280
 Dey DK, Srinivasan C (1985) Estimation of a covariance matrix under Stein's loss. *Ann Statist* 13: 1581–1591
 James W, Stein C (1961) Estimation with quadratic loss. *Proc Fourth Berkeley Symp Math Statist Prob* 1: 361–379
 Lange KL, Little RJ, Taylor MG (1989) Robust statistical modelling using the t distribution. *J Amer Statist Assoc* 84: 881–896
 Muirhead RJ (1982) Aspects of multivariate statistical theory. Wiley, New York
 Stein C (1975) Reitz Lecture, 39th Annual Meeting IMS. Atlanta, Georgia
 Sutradhar BC, Ali MM (1989) A generalization of the Wishart distribution for the elliptical model and its moments for the multivariate t model. *J Mult Anal* 29: 155–162
 Zellner A (1976) Bayesian and non-Bayesian analysis of the regression model with multivariate Student- t error term. *J Amer Statist Assoc* 71: 400–405

Received 04.08.1995