Estimation of the Precision Matrix of Multivariate Pearson type II Model

Amadou Sarr[1] and **Arjun K. Gupta**[2]
Department of Mathematics and Statistics,
Bowling Green State University,
Bowling Green, OH 43403, USA.

Anwar H. Joarder[3]
Department of Mathematical Sciences
King Fhad University of Petroleum and Minerals
Dhahran 31261, Saudi Arabia.

Abstract
In this paper, the problem of estimating the precision matrix of a multivariate Pearson type II-model is considered. A new class of estimators is proposed. Moreover, the risk functions of the usual and the proposed estimators are explicitly derived. It is shown that the proposed estimator dominates the MLE and the unbiased estimator, under the quadratic loss function. A simulation study is carried out and confirms these results. Improved estimator of $tr(\mathbf{\Sigma}^{-1})$ is also obtained.

1 Introduction

There has been considerable research on the problem of estimating the precision matrix, namely the inverse of the scale matrix $\mathbf{\Sigma}$, in a multivariate normal model using the decision-theoretic approach. Works along this direction can be found in Efron and Morris (1976), Haff (1979), Krishnamoorthy and Gupta (1989), Pal (1993) among others. Several kinds of estimators of $\mathbf{\Sigma}^{-1}$ have been proposed in the literature, and the most of them can be found in the recent paper of Kubokawa (2005). In order to extend the above quoted results, much attention has been paid to the problem of estimating the precision matrix, under multivariate elliptical models. Indeed, Joarder and Ahmed (1998) obtained improved estimators of $\mathbf{\Sigma}^{-1}$ under the quadratic loss function, and for a specific

---

[1]E-mail: asarr@bgnet.bgsu.edu (A. Sarr)
[2]E-mail: gupta@bgnet.bgsu.edu (A.K. Gupta)
[3]E-mail: anwarj@kfupm.edu.sa (A.H. Joarder)

multivariate elliptical model. More recently, Tsukuma (2005) established similar results for an elliptically contoured distribution, under the Stein's loss function. In this paper, the problem of estimating the precision matrix of a multivariate Pearson type II-model with respect to the quadratic loss function is investigated. First, we give some definitions and results that are taken from Fang et al (1990).

Definition 1.1. *The p-dimensional random vector $\mathbf{x}$ is said to have a symmetric multivariate* **Pearson type II** *distribution with parameters $q > -1$, $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma} : p \times p$ with $\boldsymbol{\Sigma} > \mathbf{O}$ if its probability density function (p.d.f.) is given by*

$$f(\mathbf{x}) = \frac{\Gamma\left(\frac{p}{2} + q + 1\right)}{\pi^{\frac{p}{2}} \Gamma(q+1) \mid \boldsymbol{\Sigma} \mid^{\frac{1}{2}}} \left[1 - (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right]^q, \tag{1}$$

where $0 \leq (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \leq 1$.

This distribution was introduced by Kotz (1975) and will be denoted by $MPII_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Note that the first detailed discussion of this subclass of elliptically contoured distributions was presented in the book by Johnson (1987). As a member of the elliptical family, this distribution admits the stochastic representation which is illustrated by the following relation $\mathbf{x} \overset{d}{=} \boldsymbol{\mu} + r \boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{u}^{(p)}$, where $\mathbf{u}^{(p)}$ is uniformly distributed on the unit sphere surface in $\mathcal{R}^p$, and r is independent of $\mathbf{u}^{(p)}$. Further, it is shown (see Fang et al (1990), p.89) that r^2 has a Beta type I distribution, say $r^2 \sim B_I(\frac{p}{2}, q+1)$. Note that a closed form of the characteristic function (c.f.) of the multivariate Pearson type II distribution has been obtained by Joarder (1997). Now, we define the multivariate Pearson type II-model. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ be p-dimensional random vectors, such that $n > p$ and $\mathbf{x}_i \sim MPII_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $i = 1, \ldots, n$. Moreover, assume that $\mathbf{x}_i$, $i = 1, \ldots, n$ are uncorrelated,(but not necessarily independent), and their joint *p.d.f.* is given by

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{\Gamma(\frac{pn}{2} + q + 1)}{\pi^{\frac{pn}{2}} \Gamma(q+1) \mid \boldsymbol{\Sigma} \mid^{\frac{n}{2}}} \left(1 - \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu})\right)^q. \tag{2}$$

Here the parameter q is assumed to be known, and each p-dimentional random vector $\mathbf{x}_i$, $i = 1, \ldots, n$ is distributed with mean vector $\boldsymbol{\mu}$ and covariance matrix $\frac{1}{2q+p+2} \boldsymbol{\Sigma}$. So the relation (2) represents the multivariate Pearson type II-model. Our interest is to estimate $\boldsymbol{\Sigma}^{-1}$ and its trace, under a decision-theoretic viewpoint.

Remark 1.1.
Let $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ be the $p \times n$ matrix of the observations vectors, then the expression given by (2) is the p.d.f. of $\mathbf{X}$. That leads us to some preliminaries related to the matrix variate elliptically contoured distributions (ECD).

2 Some Preliminaries

The following definitions and results presented in this section, and that will be required in the sequel, are taken from Gupta and Varga (1993).

Definition 2.1. *Let* $\mathbf{X}$ *be a* $p \times n$ *random matrix. Then,* $\mathbf{X}$ *is said to have a matrix variate elliptically contoured distribution if its characteristic function has the form*

$$\phi_{\mathbf{X}}(\mathbf{L}) \;=\; etr(i\mathbf{L}'\mathbf{M})\psi(tr(\mathbf{L}'\boldsymbol{\Sigma}\mathbf{L}\boldsymbol{\Phi})) \; with \; \mathbf{L} \in \mathcal{R}^{p \times n}, \; \mathbf{M} \in \mathcal{R}^{p \times n}, \; \boldsymbol{\Sigma} \in \mathcal{R}^{p \times p}, \; \boldsymbol{\Phi} \in \mathcal{R}^{n \times n},$$

$$\boldsymbol{\Sigma} \;\geq\; \mathbf{O}, \quad \boldsymbol{\Phi} \;\geq\; \mathbf{O} \; and \; \psi : \; [\,0,\; \infty[\; \to \mathbb{R}. \tag{3}$$

The matrices $\mathbf{M}$, $\boldsymbol{\Sigma}$ *and* $\boldsymbol{\Phi}$ *are the parameters of the distribution.*

This distribution is denoted by $\mathbf{X} \;\sim\; E_{p,n}(\mathbf{M}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Phi}, \psi)$.
The function ψ is called the characteristic generator (c.g.). As a special case, when $\psi(\cdot)$ is specified by $\psi(z) \;=\; \exp(-\frac{z}{2})$, then $\mathbf{X}$ has a matrix variate normal distribution. If $n = 1$, then $\mathbf{x} \sim E_p(\mathbf{m}, \boldsymbol{\Sigma}, \psi)$ is said to have a vector variate elliptical distribution. The relationship, in term of the distributions, of the matrix and the vector is illustrated as follows:

$$\mathbf{X} \;\sim\; E_{p,n}(\mathbf{M}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Phi}, \psi) \; if \; and \; only \; if \; \mathbf{x} = vec(\mathbf{X}') \;\sim\; E_{pn}(vec(\mathbf{M}'), \boldsymbol{\Sigma} \otimes \boldsymbol{\Phi}, \psi), \tag{4}$$

Here, $vec(\mathbf{A})$ is defined by:

$$vec(\mathbf{A}) = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix},$$

where $\mathbf{a}_1, \ldots, \mathbf{a}_n$ denote the columns of the $p \times n$ matrix $\mathbf{A}$. Anderson and Fang (1982), (see [4] pp.1-23), derived the stochastic representation of matrix variate ECD in the theorem below.

Theorem 2.1. *Let* $\mathbf{X}$ *be a* $p \times n$ *random matrix. Further let* $\mathbf{M}$ *be* $p \times n$, $\boldsymbol{\Sigma}$ *be* $p \times p$ *and* $\boldsymbol{\Phi}$ *be* $n \times n$ *constant matrices, with* $\boldsymbol{\Sigma} \geq \mathbf{O}$, $\boldsymbol{\Phi} \geq \mathbf{O}$, $rank(\boldsymbol{\Sigma}) = p_1$, *and* $rank(\boldsymbol{\Phi}) = n_1$. *Then*

$$\mathbf{X} \;\sim\; E_{p,n}(\mathbf{M}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Phi}, \psi)$$

if and only if

$$\mathbf{X} \;\overset{d}{=}\; \mathbf{M} + R\mathbf{DUB}', \tag{5}$$

where $\mathbf{U}$ *is* $p_1 \times n_1$ *and* $vec(\mathbf{U}')$ *is uniformly distributed on the unit sphere* $S_{p_1 n_1}$, R *is a nonnegative random variable,* R *and* $\mathbf{U}$ *are independent,* $\boldsymbol{\Sigma} = \mathbf{DD}'$ *and* $\boldsymbol{\Phi} = \mathbf{BB}'$ *are rank factorizations of* $\boldsymbol{\Sigma}$ *and* $\boldsymbol{\Phi}$.

Note that an immediate consequence of relation (4) is the fact that the moments of the random variable R (generating variate matrix) can be obtained from that of r (generating variate vector) by writing pn instead of n.
Assume that the matrix $\mathbf{X}$ has a p.d.f., with $\boldsymbol{\Sigma} > \mathbf{O}$, $\boldsymbol{\Phi} > \mathbf{O}$. In that case, its p.d.f. takes the form (see Gupta and Varga(1993), p.26)

$$f(\mathbf{X}) \;=\; |\,\boldsymbol{\Sigma}\,|^{-\frac{n}{2}}|\,\boldsymbol{\Phi}\,|^{-\frac{p}{2}}h(tr((\mathbf{X} - \mathbf{M})'\boldsymbol{\Sigma}^{-1}(\mathbf{X} - \mathbf{M})\boldsymbol{\Phi}^{-1})). \tag{6}$$

Here, $h(\cdot)$ is called the density generator of the distribution. In the particular case where $\mathbf{\Phi} = \mathbf{I}_n$ and $\mathbf{M} = \boldsymbol{\mu}\mathbf{e}_n'$, with $\mathbf{e}_n' = (1, 1, \ldots, 1)$, the p.d.f.(6) simplifies to

$$f(\mathbf{X}) \;=\; \frac{1}{|\mathbf{\Sigma}|^{\frac{n}{2}}} h\left(\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu})\right). \tag{7}$$

Now, with the above notations and results, we remark that the multivariate Pearson type II-model (2) can be expressed as

$$\mathbf{X} \;\sim\; E_{p,n}(\boldsymbol{\mu}\mathbf{e}_n', \mathbf{\Sigma}\otimes\mathbf{I}_n, \psi).$$

Indeed, for such a model, Anderson et al. (1986) showed that the maximum likelihood estimator (MLE) of the scale matrix $\mathbf{\Sigma}$ has the form $c_0\mathbf{A}$, where $\mathbf{A} = \sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})'$ is the sample sum of product matrix, $\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i$ is the sample mean, and c_0 is a positive constant that depends on the specific multivariate elliptical-model. Consequently a natural (or usual) estimator of the precision matrix $\mathbf{\Sigma}^{-1}$ has the form $\alpha_0\mathbf{A}^{-1}$. For example, for the multivariate Pearson type II-model, we can easily show that $\alpha_0 = \dfrac{n}{2q + pn}$, under the restriction $q > 0$. However, this estimator of $\mathbf{\Sigma}^{-1}$ is generally not optimal from a decision-theoretic viewpoint. The aim of this paper is to propose new estimators of $\mathbf{\Sigma}^{-1}$ and $tr(\mathbf{\Sigma}^{-1})$ which dominate the usual ones, under the quadratic loss and the squared error loss functions. Note that our motivation arises from the paper of Joarder and Ahmed (1998), who established dominance results in the subclass of scale mixture of normal distributions. Here, it is important to mention that Pearson type II distribution cannot be expressed as mixture of normal distribution (see Kano (1994)). Following Joarder and Ahmed, we will make use of the loss functions given as

$$L(\widehat{\mathbf{T}}, \mathbf{T}) \;=\; tr\left[(\widehat{\mathbf{T}} - \mathbf{T})^2\right], \tag{8}$$

$$L(\hat{\delta}, \delta) \;=\; (\hat{\delta} - \delta)^2, \tag{9}$$

for the estimation of $\mathbf{\Sigma}^{-1}$ and $tr(\mathbf{\Sigma}^{-1})$ respectively. Their corresponding risk functions are obtained by taking the expectations of the loss functions. The usual estimator of $\mathbf{T} = \mathbf{\Sigma}^{-1}$ will be denoted by $\widehat{\mathbf{T}}_u = \alpha_0\mathbf{A}^{-1}$. Now we propose a class of **improved estimator** of $\mathbf{\Sigma}^{-1}$ denoted by $\widehat{\mathbf{T}}$, and defined as

$$\widehat{\mathbf{T}} \;=\; \alpha_0\mathbf{A}^{-1} - \alpha\,|\,\mathbf{A}\,|^{-\frac{1}{p}}\,\mathbf{I}_p, \tag{10}$$

where α is chosen such that $\widehat{\mathbf{T}}$ is positive definite and $\mathbf{I}_p$ is the $p \times p$ dimensional identity matrix. The form of the improved estimator $\widehat{\mathbf{T}}$ is motivated by Dey's result (1988), who developed simultaneous estimators of the eigenvalues of the covariance matrix of the multivariate normal distribution by **shrinking** sample eigenvalues towards their geometric mean. Before giving the main results, we need some auxiliary lemmas. These will be proved later with the help of the following Gupta and Varga's (1993) result.

Theorem 2.2. *Let* $\mathbf{X} \sim E_{p,n}(\mathbf{O}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Phi}, \psi)$. *Let* $l = rank(\boldsymbol{\Sigma})$, $m = rank(\boldsymbol{\Phi})$, *and* $R\mathbf{DUB}'$ *be the stochastic representation of* $\mathbf{X}$. *Assume that* $\mathbf{Y} \sim N_{p,n}(\mathbf{O}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Phi})$. *Let* $K(\mathbf{Z})$ *be a function defined on* $\mathcal{R}^{p \times n}$ *such that if* $\mathbf{Z} \in \mathcal{R}^{p \times n}$ *and* $a \geq 0$ *then* $K(a\mathbf{Z}) = a^k K(\mathbf{Z})$ *where* $k > -lm$. *Assume* $E(K(\mathbf{X}))$ *and* $E(K(\mathbf{Y}))$ *exist. Then,*

$$E(K(\mathbf{X})) = E(K(\mathbf{Y})) \frac{E(R^k)\Gamma\left(\frac{lm}{2}\right)}{2^{\frac{k}{2}}\Gamma\left(\frac{lm+k}{2}\right)}. \tag{11}$$

The full version and the proof of this theorem can be found in Gupta and Varga (1993), p.100.

Remark 2.1. *Let* $\mathbf{X} \sim E_{p,n}(\boldsymbol{\mu}\mathbf{e}'_n, \boldsymbol{\Sigma} \otimes \mathbf{I}_n, \psi)$, $\boldsymbol{\Sigma} > \mathbf{O}$, *then the sample sum of product matrix* $\mathbf{A}$ *can be expressed as a function of* $\mathbf{X}$, *say*

$$\mathbf{A} = \mathbf{X}\left(\mathbf{I}_n - \frac{1}{n}\mathbf{e}_n\mathbf{e}'_n\right)\mathbf{X}'.$$

Now, by setting

$$K(\mathbf{X}) = \left[\mathbf{X}\left(\mathbf{I}_n - \frac{1}{n}\mathbf{e}_n\mathbf{e}'_n\right)\mathbf{X}'\right]^{-1},$$

and using the relation (11), we get the unbiased estimator of $\boldsymbol{\Sigma}^{-1}$ *given below, provided that* $0 < E(R^{-2}) < \infty$ *and* $n > p + 2$,

$$\widehat{\boldsymbol{\Sigma}}^{-1} = \frac{n - p - 2}{(pn - 2)E(R^{-2})}\mathbf{A}^{-1}. \tag{12}$$

The following lemmas provide some moments of the sample sum of product matrix of the multivariate Pearson type II-model. In sequel, we assume $n > p + 2$ except for some cases where we need $n > p + 4$.

Lemma 2.1. *The following relations hold.*

$$E(\mathbf{A}^{-1}) = \frac{np + 2q}{n - p - 2}\boldsymbol{\Sigma}^{-1}. \tag{13}$$

$$E(\mathbf{A}^{-2}) = (np + 2q)(np + 2q - 2)E\left(\mathbf{W}^{-2}\right), \tag{14}$$

where $n > p + 4$, $\mathbf{W} \stackrel{d}{=} \mathbf{W}_p(\boldsymbol{\Sigma}, n - 1)$ *and* $E\left(\mathbf{W}^{-2}\right)$ *is given by (see Gupta and Nagar (2000), p.99)*

$$
\begin{aligned}
E\left(\mathbf{W}^{-2}\right) &= [(n - p - 1)(n - p - 4)]^{-1}\boldsymbol{\Sigma}^{-2} \\
&+ [(n - p - 1)(n - p - 2)(n - p - 4)]^{-1}\boldsymbol{\Sigma}^{-1}tr(\boldsymbol{\Sigma}^{-1}).
\end{aligned} \tag{15}
$$

Proof: Since R^2 is distributed according to a Beta type I distribution with parameters $\frac{np}{2}$ and $q+1$, we have

$$E\left(R^{-2t}\right) = \frac{\Gamma\left(\frac{np}{2} - t\right)\Gamma\left(\frac{np}{2} + q + 1\right)}{\Gamma\left(\frac{np}{2}\right)\Gamma\left(\frac{np}{2} + q - t + 1\right)}. \tag{16}$$

Now, using (16) with $t = 1$ and (12), we obtain (13).

To prove (14), we consider the function $K(\cdot)$ defined on $\mathcal{R}^{p \times n}$ by

$$K(\mathbf{X}) = \left[\mathbf{X}\left(\mathbf{I}_n - \frac{1}{n}\mathbf{e}_n\mathbf{e}_n'\right)\mathbf{X}'\right]^{-2}.$$

Hence, from relation (11), we have

$$E(\mathbf{A}^{-2}) = \frac{4E(R^{-4})\Gamma\left(\frac{np}{2}\right)}{\Gamma\left(\frac{np-4}{2}\right)}E(\mathbf{W}^{-2}).$$

On the other hand, $E(R^{-4})$ is obtained from (16), with $t = 2$

$$E(R^{-4}) = \frac{(np + 2q)(np + 2q - 2)}{(np - 2)(np - 4)},$$

and the desired result follows directly.$\square$

Lemma 2.2. *For any t such that $n - p - 2t - 2 > 0$ we have*

$$E\left(\mid \mathbf{A} \mid^{-t}\right) = \frac{\Gamma\left(\frac{np}{2} + q + 1\right)\Gamma_p\left(\frac{n-1}{2} - t\right)}{\Gamma\left(\frac{np}{2} + q - tp + 1\right)\Gamma_p\left(\frac{n-1}{2}\right)}\mid \mathbf{\Sigma} \mid^{-t} \tag{17}$$

$$E\left(\mid \mathbf{A} \mid^{-t} \mathbf{A}^{-1}\right) = \frac{2\Gamma\left(\frac{np}{2} + q + 1\right)\Gamma_p\left(\frac{n-1}{2} - t\right)}{(n - p - 2t - 2)\Gamma\left(\frac{np}{2} + q - tp\right)\Gamma_p\left(\frac{n-1}{2}\right)}\mid \mathbf{\Sigma} \mid^{-t}\mathbf{\Sigma}^{-1}. \tag{18}$$

Proof: From relation (11) with the function $K(\cdot)$ defined on $\mathcal{R}^{p \times n}$ by

$$K(\mathbf{X}) = \mid \mathbf{X}(\mathbf{I}_n - \frac{1}{n}\mathbf{e}_n\mathbf{e}_n')\mathbf{X}' \mid^{-t},$$

we have

$$E\left(\mid \mathbf{A} \mid^{-t}\right) = \frac{E(R^{-2tp})\Gamma\left(\frac{np}{2}\right)}{2^{-tp}\Gamma\left(\frac{np-2tp}{2}\right)}E(\mid \mathbf{W} \mid^{-t}), \tag{19}$$

where $E(\mid \mathbf{W} \mid^{-t})$ is given by the following well-known relation (see Gupta and Nagar (2000), p.105)

$$E\left(\mid \mathbf{W} \mid^{-t}\right) = \frac{2^{-tp}\Gamma_p\left(\frac{n-1}{2} - t\right)}{\Gamma_p\left(\frac{n-1}{2}\right)}\mid \mathbf{\Sigma} \mid^{-t}, \tag{20}$$

6

and $E(R^{-2tp})$ is given by (see 16)

$$E(R^{-2tp}) = \frac{\Gamma\left(\frac{np}{2} - tp\right)\Gamma\left(\frac{np}{2} + q + 1\right)}{\Gamma\left(\frac{np}{2}\right)\Gamma\left(\frac{np}{2} + q - tp + 1\right)}. \tag{21}$$

Consequently, we obtain (17) by substituting (21) and (20) in equation (19).
In the same way, by using (11) and the function $K(\cdot)$ defined on $\mathcal{R}^{p \times n}$ by

$$\begin{aligned} K(\mathbf{X}) &= |\mathbf{X}(\mathbf{I}_n - \frac{1}{n}\mathbf{e}_n\mathbf{e}_n')\mathbf{X}'|^{-t}\left[\mathbf{X}(\mathbf{I}_n - \frac{1}{n}\mathbf{e}_n\mathbf{e}_n')\mathbf{X}'\right]^{-1} \\ &= |\mathbf{A}|^{-t}\mathbf{A}^{-1}, \end{aligned}$$

we have

$$E\left(|\mathbf{A}|^{-t}\mathbf{A}^{-1}\right) = \frac{E(R^{-2(tp+1)})\Gamma\left(\frac{np}{2}\right)}{2^{-(tp+1)}\Gamma\left(\frac{np}{2} - tp - 1\right)}E(|\mathbf{W}|^{-t}\mathbf{W}^{-1}). \tag{22}$$

Here, $E(R^{-2(tp+1)})$ is given by

$$E(R^{-2(tp+1)}) = \frac{\Gamma\left(\frac{np}{2} - tp - 1\right)\Gamma\left(\frac{np}{2} + q + 1\right)}{\Gamma\left(\frac{np}{2}\right)\Gamma\left(\frac{np}{2} + q - tp\right)}. \tag{23}$$

Since $E\left(|\mathbf{W}|^{-t}\mathbf{W}^{-1}\right)$ is given (see Dey (1988)) by

$$E\left(|\mathbf{W}|^{-t}\mathbf{W}^{-1}\right) = \frac{2^{-tp}\Gamma_p\left(\frac{n-1}{2} - t\right)}{(n - p - 2t - 2)\Gamma_p\left(\frac{n-1}{2}\right)}|\mathbf{\Sigma}|^{-t}\mathbf{\Sigma}^{-1}, \tag{24}$$

then relation (18) follows from (24) and (23).□
As special cases, when $t = \frac{1}{p}$ and $t = \frac{2}{p}$ in the relation (17), we obtain respectively

$$E\left(|\mathbf{A}|^{-\frac{1}{p}}\right) = \frac{np + 2q}{2}\frac{\Gamma_p\left(\frac{n-1}{2} - \frac{1}{p}\right)}{\Gamma_p\left(\frac{n-1}{2}\right)}|\mathbf{\Sigma}|^{-\frac{1}{p}}, \tag{25}$$

$$E\left(|\mathbf{A}|^{-\frac{2}{p}}\right) = \frac{(np + 2q)(np + 2q - 2)}{4}\frac{\Gamma_p\left(\frac{n-1}{2} - \frac{2}{p}\right)}{\Gamma_p\left(\frac{n-1}{2}\right)}|\mathbf{\Sigma}|^{-\frac{2}{p}}. \tag{26}$$

When $t = \frac{1}{p}$ in the relation (18), we obtain

$$E\left(|\mathbf{A}|^{-\frac{1}{p}}\mathbf{A}^{-1}\right) = \frac{(np + 2q)(np + 2q - 2)\Gamma_p\left(\frac{n-1}{2} - \frac{1}{p}\right)}{2(n - p - \frac{2}{p} - 2)\Gamma_p\left(\frac{n-1}{2}\right)}|\mathbf{\Sigma}|^{-\frac{1}{p}}\mathbf{\Sigma}^{-1}. \tag{27}$$

These last three expressions will allow us to prove the main results of this paper.

3 Main Results

The main results are presented in this section in the form of three theorems.

Theorem 3.1. *Consider the multivariate Pearson type II-model (2). Then, under the loss function given by (8), the improved estimator $\widehat{\mathbf{T}} = \alpha_0 \mathbf{A}^{-1} - \alpha \mid \mathbf{A} \mid^{-\frac{1}{p}} \mathbf{I}_p$ dominates the usual estimator $\widehat{\mathbf{T}}_u = \alpha_0 \mathbf{A}^{-1}$ for any α satisfying the conditions:*

$$d_p \; < \; \alpha \; < \; 0, \quad with \; \alpha_0 \; < \; \frac{n - p - \frac{2}{p} - 2}{np + 2q - 2}, \tag{28}$$

$$or \; 0 \; < \; \alpha \; < \; d_p, \quad with \; \alpha_0 \; > \; \frac{n - p - \frac{2}{p} - 2}{np + 2q - 2}, \tag{29}$$

where d_p is given by

$$d_p \; = \; \left[\frac{\alpha_0}{n - p - \frac{2}{p} - 2} - \frac{1}{np + 2q - 2} \right] \frac{4\Gamma_p\left(\frac{n-1}{2} - \frac{1}{p}\right)}{\Gamma_p\left(\frac{n-1}{2} - \frac{2}{p}\right)}. \tag{30}$$

Proof: Consider the risk difference $D(\mathbf{T}, \alpha) = R(\widehat{\mathbf{T}}, \mathbf{T}; \alpha) - R(\widehat{\mathbf{T}}_u, \mathbf{T})$. We thus have

$$D(\mathbf{T}, \alpha) \; = \; -2\alpha_0 \alpha tr \left[E\left(\mid \mathbf{A} \mid^{-\frac{1}{p}} \mathbf{A}^{-1} \right) \right] + \alpha^2 p E\left[\mid \mathbf{A} \mid^{-\frac{2}{p}} \right] + 2\alpha E\left[\mid \mathbf{A} \mid^{-\frac{1}{p}} \right] tr(\mathbf{\Sigma}^{-1}),$$

and from relations (25), (26) and (27), it can be written as

$$
\begin{aligned}
D(\mathbf{T}, \alpha) \; = \; & \frac{-\alpha_0 \alpha p(np + 2q)(np + 2q - 2)\Gamma_p\left(\frac{n-1}{2} - \frac{1}{p}\right)}{(n - p - \frac{2}{p} - 2)\Gamma_p\left(\frac{n-1}{2}\right)} \mid \mathbf{\Sigma} \mid^{-\frac{2}{p}} \frac{tr(\mathbf{\Sigma}^{-1}/p)}{\mid \mathbf{\Sigma} \mid^{-\frac{1}{p}}} \\
& + \frac{\alpha^2 p(np + 2q)(np + 2q - 2)\Gamma_p\left(\frac{n-1}{2} - \frac{2}{p}\right)}{4\Gamma_p\left(\frac{n-1}{2}\right)} \mid \mathbf{\Sigma} \mid^{-\frac{2}{p}} \\
& + \frac{\alpha p(np + 2q)\Gamma_p\left(\frac{n-1}{2} - \frac{1}{p}\right)}{\Gamma_p\left(\frac{n-1}{2}\right)} \mid \mathbf{\Sigma} \mid^{-\frac{2}{p}} \frac{tr(\mathbf{\Sigma}^{-1}/p)}{\mid \mathbf{\Sigma} \mid^{-\frac{1}{p}}} \\
= \; & \frac{p(np + 2q) \mid \mathbf{\Sigma} \mid^{-\frac{2}{p}}}{\Gamma_p\left(\frac{n-1}{2}\right)} \alpha \\
& \times \left[\left(1 - \frac{\alpha_0(np + 2q - 2)}{(n - p - \frac{2}{p} - 2)} \right) \Gamma_p\left(\frac{n-1}{2} - \frac{1}{p}\right) \frac{tr(\mathbf{\Sigma}^{-1}/p)}{\mid \mathbf{\Sigma} \mid^{-\frac{1}{p}}} \right. \\
& \left. + \frac{(np + 2q - 2)\Gamma_p\left(\frac{n-1}{2} - \frac{2}{p}\right)}{4} \alpha \right].
\end{aligned}
$$

8

Consequently, we have

$$
\begin{aligned}
D(\mathbf{T}, \alpha) \;=\; & \frac{p(np + 2q)(np + 2q - 2)\Gamma_p\left(\frac{n-1}{2} - \frac{2}{p}\right) \mid \mathbf{\Sigma} \mid^{-\frac{2}{p}}}{4\Gamma_p\left(\frac{n-1}{2}\right)}\alpha \\
& \times \left[\alpha + \left(\frac{1}{np + 2q - 2} - \frac{\alpha_0}{n - p - \frac{2}{p} - 2}\right)\frac{4\Gamma_p\left(\frac{n-1}{2} - \frac{1}{p}\right)}{\Gamma_p\left(\frac{n-1}{2} - \frac{2}{p}\right)}\frac{tr(\mathbf{\Sigma}^{-1}/p)}{\mid \mathbf{\Sigma} \mid^{-\frac{1}{p}}}\right].
\end{aligned}
$$

Now, note that $D(\mathbf{T}, \alpha)$ can be written as

$$
D(\mathbf{T}, \alpha) \;=\; C\alpha\left(\alpha - d_p\frac{tr(\mathbf{\Sigma}^{-1}/p)}{\mid \mathbf{\Sigma} \mid^{-\frac{1}{p}}}\right), \tag{31}
$$

where d_p is given by (30) and C is the positive constant defined by

$$
C \;=\; \frac{p(np + 2q)(np + 2q - 2)\Gamma_p\left(\frac{n-1}{2} - \frac{2}{p}\right) \mid \mathbf{\Sigma} \mid^{-\frac{2}{p}}}{4\Gamma_p\left(\frac{n-1}{2}\right)}.
$$

Now, $\widehat{\mathbf{T}}$ dominates $\widehat{\mathbf{T}}_u$ if and only if $D(\mathbf{T}, \alpha) < 0$. First, we consider the case when $d_p > 0$. From (30) we deduce

$$
d_p > 0 \;\; \text{if and only if} \;\; \alpha_0 > \frac{n - p - \frac{2}{p} - 2}{np + 2q - 2}.
$$

In this case, $D(\mathbf{T}, \alpha) < 0$ if and only if $0 < \alpha < d_p\frac{tr(\mathbf{\Sigma}^{-1}/p)}{\mid\mathbf{\Sigma}\mid^{-\frac{1}{p}}}$. Now, using the following relation

$$
\frac{tr(\mathbf{\Sigma}^{-1}/p)}{\mid \mathbf{\Sigma} \mid^{-\frac{1}{p}}} \;\geq\; 1, \tag{32}
$$

(inequality between the arithmetic and geometric means) we remark that if $0 < \alpha < d_p$ then $0 < \alpha < d_p\frac{tr(\mathbf{\Sigma}^{-1}/p)}{\mid\mathbf{\Sigma}\mid^{-\frac{1}{p}}}$ (i.e. $D(\mathbf{T}, \alpha) < 0$). So the first case is thus proved.

The proof of the case when $d_p < 0$ is easily obtained. Indeed,

$$
d_p < 0 \;\; \text{if and only if} \;\; \alpha_0 < \frac{n - p - \frac{2}{p} - 2}{np + 2q - 2}.
$$

In the same way, we remark that if $d_p < \alpha < 0$ then $d_p\frac{tr(\mathbf{\Sigma}^{-1}/p)}{\mid\mathbf{\Sigma}\mid^{-\frac{1}{p}}} < \alpha < 0$,

(i.e. $D(\mathbf{T}, \alpha) < 0$).$\square$

The explicit expressions of the risk functions of the estimators $\widehat{\mathbf{T}}$ and $\widehat{\mathbf{T}}_u$ are derived in the following theorem.

Theorem 3.2. *The risk functions of the usual estimator $\widehat{\mathbf{T}}_u = \alpha_0 \mathbf{A}^{-1}$ and the improved estimator $\widehat{\mathbf{T}} = \alpha_0 \mathbf{A}^{-1} - \alpha \mid \mathbf{A} \mid^{-\frac{1}{p}} \mathbf{I}_p$ are respectively given by*

$$R(\widehat{\mathbf{T}}_u, \mathbf{T}) = \gamma_1 tr(\mathbf{\Sigma}^{-2}) + \gamma_2 [tr(\mathbf{\Sigma}^{-1})]^2 , \tag{33}$$

where γ_1 and γ_2 are defined as follows

$$\gamma_1 = \frac{\alpha_0^2(np+2q)(np+2q-2)}{(n-p-1)(n-p-4)} - \frac{2\alpha_0(np+2q)}{n-p-2} + 1 , \tag{34}$$

$$\gamma_2 = \frac{\alpha_0^2(np+2q)(np+2q-2)}{(n-p-1)(n-p-2)(n-p-4)} , \tag{35}$$

and

$$R(\widehat{\mathbf{T}}, \mathbf{T}; \alpha) = \frac{p(np+2q)(np+2q-2)\Gamma_p\left(\frac{n-1}{2} - \frac{2}{p}\right) \mid \mathbf{\Sigma} \mid^{-\frac{2}{p}}}{4\Gamma_p\left(\frac{n-1}{2}\right)} \alpha \left(\alpha - d_p \frac{tr(\mathbf{\Sigma}^{-1}/p)}{\mid \mathbf{\Sigma} \mid^{-\frac{1}{p}}} \right)$$
$$+ \gamma_1 tr(\mathbf{\Sigma}^{-2}) + \gamma_2 [tr(\mathbf{\Sigma}^{-1})]^2 , \tag{36}$$

where d_p is given by (30).

Proof: Indeed, by direct calculations, we get

$$R(\widehat{\mathbf{T}}_u, \mathbf{T}) = \alpha_0^2 tr\left[E(\mathbf{A}^{-2})\right] - 2\alpha_0 tr[E(\mathbf{A}^{-1})\mathbf{\Sigma}^{-1}] + tr(\mathbf{\Sigma}^{-2}).$$

Since $E(\mathbf{A}^{-1})$ and $E(\mathbf{A}^{-2})$ are given by, (see (13) and (14)), $E(\mathbf{A}^{-1}) = m_1\mathbf{\Sigma}^{-1}$ and $E(\mathbf{A}^{-2}) = m_2 E(\mathbf{W}^{-2})$, where $m_1 = \dfrac{np+2q}{n-p-2}$, $m_2 = (np+2q)(np+2q-2)$ and $E(\mathbf{W}^{-2})$ is given by (15), we find

$$R(\widehat{\mathbf{T}}_u, \mathbf{T}) = \frac{\alpha_0^2 m_2 tr(\mathbf{\Sigma}^{-2})}{(n-p-1)(n-p-4)} + \frac{\alpha_0^2 m_2}{(n-p-1)(n-p-2)(n-p-4)}[tr(\mathbf{\Sigma}^{-1})]^2$$
$$- 2\alpha_0 m_1 tr(\mathbf{\Sigma}^{-2}) + tr(\mathbf{\Sigma}^{-2})$$
$$= [\frac{\alpha_0^2 m_2}{(n-p-1)(n-p-4)} - 2\alpha_0 m_1 + 1]tr(\mathbf{\Sigma}^{-2})$$
$$+ \frac{\alpha_0^2 m_2}{(n-p-1)(n-p-2)(n-p-4)}[tr(\mathbf{\Sigma}^{-1})]^2 .$$

Now, we easily verify that

$$\begin{cases} \frac{\alpha_0^2 m_2}{(n-p-1)(n-p-4)} - 2\alpha_0 m_1 + 1 &= \gamma_1 \\ \frac{\alpha_0^2 m_2}{(n-p-1)(n-p-2)(n-p-4)} &= \gamma_2. \end{cases}$$

On the other hand, the proof of (36) follows directly from (33) and (31). The theorem is thus proved. $\square$

3.0.1 Estimation of $tr(\boldsymbol{\Sigma}^{-1})$

Let $\hat{\delta}_u = \alpha_0 tr(\mathbf{A}^{-1})$ and $\hat{\delta}_i = \alpha_0 tr(\mathbf{A}^{-1}) - \alpha p \mid \mathbf{A} \mid^{-\frac{1}{p}}$ be the usual and the proposed estimator of $\delta = tr(\boldsymbol{\Sigma}^{-1})$ respectively. Then the dominance result between $\hat{\delta}_i$ and $\hat{\delta}_u$ is given in the following theorem.

Theorem 3.3. *Consider the multivariate Pearson type II-model (2). Then $\hat{\delta}_i$ dominates $\hat{\delta}_u$ under the squared error loss function given by (9), if one of the following conditions holds:*

$$d_p \; < \; \alpha \; < \; 0, \quad with \; \alpha_0 \; < \; \frac{n - p - \frac{2}{p} - 2}{np + 2q - 2}, \tag{37}$$

$$or \; 0 \; < \; \alpha \; < \; d_p, \quad with \; \alpha_0 \; > \; \frac{n - p - \frac{2}{p} - 2}{np + 2q - 2}, \tag{38}$$

where d_p is given by (30).

Proof: Let $D(\delta, \alpha)$ be the risk difference between the proposed and the usual estimators of δ. Then simple calculations show that $D(\delta, \alpha) = pD(\mathbf{T}, \alpha)$, where $D(\mathbf{T}, \alpha)$ is the difference between the risks of $\widehat{\mathbf{T}}$ and $\widehat{\mathbf{T}}_u$. Consequently, the rest of the proof follows directly from that of theorem 3.1. $\square$

Remark 3.1. *The risk difference $D(\mathbf{T}, \alpha) = C\alpha \left(\alpha - d_p \frac{tr(\boldsymbol{\Sigma}^{-1}/p)}{|\boldsymbol{\Sigma}|^{-\frac{1}{p}}} \right)$ is a convex parabolic function of α, and the optimal value of α (denoted by α_{op}) that minimize the risk difference is given by*

$$\alpha_{op} \; = \; \frac{d_p}{2} \frac{tr(\boldsymbol{\Sigma}^{-1}/p)}{\mid \boldsymbol{\Sigma} \mid^{-\frac{1}{p}}}, \tag{39}$$

which depends on the unkown scale matrix $\boldsymbol{\Sigma}$ and therefore this is not usable in practice. However, from (32), we have the following inequality

$$D(\mathbf{T}, \alpha) \; \leq \; C\alpha(\alpha - d_p) \quad where \; 0 < \alpha < d_p \quad or \; d_p < \alpha < 0,$$

and $\alpha'_{op} = \dfrac{d_p}{2}$, where d_p is given by (30), minimize the expression $C\alpha(\alpha - d_p)$. So this value for α will be used in a simulation study.

4 Simulation Study and Conclusion

In this section, a Monte Carlo simulation study is carried out to evaluate the risk performance of the proposed estimator $\widehat{\mathbf{T}}$ (corresponding to the choice $\alpha = \frac{d_p}{2}$), over the usual ones (i.e. the unbiased and MLE).
First, n p-dimensional random vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are generated from the multivariate

Pearson type II distribution with three different $\mathbf{\Sigma}$ and various q, by using the following stochastic representation

$$\mathbf{x}_i \stackrel{d}{=} r\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{u}^{(p)}, \quad i = 1, \ldots, n, \tag{40}$$

where $r^2 \stackrel{d}{=} Beta\left(\frac{p}{2}, q+1\right)$, and $\mathbf{u}^{(p)}$ is uniformly distributed on the unit sphere S_p. A method to generate $\mathbf{u}^{(p)}$ can be found in Fang and Wang (1994).

Let $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ be the $p \times n$ sample matrix. Define the sample covariance matrix $\mathbf{A}$ as $\mathbf{A} = \mathbf{X}\mathbf{X}'$. Here the condition $n > p + 4$ is required. Then the estimators $\widehat{\mathbf{T}}$ and $\widehat{\mathbf{T}}_u$, and their corresponding losses are computed.

In our simulation, $p = 3$ and we choose $n = 8, 20, 30$. This procedure is repeated 2000 times and the average loss is used to estimate the risk of the corresponding estimators. Finally, the percentage reductions in average loss (PRIAL) for $\widehat{\mathbf{T}}$ compared with the MLE and the unbiased estimator, is computed. Which is an estimate of

$$\frac{E\left[L(\mathbf{T}, \widehat{\mathbf{T}}_u) - L(\mathbf{T}, \widehat{\mathbf{T}})\right]}{E\left[L(\mathbf{T}, \widehat{\mathbf{T}}_u)\right]} \times 100$$

Table 1 contains the PRIAL of $\widehat{\mathbf{T}}$ over the MLE $\widehat{\mathbf{T}}_u = \alpha_0 \mathbf{A}^{-1}$, with $\alpha_0 = \dfrac{n}{2q + np}$. While

				diag(1,1,1)	diag(4,2,1)	diag(25,1,1)
$n = 8$	q	$=$	5	23.96	65.08	45.26
$n = 8$	q	$=$	200	57.62	59.31	48.68
$n = 8$	q	$=$	1000	12.49	56.07	49.15
$n = 20$	q	$=$	11	31.90	49.59	25.85
$n = 20$	q	$=$	100	60.44	43.99	24.68
$n = 20$	q	$=$	1000	59.07	42.82	28.64
$n = 30$	q	$=$	19	39.43	28.11	15.68
$n = 30$	q	$=$	150	40.07	24.89	18.42
$n = 30$	q	$=$	1000	47.97	42.34	23.81

Table 1: PRIAL of $\widehat{\mathbf{T}}$ over the MLE

Table 2 gives the PRIAL of $\widehat{\mathbf{T}}$ over the unbiased estimator, with $\alpha_0 = \dfrac{n - p - 2}{2q + np}$.

The results indicate that for the choices of the matrix $\mathbf{\Sigma}$, our proposed estimator $\widehat{\mathbf{T}}$ provides a substantial improvement over the MLE and the unbiased estimator, especially when the sample size n is small. On the other hand, the improvement over the unbiased estimator seems to be a decreasing function on n.

				diag(1,1,1)	diag(4,2,1)	diag(25,1,1)
$n = 8$	q	$=$	5	39.31	34.55	21.64
$n = 8$	q	$=$	200	46.29	38.69	22.47
$n = 8$	q	$=$	1000	46.97	40.00	23.89
$n = 20$	q	$=$	11	6.05	4.68	2.46
$n = 20$	q	$=$	100	8.37	4.15	3.21
$n = 20$	q	$=$	1000	9.97	7.30	3.84
$n = 30$	q	$=$	19	3.29	2.44	2.50
$n = 30$	q	$=$	150	7.71	6.06	2.76
$n = 30$	q	$=$	1000	10.33	8.91	2.75

Table 2: PRIAL of $\widehat{\mathbf{T}}$ over the unbiased estimator

References

[1] Anderson, T.W., Fang, K.T. and Hsu, H. (1986). *Maximum likelihood estimates and likelihood-ratio criteria for multivariate elliptically contoured distributions.* Canad. J. Statist. **14**(1), 55-59.

[2] Dey, D.K. (1988). *Simultaneous estimation of eigenvalues.* Ann. Inst. Stat. Math. **40**, 137-147.

[3] Efron, B. and Morris, C. (1976). *Multivariate empirical Bayes and estimation of covariance matrices.* Ann. Statist. **4**, 22-32.

[4] Fang, K.T. and Anderson, T.W. (1990). *Statistical Inference in Elliptically Contoured and Related Distributions.* Allerton Press, New York.

[5] Fang, K.T., Kotz, S. and Ng, K.W. (1990). *Symmetric Multivariate and Related Distributions.* Chapman and Hall, London, New York

[6] Fang, K.T. and Wang, Y. (1994). *Number-Theoretic Methods in Statistics.* Chapman and Hall, London.

[7] Gupta, A.K. and Nagar, D.K. (2000). *Matrix Variate Distributions*. Boca Raton: Chapman and Hall / CRC

[8] Gupta, A.K. and Varga, T. (1993). *Elliptically Contoured Models in Statistics*. Kluwer Academic Publishers, Dordrecht.

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

[10] Joarder, A.H. (1997). *On the characteristic function of the multivariate Pearson type II distribution*. J. Info. Opt. Sci. **18**(1), 177-182.

[11] Joarder, A.H. and Ahmed, S.E. (1998). *Estimation of the scale matrix of a class of elliptical distributions*. Metrika, **48**, 149-160.

[12] Johnson, M.E. (1987). *Multivariate Statistical Simulation*. Wiley, New York.

[13] Kano, Y. (1994). *Consistency properties of elliptical probability density functions*. J. Multivariate Anal. **51**, 139-147.

[14] Kotz, S. (1975). *Multivariate distributions at a cross-road*. Statistical Distributions in Scientific Work, **1**, 247-270, Reidel, Dordrecht.

[15] Krishnamoorthy, K. and Gupta, A.K. (1989). *Improved minimax estimation of a normal precision matrix*. Canad. J. Statist., **17**, 91-102.

[16] Kubokawa, T. (2005). *A revisit to estimation of the precision matrix of the Wishart distribution*. J. Statist. Res., **39**, 91-114.

[17] Pal, N. (1993). *Estimating the normal dispersion matrix and the precision matrix from a decision-theoretic point of view*. Statist. Papers, **34**, 1-26.

[18] Tsukuma, H. (2005). *Estimating the inverse matrix of scale parameters in an elliptically contoured distribution*. J. Japan Statist. Soc., **35**(1), 21-39.