Ann. Inst. Statist. Math.
Vol. 57, No. 3, 455–484 (2005)

ESTIMATION OF THE MULTIVARIATE NORMAL PRECISION AND COVARIANCE MATRICES IN A STAR-SHAPE MODEL*

DONGCHU SUN AND XIAOQIAN SUN

Department of Statistics, University of Missouri, Columbia, MO 65211, U.S.A.

(Received March 31, 2004; revised September 27, 2004)

Abstract. In this paper, we introduce the star-shape models, where the precision matrix $\mathbf{\Omega}$ (the inverse of the covariance matrix) is structured by the special conditional independence. We want to estimate the precision matrix under entropy loss and symmetric loss. We show that the maximal likelihood estimator (MLE) of the precision matrix is biased. Based on the MLE, an unbiased estimate is obtained. We consider a type of Cholesky decomposition of $\mathbf{\Omega}$, in the sense that $\mathbf{\Omega} = \mathbf{\Psi}'\mathbf{\Psi}$, where $\mathbf{\Psi}$ is a lower triangular matrix with positive diagonal elements. A special group $\mathcal{G}$, which is a subgroup of the group consisting all lower triangular matrices, is introduced. General forms of equivariant estimates of the covariance matrix and precision matrix are obtained. The invariant Haar measures on $\mathcal{G}$, the reference prior, and the Jeffreys prior of $\mathbf{\Psi}$ are also discussed. We also introduce a class of priors of $\mathbf{\Psi}$, which includes all the priors described above. The posterior properties are discussed and the closed forms of Bayesian estimators are derived under either the entropy loss or the symmetric loss. We also show that the best equivariant estimators with respect to $\mathcal{G}$ is the special case of Bayesian estimators. Consequently, the MLE of the precision matrix is inadmissible under either entropy or symmetric loss. The closed form of risks of equivariant estimators are obtained. Some numerical results are given for illustration.

Key words and phrases: Star-shape model, maximum likelihood estimator, precision matrix, covariance matrix, Jeffreys prior, reference prior, invariant Haar measure, Bayesian estimator, entropy loss, symmetric loss, inadmissibility.

1. Introduction

Multivariate normal distributions play an important role in multivariate statistical analysis. There is a large literature on estimating the covariance matrix and precision matrix in the saturated multivariate normal distribution, that is, with no restriction to the covariance matrix except assuming to be positive definite. See Haff (1980), Sinha and Ghosh (1987), Krishnamoorthy and Gupta (1989), Pal (1993), Yang and Berger (1994), and others. However, as the number of variables p in a multivariate distribution increases, the number of parameters $p(p+1)/2$ to be estimated increases fast. Unless the number of observations, n, is very large, estimation is often inefficient, and models with many parameters are, in general, difficult to interpret. In many practical situations, there will be some manifest inter-relationships among several variables. One important

*The project is supported by the National Science Foundation grants DMS-9972598, SES-0095919, and SES-0351523, and a grant from Federal Aid in Wildlife Restoration Project W-13-R through Missouri Department of Conservation.

case uses several pair variables that are conditionally independent given other remaining variables. For multivariate normal distribution, this will correspond to some zeros among the entries of the precision matrix. See Dempster (1972), Whittaker (1990), or Lauritzen (1996). Bayesian model selection of detecting zeros in precision matrix can be found in Wong *et al.* (2003).

Assume that $\boldsymbol{X} \sim N_p(0, \boldsymbol{\Sigma})$. The vector $\boldsymbol{X}$ is partitioned into k groups, that is, $\boldsymbol{X} = (\boldsymbol{X}_1', \boldsymbol{X}_2', \ldots, \boldsymbol{X}_k')'$, where $\boldsymbol{X}_i$ is p_i-dimensional and $\sum_{i=1}^k p_i = p$. We assume that for giving $\boldsymbol{X}_1$, the other subvectors $\boldsymbol{X}_2, \ldots, \boldsymbol{X}_k$ are mutually conditionally independent. From Whittaker (1990) and Lauritzen (1996), the precision matrix $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}$ has the following special structure:

$$\boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\Omega}_{11} & \boldsymbol{\Omega}_{12} & \boldsymbol{\Omega}_{13} & \cdots & \boldsymbol{\Omega}_{1k} \\ \boldsymbol{\Omega}_{21} & \boldsymbol{\Omega}_{22} & \mathbf{0} & \cdots & \mathbf{0} \\ \boldsymbol{\Omega}_{31} & \mathbf{0} & \boldsymbol{\Omega}_{33} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Omega}_{k1} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Omega}_{kk} \end{pmatrix}. \tag{1.1}$$

Now we state several examples in statistics and other fields.

Example 1. Let X_1 be the prime interest rate, which is a global variable, and $X_2, \ldots, X_k$ be the medium house prices in $k-1$ cities in different states, which are local variables. Then $X_2, \ldots, X_k$ are often conditionally independent given X_1 and normally distributed. Hence the precision matrix of $(X_1, \ldots, X_k)$ would have the structure (1.1).

Example 2. Conditional independence assumption is very common in graphical models. The case of $k = 3$ is considered in detail by Whittaker (1990) and is called a "butterfly model." For general $k \geq 3$, we call a normal model with precision matrix (1.1) a *star-shape* model because the graphical shape of the relationship among the variables when $k \geq 4$ looks like a star. The cases when $k = 3, 4, 7$ are illustrated in Fig. 1.

Example 3. *Balanced one-way random effect model.* Suppose that Y_{ij} follow the one-way random effect model,

$$Y_{ij} = \alpha_i + e_{ij}, \qquad j = 1, \ldots, J, \quad i = 1, \ldots, n,$$

where $\alpha_1, \ldots, \alpha_n$ are iid $N(\mu, \tau^2)$, and e_{ij} are iid $N(0, \sigma^2)$. In the case, α_i are often treated as unobserved latent variables. Clearly the joint distribution of $(\alpha_i, Y_{i1}, \ldots, Y_{iJ})$ follow a star-shape model with $k = J + 1$. (Considering latent variable is common in Bayesian context since latent variable is often used in computation.)

The star-shape model is a special case of lattice conditional independence models introduced by Andersson and Perlman (1993). Although star-shaped models or general graphical models have been used widely, as far as we know, fewer theoretic results are obtained on estimating the covariance matrix and the precision matrix in lattice conditional independence models. Andersson and Perlman (1993) gave the form of the maximal likelihood estimator (MLE) of the covariance matrix $\boldsymbol{\Sigma}$. Konno (2001) considered the estimation of the covariance matrix under the Stein loss

$$L_0^*(\hat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) = \mathrm{tr}(\hat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}) - \log|\hat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}| - p \tag{1.2}$$

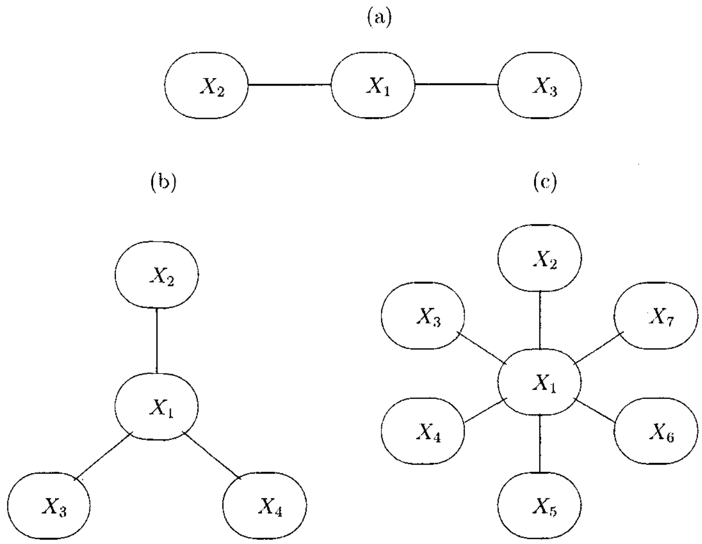

Fig. 1. Several graphs of conditional independence: (a) $k = 3$, (b) $k = 4$, and (c) $k = 7$.

and proved that the MLE of $\boldsymbol{\Sigma}$ is inadmissible.

The Stein loss for estimating the covariance matrix $\boldsymbol{\Sigma}$ is equivalent to the following loss for estimating the precision matrix $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}$:

$$L_0(\hat{\boldsymbol{\Omega}}, \boldsymbol{\Omega}) = \mathrm{tr}(\hat{\boldsymbol{\Omega}}^{-1}\boldsymbol{\Omega}) - \log|\hat{\boldsymbol{\Omega}}^{-1}\boldsymbol{\Omega}| - p. \tag{1.3}$$

Of course, the Stein loss is related to the commonly used entropy loss. See Robert (1994). In this paper, we will consider the estimation of the precision matrix and the covariance matrix in the star-shape model under two loss functions. Let $f(\boldsymbol{x} \mid \boldsymbol{\Omega})$ be the density of $\boldsymbol{X}$ under $\boldsymbol{\Omega}$. The entropy loss for $\boldsymbol{\Omega}$ is

$$\begin{aligned} L_1(\hat{\boldsymbol{\Omega}}, \boldsymbol{\Omega}) &= 2\int \log\left\{\frac{f(\boldsymbol{X} \mid \boldsymbol{\Omega})}{f(\boldsymbol{X} \mid \hat{\boldsymbol{\Omega}})}\right\} f(\boldsymbol{X} \mid \boldsymbol{\Omega})d\boldsymbol{X} \\ &= \mathrm{tr}(\hat{\boldsymbol{\Omega}}\boldsymbol{\Omega}^{-1}) - \log|\hat{\boldsymbol{\Omega}}\boldsymbol{\Omega}^{-1}| - p. \end{aligned} \tag{1.4}$$

The Stein loss is obtained from the entropy loss by switching the role of two arguments, $\hat{\boldsymbol{\Omega}}$ and $\boldsymbol{\Omega}$. The loss function L_1 is typical entropy loss and has been studied by many authors such as Sinha and Ghosh (1987), Krishnamoorthy and Gupta (1989), and others for various contexts.

Note that because neither L_0 nor L_1 is symmetric, we could consider a symmetric version by adding the Stein loss L_0 and entropy loss L_1:

$$L_2(\hat{\boldsymbol{\Omega}}, \boldsymbol{\Omega}) = L_0(\hat{\boldsymbol{\Omega}}, \boldsymbol{\Omega}) + L_1(\hat{\boldsymbol{\Omega}}, \boldsymbol{\Omega}) = \mathrm{tr}(\hat{\boldsymbol{\Omega}}\boldsymbol{\Omega}^{-1}) + \mathrm{tr}(\hat{\boldsymbol{\Omega}}^{-1}\boldsymbol{\Omega}) - 2p. \tag{1.5}$$

The symmetric loss L_2 for $\boldsymbol{\Omega}$ introduced by Kubokawa and Konno (1990) and Gupta and Ofori-Nyarko (1995), can be seen as estimating the covariance matrix and the precision matrix simultaneously. The corresponding entropy loss and symmetric loss for the covariance $\boldsymbol{\Sigma}$ can be obtained by replacing $\boldsymbol{\Omega}^{-1}$ by $\boldsymbol{\Sigma}$ and $\hat{\boldsymbol{\Omega}}^{-1}$ by $\hat{\boldsymbol{\Sigma}}$. We will focus on estimation of the precision matrix $\boldsymbol{\Omega}$ because the results for $\boldsymbol{\Sigma}$ can be derived as corollaries of the corresponding results for $\boldsymbol{\Omega}$.

The remainder of this paper is organized as follows. In Section 2, we give the MLE of the precision matrix in the star-shape model and prove that it is not unbiased. Based on the MLE, an unbiased estimate of the precision matrix is given. In Section 3, we first consider a type of Cholesky decomposition of the precision matrix, that is, $\boldsymbol{\Omega} = \boldsymbol{\Psi}'\boldsymbol{\Psi}$, where $\boldsymbol{\Psi}$ is a lower triangular matrix with positive diagonal elements and then get the special group $\mathcal{G}$ of lower-triangular block matrices. Interestingly, the problem is invariant under the matrix operation from $\mathcal{G}$, instead of traditional Cholesky decomposition. The invariant Haar measures of this group are given, and we also prove that the Jeffreys prior of $\boldsymbol{\Psi}$ matrix is exactly the same as the right invariant Haar measure on the group $\mathcal{G}$. A reference prior is obtained by using the algorithm in Berger and Bernardo (1992). The closed form of equivariant estimators of $\boldsymbol{\Omega}$ also is derived.

In Section 4, we introduce a class of priors of $\boldsymbol{\Psi}$, which includes all priors such as the left and right Haar measures and the reference prior as special cases. Some properties on the posterior under such class of priors are discussed. In Section 5, the closed form of Bayesian estimators with respect to such a class of priors is obtained under the entropy loss. We find that these Bayesian estimators include a lot of usual estimators such as the MLE, the best equivariant estimates under the group $\mathcal{G}$. From this, we also know that the MLE of $\boldsymbol{\Omega}$ is inadmissible under the entropy loss. Results on the symmetric loss are shown in Section 6. The risks of equivariant estimators are given in Section 7. The results on estimating covariance matrix are given in Section 8. Finally, some numerical results also are given in Section 9.

2. The MLE and unbiased estimator

Let $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_n$ be a random sample from $N_p(0, \boldsymbol{\Omega}^{-1})$, where $\boldsymbol{\Omega}$ satisfies (1.1). Let $\boldsymbol{S} = \sum_{i=1}^n \boldsymbol{Y}_i \boldsymbol{Y}_i'$. Assume $n > p+1$ throughout this paper. Then $\boldsymbol{S}$ is a sufficient statistic of $\boldsymbol{\Sigma}$ or $\boldsymbol{\Omega}$ and has a Wishart distribution with parameters n and $\boldsymbol{\Omega}^{-1}$. Write $\boldsymbol{S} = (\boldsymbol{S}_{ij})$, where $\boldsymbol{S}_{ij}$ is the $p_i \times p_j$ sub-matrix. Whittaker (1990) gives the expression of MLE of the covariance matrix $\boldsymbol{\Sigma}$ for $k=3$. For general k, it is easy to show that the MLE of $\boldsymbol{\Sigma}$ has the expression

$$\hat{\boldsymbol{\Sigma}}_M = \frac{1}{n}\begin{pmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} & \boldsymbol{S}_{13} & \cdots & \boldsymbol{S}_{1k} \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} & \boldsymbol{S}_{21}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{13} & \cdots & \boldsymbol{S}_{21}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1k} \\ \boldsymbol{S}_{31} & \boldsymbol{S}_{31}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12} & \boldsymbol{S}_{33} & \cdots & \boldsymbol{S}_{31}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{S}_{k1} & \boldsymbol{S}_{k1}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12} & \boldsymbol{S}_{k1}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{13} & \cdots & \boldsymbol{S}_{kk} \end{pmatrix}. \tag{2.1}$$

Define

$$\boldsymbol{S}_{ii\cdot 1} = \boldsymbol{S}_{ii} - \boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{i1}, \qquad i = 2, \ldots, k. \tag{2.2}$$

By (2.1), the MLE $\hat{\boldsymbol{\Omega}}_M$ of the precision matrix $\boldsymbol{\Omega}$ can be obtained by

$$\hat{\boldsymbol{\Omega}}_M = \begin{pmatrix} \hat{\boldsymbol{\Omega}}_{11}^M & \hat{\boldsymbol{\Omega}}_{12}^M & \cdots & \hat{\boldsymbol{\Omega}}_{1k}^M \\ \hat{\boldsymbol{\Omega}}_{21}^M & \hat{\boldsymbol{\Omega}}_{22}^M & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\boldsymbol{\Omega}}_{k1}^M & \mathbf{0} & \cdots & \hat{\boldsymbol{\Omega}}_{kk}^M \end{pmatrix}, \tag{2.3}$$

where

$$\begin{aligned}\hat{\boldsymbol{\Omega}}_{11}^{M} &= n\boldsymbol{S}_{11}^{-1} + n\sum_{i=2}^{k}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1i}\boldsymbol{S}_{ii\cdot 1}^{-1}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1};\\ \hat{\boldsymbol{\Omega}}_{1i}^{M} &= -n\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1i}\boldsymbol{S}_{ii\cdot 1}^{-1},\\ \hat{\boldsymbol{\Omega}}_{ii}^{M} &= n\boldsymbol{S}_{ii\cdot 1}^{-1}, \qquad i = 2,\dots,k.\end{aligned}$$

The following proposition shows that $\hat{\boldsymbol{\Sigma}}_M$ is an unbiased estimate of $\boldsymbol{\Sigma}$ while $\hat{\boldsymbol{\Omega}}_M$ is not an unbiased estimate of $\boldsymbol{\Omega}$.

PROPOSITION 2.1. *Consider the star-shape model.*
(a) *The MLE* $\hat{\boldsymbol{\Sigma}}_M$ *in* (2.1) *is an unbiased estimate of* $\boldsymbol{\Sigma}$;
(b) *The MLE* $\hat{\boldsymbol{\Omega}}_M$ *in* (2.3) *is not an unbiased estimate of* $\boldsymbol{\Omega}$.

PROOF. It is well-known that $E(\boldsymbol{S}_{ij}) = n\boldsymbol{\Sigma}_{ij}$, $i,j = 1,2,\dots,k$ because $\boldsymbol{S} \sim W_p(n,\boldsymbol{\Sigma})$. For (a), we need to prove

$$E(\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1j}) = n\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1j}, \qquad 2 \le i < j \le k. \tag{2.4}$$

We will prove (2.4) for the case when $(i,j) = (2,3)$ only. Other cases are similar.

For convenience, let

$$\boldsymbol{\Sigma}_{ii\cdot 1} = \boldsymbol{\Sigma}_{ii} - \boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}, \qquad i = 2,\cdots,k. \tag{2.5}$$

By

$$\begin{pmatrix}\boldsymbol{S}_{11} & \boldsymbol{S}_{12} & \boldsymbol{S}_{13}\\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} & \boldsymbol{S}_{23}\\ \boldsymbol{S}_{31} & \boldsymbol{S}_{32} & \boldsymbol{S}_{33}\end{pmatrix} \sim W_{p_1+p_2+p_3}\left(n, \begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \boldsymbol{\Sigma}_{13}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{13}\\ \boldsymbol{\Sigma}_{31} & \boldsymbol{\Sigma}_{31}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} & \boldsymbol{\Sigma}_{33}\end{pmatrix}\right)$$

and using the property of the marginal and conditional distributions for Wishart distribution (see Theorem 3.3.9 of Gupta and Nagar (2000)), we have

$$\begin{pmatrix}\boldsymbol{S}_{21}\\ \boldsymbol{S}_{31}\end{pmatrix} \mid \boldsymbol{S}_{11} \sim N_{p_2+p_3,p_1}\left(\begin{pmatrix}\boldsymbol{\Sigma}_{21}\\ \boldsymbol{\Sigma}_{31}\end{pmatrix}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{S}_{11}, \begin{pmatrix}\boldsymbol{\Sigma}_{22\cdot 1} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{\Sigma}_{33\cdot 1}\end{pmatrix} \otimes \boldsymbol{S}_{11}\right). \tag{2.6}$$

Here we exploit the notation for matrix variate normal distribution given by Definition 2.2.1 of Gupta and Nagar (2000), that is, $\boldsymbol{X}_{p\times n} \sim N_{p,n}(\boldsymbol{M}_{p\times n}, \boldsymbol{\Sigma}\otimes\boldsymbol{\Psi})$ if and only if $\mathrm{vec}(\boldsymbol{X}') \sim N_{pn}(\mathrm{vec}(\boldsymbol{M}'), \boldsymbol{\Sigma}\otimes\boldsymbol{\Psi})$. Applying Theorem 2.3.5 of Gupta and Nagar (2000), it follows that

$$\begin{aligned}E(\boldsymbol{S}_{21}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{13} \mid \boldsymbol{S}_{11}) &= (\boldsymbol{I}_{p_2}\ \ \boldsymbol{0})E\left\{\begin{pmatrix}\boldsymbol{S}_{21}\\ \boldsymbol{S}_{31}\end{pmatrix}\boldsymbol{S}_{11}^{-1}\begin{pmatrix}\boldsymbol{S}_{21}\\ \boldsymbol{S}_{31}\end{pmatrix}'\right\}\begin{pmatrix}\boldsymbol{0}\\ \boldsymbol{I}_{p_3}\end{pmatrix}\\ &= \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{S}_{11}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{13}.\end{aligned}$$

Because $E(\boldsymbol{S}_{11}) = n\boldsymbol{\Sigma}_{11}$, we get $E(\boldsymbol{S}_{21}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{13}) = n\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{13}$. Thus (2.4) holds for $(i,j) = (2,3)$. For (b), because

$$\begin{pmatrix}\boldsymbol{S}_{11} & \boldsymbol{S}_{1i}\\ \boldsymbol{S}_{i1} & \boldsymbol{S}_{ii}\end{pmatrix} \sim W_{p_1+p_i}\left(n, \begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{1i}\\ \boldsymbol{\Sigma}_{i1} & \boldsymbol{\Sigma}_{ii}\end{pmatrix}\right), \tag{2.7}$$

we have

$$\boldsymbol{S}_{11} \sim W_{p_1}(n, \boldsymbol{\Sigma}_{11}), \qquad \boldsymbol{S}_{ii\cdot 1} \sim W_{p_i}(n - p_1, \boldsymbol{\Sigma}_{ii\cdot 1}), \qquad i = 2, \ldots, k. \tag{2.8}$$

Thus

$$E(\boldsymbol{S}_{11}^{-1}) = \frac{1}{n - p_1 - 1}\boldsymbol{\Sigma}_{11}^{-1}, \qquad E(\boldsymbol{S}_{ii\cdot 1}^{-1}) = \frac{1}{n - p_1 - p_i - 1}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1}, \qquad i = 2, \ldots, k.$$

The proof is completed.

Based on the MLE $\hat{\boldsymbol{\Omega}}_M$ in (2.3), we create an unbiased estimate of $\boldsymbol{\Omega}$ in Proposition 2.2.

PROPOSITION 2.2. *Under the star-shape model, an unbiased estimate of $\boldsymbol{\Omega}$ is*

$$\hat{\boldsymbol{\Omega}}_U = \begin{pmatrix} \hat{\boldsymbol{\Omega}}_{11}^U & \hat{\boldsymbol{\Omega}}_{12}^U & \cdots & \hat{\boldsymbol{\Omega}}_{1k}^U \\ \hat{\boldsymbol{\Omega}}_{21}^U & \hat{\boldsymbol{\Omega}}_{22}^U & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\boldsymbol{\Omega}}_{k1}^U & \mathbf{0} & \cdots & \hat{\boldsymbol{\Omega}}_{kk}^U \end{pmatrix}, \tag{2.9}$$

where

$$\begin{aligned} \hat{\boldsymbol{\Omega}}_{11}^U &= (n - p - 1)\boldsymbol{S}_{11}^{-1} + \sum_{i=2}^{k}(n - p_1 - p_i - 1)\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1i}\boldsymbol{S}_{ii\cdot 1}^{-1}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1}; \\ \hat{\boldsymbol{\Omega}}_{1i}^U &= -(n - p_1 - p_i - 1)\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1i}\boldsymbol{S}_{ii\cdot 1}^{-1}, \\ \hat{\boldsymbol{\Omega}}_{ii}^U &= (n - p_1 - p_i - 1)\boldsymbol{S}_{ii\cdot 1}^{-1}, \qquad i = 2, \ldots, k. \end{aligned}$$

PROOF. First, note that the following relationships between $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}$ and $\boldsymbol{\Sigma}$,

$$\begin{aligned} \boldsymbol{\Omega}_{11} &= \boldsymbol{\Sigma}_{11}^{-1} + \sum_{i=2}^{k}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1}\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}; \\ \boldsymbol{\Omega}_{1i} &= -\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1}, \\ \boldsymbol{\Omega}_{ii} &= \boldsymbol{\Sigma}_{ii\cdot 1}^{-1}, \qquad i = 2, \ldots, k. \end{aligned}$$

From (2.7), we have

$$\boldsymbol{S}_{i1} \mid \boldsymbol{S}_{11} \sim N_{p_i, p_1}(\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{S}_{11}, \boldsymbol{\Sigma}_{ii\cdot 1} \otimes \boldsymbol{S}_{11}), \tag{2.10}$$

and $(\boldsymbol{S}_{i1}, \boldsymbol{S}_{11})$ is independent of $\boldsymbol{S}_{ii\cdot 1}$, where $\boldsymbol{S}_{ii\cdot 1}$ is defined by (2.2). Therefore, we get

$$\begin{aligned} E(\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1i}\boldsymbol{S}_{ii\cdot 1}^{-1} \mid \boldsymbol{S}_{11}) &= \boldsymbol{S}_{11}^{-1}E(\boldsymbol{S}_{1i} \mid \boldsymbol{S}_{11})E(\boldsymbol{S}_{ii\cdot 1}^{-1}) \\ &= \frac{1}{n - p_1 - p_i - 1}\boldsymbol{S}_{11}^{-1} \cdot (\boldsymbol{S}_{11}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}) \cdot \boldsymbol{\Sigma}_{ii\cdot 1}^{-1} \\ &= \frac{1}{n - p_1 - p_i - 1}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1} \end{aligned}$$

and

$$
\begin{aligned}
E(\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1i}\boldsymbol{S}_{ii\cdot 1}^{-1}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1} \mid \boldsymbol{S}_{11}) &= \frac{1}{n-p_1-p_i-1}\boldsymbol{S}_{11}^{-1}E(\boldsymbol{S}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1}\boldsymbol{S}_{i1} \mid \boldsymbol{S}_{11})\boldsymbol{S}_{11}^{-1} \\
&= \frac{p_i}{n-p_1-p_i-1}\boldsymbol{S}_{11}^{-1} \\
&\quad + \frac{1}{n-p_1-p_i-1}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1}\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}.
\end{aligned}
$$

So we get

$$
E(\hat{\boldsymbol{\Omega}}_{11}^{U}) = \left(n-p-1+\sum_{i=2}^{k}p_i\right)E(\boldsymbol{S}_{11}^{-1}) + \sum_{i=2}^{k}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1}\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}.
$$

Because $E(\boldsymbol{S}_{11}^{-1}) = \boldsymbol{\Sigma}_{11}^{-1}/(n-p_1-1)$, we have $E(\hat{\boldsymbol{\Omega}}_{11}^{U}) = \boldsymbol{\Omega}_{11}$. For $i = 2,\ldots,k$,

$$
\begin{aligned}
E(\hat{\boldsymbol{\Omega}}_{1i}^{U}) &= -(n-p_1-p_i-1)E(\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1i}\boldsymbol{S}_{ii\cdot 1}^{-1}) = -\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1} = \boldsymbol{\Omega}_{1i}, \\
E(\hat{\boldsymbol{\Omega}}_{ii}^{U}) &= (n-p_1-p_i-1)E(\boldsymbol{S}_{ii\cdot 1}^{-1}) = \boldsymbol{\Sigma}_{ii\cdot 1}^{-1} = \boldsymbol{\Omega}_{ii}.
\end{aligned}
$$

The proof is completed.

3. Cholesky decomposition and noninformative priors

3.1 *The role of group invariance*

It is well known that group invariance plays an important role in finding better estimates of the covariance and the precision matrices in multivariate normal models. With an appropriate group $\mathcal{G}$, if the model is invariant under the transformation using the element from the group to the data, we can restrict our attention to a class of equivariant estimators obtained by the group invariance principle. This class often includes the MLE under some mild conditions. Then a frequentist method may be applied to get the best equivariant estimator in this class. The procedure is to characterize the functional form of any equivariant estimator with respect to $\mathcal{G}$. The best equivariant estimator can be obtained by minimizing the risk of an equivariant estimator with respect to the given loss function. The best equivariant estimator is often superior to the MLE. See for example, James and Stein (1961), Olkin and Selliah (1977), Sharma and Krishnamoorthy (1983), etc.

Two drawbacks exist with this frequentist method. One drawback is that it might be difficult to get the closed form of equivariant estimators with respect to the group. Furthermore, it could be more difficult to calculate risks for an equivariant estimator. Alternatively, a Bayesian method may be considered. Eaton (1989) showed that under some conditions, the best equivariant estimator will be a Bayesian estimator if we take the right Haar measure on the group as a prior. See Chapter 6 of Eaton (1989) for details. With the Bayesian method, Eaton (1970) successfully obtained the best equivariant estimator of the covariance matrix with respect to the group of lower triangular matrices in multivariate normal distribution with missing data under the Stein loss $L_0^*(\hat{\boldsymbol{\Sigma}},\boldsymbol{\Sigma})$ although the general form of equivariant estimators is not derived because of complication. Konno (2001) got a similar result in lattice conditional independence models introduced by Andersson and Perlman (1993).

Note that to exploit the frequentist method or the Bayesian method described above, it is crucial to choose an appropriate group that makes the model invariant. It could be difficult to get an appropriate group when the precision matrix or the covariance matrix is structured. Fortunately, such a group does exist under the star-shape model by Andersson and Perlman (1993). We will describe this group in detail for our statistical inference.

3.2 *Cholesky decomposition*

We now consider how to get better estimators over the MLE $\hat{\boldsymbol{\Omega}}_M$ and the unbiased estimator $\hat{\boldsymbol{\Omega}}_U$ of the precision matrix under the entropy loss L_1 and the symmetric loss L_2. First, we decompose the precision matrix or the covariance matrix as

$$\boldsymbol{\Omega} = \boldsymbol{\Psi}'\boldsymbol{\Psi} \quad \text{or} \quad \boldsymbol{\Sigma} = \boldsymbol{\Delta}\boldsymbol{\Delta}', \tag{3.1}$$

where both $\boldsymbol{\Psi}$ and $\boldsymbol{\Delta}$ are p by p lower-triangular matrices with positive diagonal entries. For convenience, $\boldsymbol{\Delta}$ will be viewed as the Cholesky decomposition of $\boldsymbol{\Sigma}$. From the structure of $\boldsymbol{\Omega}$ given by (1.1), it is easy to show that $\boldsymbol{\Psi}$ has the following block structure:

$$\boldsymbol{\Psi} = \begin{pmatrix} \boldsymbol{\Psi}_{11} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \boldsymbol{\Psi}_{21} & \boldsymbol{\Psi}_{22} & \mathbf{0} & \cdots & \mathbf{0} \\ \boldsymbol{\Psi}_{31} & \mathbf{0} & \boldsymbol{\Psi}_{33} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}_{k1} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Psi}_{kk} \end{pmatrix}, \tag{3.2}$$

where $\boldsymbol{\Psi}_{ii}$ is a $p_i \times p_i$ lower-triangular matrix. Note that there is no restriction on $\boldsymbol{\Psi}_{ij}$ $(i \geq j)$ except requiring that all diagonal elements of $\boldsymbol{\Psi}_{ii}$ are positive. Define

$$\mathcal{G} = \{\boldsymbol{A} \in R^{p\times p} \mid \boldsymbol{A} \text{ has a structure as (3.2)}\}. \tag{3.3}$$

We have the following result:

LEMMA 3.1. (a) *$\mathcal{G}$ is a group with respect to matrix multiplication.* (b) *For any $\boldsymbol{A} \in \mathcal{G}$, $\boldsymbol{A}^{-1}$ has the expression,*

$$\boldsymbol{A}^{-1} = \begin{pmatrix} \boldsymbol{A}_{11}^{-1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{A}_{22}^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ -\boldsymbol{A}_{33}^{-1}\boldsymbol{A}_{31}\boldsymbol{A}_{11}^{-1} & \mathbf{0} & \boldsymbol{A}_{33}^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\boldsymbol{A}_{kk}^{-1}\boldsymbol{A}_{k1}\boldsymbol{A}_{11}^{-1} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{A}_{kk}^{-1} \end{pmatrix}. \tag{3.4}$$

PROOF. Note that $\mathcal{G}$ is a subset of the group of all $p \times p$ lower triangular matrices. The results then follow from Andersson and Perlman (1993) or can be verified directly.

Note that the star-shape model is invariant under the group $\mathcal{G}$ and so are the entropy loss L_1 and the symmetric loss L_2. In the following subsection, we will give the general form of equivariant estimators of $\boldsymbol{\Omega}$ or $\boldsymbol{\Sigma}$ with respect to the group $\mathcal{G}$.

3.3 *The general forms of equivariant estimates of* $\boldsymbol{\Sigma}$ *and* $\boldsymbol{\Omega}$

An estimate $\hat{\boldsymbol{\Sigma}}(\boldsymbol{S})$ of $\boldsymbol{\Sigma}$ is called equivariant under the group $\mathcal{G}$ if and only if for any $\boldsymbol{A} \in \mathcal{G}$,

$$\hat{\boldsymbol{\Sigma}}(\boldsymbol{ASA}') = \boldsymbol{A}\hat{\boldsymbol{\Sigma}}(\boldsymbol{S})\boldsymbol{A}'. \tag{3.5}$$

For estimating the precision matrix $\boldsymbol{\Omega}$, $\hat{\boldsymbol{\Omega}}(\boldsymbol{S})$ is equivariant under the group $\mathcal{G}$ if and only if

$$\hat{\boldsymbol{\Omega}}(\boldsymbol{ASA}') = (\boldsymbol{A}')^{-1}\hat{\boldsymbol{\Omega}}(\boldsymbol{S})\boldsymbol{A}^{-1} \tag{3.6}$$

for any $\boldsymbol{A} \in \mathcal{G}$.

For the star-shape model, it is easy to show that $\boldsymbol{S}_{11}, \boldsymbol{S}_{12}, \dots, \boldsymbol{S}_{1k}, \boldsymbol{S}_{22}, \dots, \boldsymbol{S}_{kk}$ are sufficient statistics for $\boldsymbol{\Sigma}$ (see (4.2) and (4.3) for details). Then $\tilde{\boldsymbol{S}} = n\hat{\boldsymbol{\Sigma}}_{MLE}$ is also a sufficient statistic for $\boldsymbol{\Sigma}$. Thus we can easily see that for any equivariant estimate $\hat{\boldsymbol{\Sigma}}(\boldsymbol{S})$ satisfying (3.5), there will be $\hat{\boldsymbol{\Sigma}}_0(\tilde{\boldsymbol{S}}) = \hat{\boldsymbol{\Sigma}}(\boldsymbol{S})$ a.s. such that

$$\hat{\boldsymbol{\Sigma}}_0(\boldsymbol{A}\tilde{\boldsymbol{S}}\boldsymbol{A}') = \boldsymbol{A}\hat{\boldsymbol{\Sigma}}_0(\tilde{\boldsymbol{S}})\boldsymbol{A}'. \tag{3.7}$$

Let $\boldsymbol{S}_{11\cdot 1} = \boldsymbol{S}_{11}$ for convenience. Define

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{T}_{11} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{S}_{21}(\boldsymbol{T}_{11}')^{-1} & \boldsymbol{T}_{22} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{S}_{31}(\boldsymbol{T}_{11}')^{-1} & \boldsymbol{0} & \boldsymbol{T}_{33} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{S}_{k1}(\boldsymbol{T}_{11}')^{-1} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{T}_{kk} \end{pmatrix}, \tag{3.8}$$

where $\boldsymbol{T}_{ii}$ is Cholesky decomposition of $\boldsymbol{S}_{ii\cdot 1}$, $i = 1, 2, \dots, k$. Then $\tilde{\boldsymbol{S}} = \boldsymbol{TT}'$ and $\boldsymbol{T} \in \mathcal{G}$. Putting $\boldsymbol{A} = \boldsymbol{T}^{-1}$ in (3.7) gives $\hat{\boldsymbol{\Sigma}}_0(\tilde{\boldsymbol{S}}) = \boldsymbol{T}\hat{\boldsymbol{\Sigma}}_0(\boldsymbol{I})\boldsymbol{T}' = \boldsymbol{TWT}'$, where $\boldsymbol{W} = \hat{\boldsymbol{\Sigma}}_0(\boldsymbol{I})$ is a constant matrix that can be expressed as $\boldsymbol{PP}'$ for a $\boldsymbol{P} \in \mathcal{G}$. In addition, for any $\boldsymbol{P} \in \mathcal{G}$, the estimate $\hat{\boldsymbol{\Sigma}}_0(\tilde{\boldsymbol{S}}) = \boldsymbol{TPP}'\boldsymbol{T}'$ satisfies (3.7) obviously. Hence the general form of equivariant estimates of $\boldsymbol{\Sigma}$ is

$$\hat{\boldsymbol{\Sigma}}(\boldsymbol{S}) = \boldsymbol{TPP}'\boldsymbol{T}', \tag{3.9}$$

where $\boldsymbol{T}$ is defined by (3.8) and $\boldsymbol{P} \in \mathcal{G}$ is a constant matrix.

Let

$$\boldsymbol{R} = \boldsymbol{T}^{-1} = \begin{pmatrix} \boldsymbol{T}_{11}^{-1} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ -\boldsymbol{T}_{22}^{-1}\boldsymbol{S}_{21}\boldsymbol{S}_{11}^{-1} & \boldsymbol{T}_{22}^{-1} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ -\boldsymbol{T}_{33}^{-1}\boldsymbol{S}_{31}\boldsymbol{S}_{11}^{-1} & \boldsymbol{0} & \boldsymbol{T}_{33}^{-1} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\boldsymbol{T}_{kk}^{-1}\boldsymbol{S}_{k1}\boldsymbol{S}_{11}^{-1} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{T}_{kk}^{-1} \end{pmatrix}. \tag{3.10}$$

Thus the general form of equivariant estimates of $\boldsymbol{\Omega}$ is

$$\hat{\boldsymbol{\Omega}}(\boldsymbol{S}) = \boldsymbol{R}'\boldsymbol{Q}'\boldsymbol{QR}, \tag{3.11}$$

where $\boldsymbol{Q} \in \mathcal{G}$ is a constant matrix.

Remark 1. Both the MLE $\hat{\boldsymbol{\Omega}}_M$ and the unbiased estimator $\hat{\boldsymbol{\Omega}}_U$ obtained in Proposition 2.2 are equivariant with respect to $\mathcal{G}$. In detail, the MLE of $\boldsymbol{\Omega}$ given by (2.3) can be expressed as

$$\hat{\boldsymbol{\Omega}}_M = n\boldsymbol{R}'\boldsymbol{R}. \tag{3.12}$$

And the unbiased estimate of $\boldsymbol{\Omega}$ given by (2.9) can be expressed as

$$\hat{\boldsymbol{\Omega}}_U = \boldsymbol{R}'\boldsymbol{U}\boldsymbol{R}, \tag{3.13}$$

where

$$\boldsymbol{U} = \mathrm{diag}\{(n-p-1)\boldsymbol{I}_{p_1}, (n-p_1-p_2-1)\boldsymbol{I}_{p_2}, \ldots, (n-p_1-p_k-1)\boldsymbol{I}_{p_k}\}. \tag{3.14}$$

Remark 2. By (3.11), any estimator that has the following expression

$$\hat{\boldsymbol{\Omega}} = \boldsymbol{R}'\boldsymbol{Q}\boldsymbol{R} \tag{3.15}$$

will be a $\mathcal{G}$-equivariant estimator of $\boldsymbol{\Omega}$, where $\boldsymbol{Q}$ is an arbitrary diagonal matrix with constant positive diagonal entries. This is an important class because we will show later that each of them is a Bayesian estimator under either L_1 or L_2.

3.4 *Invariant Haar measures and noninformative priors*

Now we discuss the invariant Haar measures on the group $\mathcal{G}$ and some other noninformative priors.

For any $i = 1, 2, \ldots, k$, let

$$\boldsymbol{\Psi}_{ii} = \begin{pmatrix} \psi_{i11} & 0 & \cdots & 0 \\ \psi_{i21} & \psi_{i22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{ip_i1} & \psi_{ip_i2} & \cdots & \psi_{ip_ip_i} \end{pmatrix}. \tag{3.16}$$

And for $i = 2, \ldots, k$, let

$$\boldsymbol{\Psi}_{i1} = \begin{pmatrix} \phi_{i11} & \phi_{i12} & \cdots & \phi_{i1p_1} \\ \phi_{i21} & \phi_{i22} & \cdots & \phi_{i2p_1} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{ip_i1} & \phi_{ip_i2} & \cdots & \phi_{ip_ip_1} \end{pmatrix}. \tag{3.17}$$

Similar to Example 1.14 of Eaton (1989), we have the lemma.

LEMMA 3.2. *The left invariant Haar measure on $\mathcal{G}$ is*

$$\nu_{\mathcal{G}}^l(d\boldsymbol{\Psi}) = \frac{d\boldsymbol{\Psi}}{\prod_{j=1}^{p_1} \psi_{1jj}^j \cdot \prod_{i=2}^k \prod_{j=1}^{p_i} \psi_{ijj}^{p_1+j}} \tag{3.18}$$

while the right invariant Haar measure on $\mathcal{G}$ is

$$\nu_{\mathcal{G}}^r(d\boldsymbol{\Psi}) = \frac{d\boldsymbol{\Psi}}{\prod_{j=1}^{p_1} \psi_{1jj}^{p-j+1} \cdot \prod_{i=2}^k \prod_{j=1}^{p_i} \psi_{ijj}^{p_i-j+1}}. \tag{3.19}$$

Remark 3. For $\boldsymbol{\Delta} = \boldsymbol{\Psi}^{-1}$, we can readily verify that $\nu_{\mathcal{G}}^{r}(d\boldsymbol{\Psi}) = \nu_{\mathcal{G}}^{l}(d\boldsymbol{\Delta})$ and $\nu_{\mathcal{G}}^{l}(d\boldsymbol{\Psi}) = \nu_{\mathcal{G}}^{r}(d\boldsymbol{\Delta})$.

The invariant Haar measures plays a crucial role in finding a better estimator of the precision matrix or the covariance matrix (see Eaton (1989) for details). The following proposition will give other two noninformative priors of $\boldsymbol{\Psi}$ in the star-shape model.

PROPOSITION 3.1. *Consider the star-shape model.*
(a) *The Jeffreys prior* $\pi_J(\boldsymbol{\Psi})$ *of* $\boldsymbol{\Psi}$ *is the same as the right invariant Haar measure of* $\mathcal{G}$ *given by* (3.19).
(b) *The reference prior of* $\boldsymbol{\Psi}$ *for the ordered group* $\{\psi_{111}, (\psi_{121}, \psi_{122}), \ldots, (\psi_{1p_1 1}, \ldots, \psi_{1p_1p_1}), (\phi_{211}, \ldots, \phi_{21p_1}, \psi_{211}), \ldots, (\phi_{kp_k1}, \ldots, \phi_{kp_kp_1}, \psi_{kp_k1}, \ldots, \psi_{kp_kp_k})\}$ *is given by*

$$\pi_R(d\boldsymbol{\Psi}) \propto \frac{d\boldsymbol{\Psi}}{\prod_{i=1}^{k}\prod_{j=1}^{p_i}\psi_{ijj}}. \tag{3.20}$$

PROOF. Let $\boldsymbol{\theta} = (\psi_{111}, \psi_{121}, \psi_{122}, \ldots, \psi_{1p_11}, \ldots, \psi_{1p_1p_1}, \phi_{211}, \ldots, \phi_{21p_1}, \psi_{211}, \ldots, \phi_{kp_k1}, \ldots, \phi_{kp_kp_1}, \psi_{kp_k1}, \ldots, \psi_{kp_kp_k})'$ and $\boldsymbol{I}_i$ be the $i \times i$ identity matrix and $\boldsymbol{e}_i$ be the $i \times 1$ vector with the i-th element 1 and others 0. Because the likelihood function of $\boldsymbol{\Psi}$ is

$$f(\boldsymbol{X} \mid \boldsymbol{\Psi}) \propto |\boldsymbol{\Psi}'\boldsymbol{\Psi}|^{1/2} \exp\left(-\frac{1}{2}\boldsymbol{X}'\boldsymbol{\Psi}'\boldsymbol{\Psi}\boldsymbol{X}\right),$$

the log-likelihood is then

$$\begin{aligned}\log f &= \text{const} + \sum_{i=1}^{k} \log|\boldsymbol{\Psi}_{ii}| \\ &\quad - \frac{1}{2}\left(\sum_{i=1}^{k} \boldsymbol{X}_i'\boldsymbol{\Psi}_{ii}'\boldsymbol{\Psi}_{ii}\boldsymbol{X}_i + 2\sum_{i=2}^{k} \boldsymbol{X}_1'\boldsymbol{\Psi}_{i1}'\boldsymbol{\Psi}_{ii}\boldsymbol{X}_i + \sum_{i=2}^{k} \boldsymbol{X}_1'\boldsymbol{\Psi}_{i1}'\boldsymbol{\Psi}_{i1}\boldsymbol{X}_1\right).\end{aligned}$$

The Fisher information matrix of $\boldsymbol{\theta}$ is

$$\begin{aligned}\boldsymbol{\Lambda}(\boldsymbol{\theta}) &= -E\left(\frac{\partial^2 \log f}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right) \\ &= \operatorname{diag}(\boldsymbol{\Lambda}_{11}, \ldots, \boldsymbol{\Lambda}_{1p_1}, \boldsymbol{\Lambda}_{21}, \ldots, \boldsymbol{\Lambda}_{2p_2}, \ldots, \boldsymbol{\Lambda}_{k1}, \ldots, \boldsymbol{\Lambda}_{kp_k}),\end{aligned} \tag{3.21}$$

where

$$\begin{aligned}\boldsymbol{\Lambda}_{1p_1} &= \operatorname{Var}(\boldsymbol{X}_1) + \frac{1}{\psi_{1p_1p_1}^2}\boldsymbol{e}_{p_1}\boldsymbol{e}_{p_1}'; \\ \boldsymbol{\Lambda}_{1j} &= (\boldsymbol{I}_j \ \ \boldsymbol{0})\operatorname{Var}(\boldsymbol{X}_1)(\boldsymbol{I}_j \ \ \boldsymbol{0})' + \frac{1}{\psi_{1jj}^2}\boldsymbol{e}_j\boldsymbol{e}_j', \qquad j = 1, 2, \ldots, p_1 - 1; \\ \boldsymbol{\Lambda}_{ip_i} &= \operatorname{Var}\begin{pmatrix}\boldsymbol{X}_1 \\ \boldsymbol{X}_i\end{pmatrix} + \frac{1}{\psi_{ip_ip_i}^2}\boldsymbol{e}_{p_1+p_i}\boldsymbol{e}_{p_1+p_i}', \qquad i = 2, \ldots, k; \\ \boldsymbol{\Lambda}_{ij} &= (\boldsymbol{I}_{p_1+j} \ \ \boldsymbol{0})\operatorname{Var}\begin{pmatrix}\boldsymbol{X}_1 \\ \boldsymbol{X}_i\end{pmatrix}\begin{pmatrix}\boldsymbol{I}_{p_1+j} \\ \boldsymbol{0}\end{pmatrix} \\ &\quad + \frac{1}{\psi_{ijj}^2}\boldsymbol{e}_{p_1+j}\boldsymbol{e}_{p_1+j}', \qquad j = 1, 2, \ldots, p_i - 1, \ i = 2, \ldots, k.\end{aligned}$$

Because $\boldsymbol{\Sigma} = \boldsymbol{\Omega}^{-1} = (\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}$, we get $\mathrm{Var}(\boldsymbol{X}_1) = (\boldsymbol{\Psi}_{11}'\boldsymbol{\Psi}_{11})^{-1}$ and for $i = 2, \ldots, k$,

$$\mathrm{Var}\begin{pmatrix}\boldsymbol{X}_1\\ \boldsymbol{X}_i\end{pmatrix} = \begin{pmatrix}\boldsymbol{\Psi}_{11} & \mathbf{0}\\ \boldsymbol{\Psi}_{i1} & \boldsymbol{\Psi}_{ii}\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{\Psi}_{11}' & \boldsymbol{\Psi}_{i1}'\\ \mathbf{0} & \boldsymbol{\Psi}_{ii}'\end{pmatrix}^{-1}.$$

Thus using $|\boldsymbol{B} + \boldsymbol{a}\boldsymbol{a}'| = |\boldsymbol{B}|(1 + \boldsymbol{a}'\boldsymbol{B}^{-1}\boldsymbol{a})$, where $\boldsymbol{B}$ is invertible and $\boldsymbol{a}$ is a vector, we can easily show that

$$|\boldsymbol{\Lambda}_{1j}| = 2\prod_{t=1}^{j}\frac{1}{\psi_{1tt}^2}, \qquad 1 \le j \le p_1; \tag{3.22}$$

$$|\boldsymbol{\Lambda}_{ij}| = 2\prod_{t=1}^{p_1}\frac{1}{\psi_{1tt}^2}\cdot\prod_{s=1}^{p_i}\frac{1}{\psi_{iss}^2}, \qquad 1 \le j \le p_i,\ 2 \le i \le k. \tag{3.23}$$

Hence the Jeffreys prior of $\boldsymbol{\Psi}$ (or $\boldsymbol{\theta}$) is $|\boldsymbol{\Lambda}(\theta)|^{1/2}$, which is proportional to that in (3.19). Based on (3.21), (3.22), and (3.23), the reference prior of $\boldsymbol{\Psi}$ for the ordered group $\{\psi_{111}, (\psi_{121}, \psi_{122}), \ldots, (\psi_{1p_11}, \ldots, \psi_{1p_1p_1}), (\phi_{211}, \ldots, \phi_{21p_1}, \psi_{211}), \ldots, (\phi_{k11}, \ldots, \phi_{kp_kp_1}, \psi_{kp_k1}, \ldots, \psi_{kp_kp_k})\}$ is easy to obtain as (3.20) according to the algorithm in Berger and Bernardo (1992).

4. Properties of posterior of $\boldsymbol{\Psi}$ under a class of priors

In this section, we consider a class of priors of $\boldsymbol{\Psi}$

$$p(\boldsymbol{\Psi}) \propto \prod_{i=1}^{k}\prod_{j=1}^{p_i}\psi_{ijj}^{\alpha_{ij}}\exp(-\beta_{ij}\psi_{ijj}^2), \tag{4.1}$$

where $\beta_{ij} \ge 0$, $j = 1, \ldots, p_i$, $i = 1, \ldots, k$. This class includes the left Haar invariant measure $\nu_{\mathcal{G}}^{l}(\boldsymbol{\Psi})$, the right Haar invariant measure $\nu_{\mathcal{G}}^{r}(\boldsymbol{\Psi})$ (the Jeffreys prior $\pi_J(\boldsymbol{\Psi})$), and the reference prior $\pi_R(\boldsymbol{\Psi})$. We have the following posterior properties:

THEOREM 4.1. *For the star-shape model, the posterior* $p(\boldsymbol{\Psi} \mid \boldsymbol{S})$ *under the prior* $p(\boldsymbol{\Psi})$ *in* (4.1) *has the following properties*:

(a) $p(\boldsymbol{\Psi} \mid \boldsymbol{S})$ *is proper if and only if* $n + \alpha_{ij} + 1 > 0$, $j = 1, \ldots, p_i$, $i = 1, \ldots, k$.

(b) $\boldsymbol{\Psi}_{11}, (\boldsymbol{\Psi}_{21}, \boldsymbol{\Psi}_{22}), \ldots, (\boldsymbol{\Psi}_{k1}, \boldsymbol{\Psi}_{kk})$ *are mutually independent*;

(c) *For* $i = 2, \ldots, k$, *conditional distribution of* $\boldsymbol{\Psi}_{i1}$ *given* $\boldsymbol{\Psi}_{ii}$ *is* $N_{p_i,p_1}(-\boldsymbol{\Psi}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1}, \boldsymbol{I}_{p_i} \otimes \boldsymbol{S}_{11}^{-1})$;

(d) *For* $i = 1, \ldots, k$,

$$\boldsymbol{\Psi}_{ii} \mid \boldsymbol{S} \sim \exp\left\{-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Psi}_{ii}\boldsymbol{S}_{ii\cdot 1}\boldsymbol{\Psi}_{ii}')\right\}\prod_{j=1}^{p_i}\psi_{ijj}^{n+\alpha_{ij}}\exp(-\beta_{ij}\psi_{ijj}^2).$$

PROOF. Because the likelihood function of $\boldsymbol{\Psi}$ is

$$f(\boldsymbol{S} \mid \boldsymbol{\Psi}) \propto |\boldsymbol{\Psi}'\boldsymbol{\Psi}|^{n/2}\exp\left\{-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Psi}'\boldsymbol{\Psi}\boldsymbol{S})\right\}, \tag{4.2}$$

then the posterior of $\boldsymbol{\Psi}$ under the prior $p(\boldsymbol{\Psi})$ in (4.1) is

$$p(\boldsymbol{\Psi} \mid \boldsymbol{S}) \propto |\boldsymbol{\Psi}'\boldsymbol{\Psi}|^{n/2} \exp\left\{-\frac{1}{2}\operatorname{tr}(\boldsymbol{\Psi}'\boldsymbol{\Psi}\boldsymbol{S})\right\} \prod_{i=1}^{k}\prod_{j=1}^{p_i} \psi_{ijj}^{\alpha_{ij}} \exp(-\beta_{ij}\psi_{ijj}^2).$$

Because $|\boldsymbol{\Psi}'\boldsymbol{\Psi}|^{n/2} = \prod_{i=1}^{k} |\boldsymbol{\Psi}_{ii}|^n$ and

$$\begin{aligned}(4.3) \qquad \operatorname{tr}(\boldsymbol{\Psi}\boldsymbol{S}\boldsymbol{\Psi}') &= \sum_{i=1}^{k} \operatorname{tr}(\boldsymbol{\Psi}_{ii}\boldsymbol{S}_{ii\cdot 1}\boldsymbol{\Psi}_{ii}') \\ &\quad + \sum_{i=2}^{k} \operatorname{tr}\{(\boldsymbol{\Psi}_{i1} + \boldsymbol{\Psi}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1})\boldsymbol{S}_{11}(\boldsymbol{\Psi}_{i1} + \boldsymbol{\Psi}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1})'\}.\end{aligned}$$

Hence it follows

$$\begin{aligned} p(\boldsymbol{\Psi} \mid \boldsymbol{S}) \propto & \prod_{i=2}^{k} \exp\left[-\frac{1}{2}\operatorname{tr}\{(\boldsymbol{\Psi}_{i1} + \boldsymbol{\Psi}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1})\boldsymbol{S}_{11}(\boldsymbol{\Psi}_{i1} + \boldsymbol{\Psi}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1})'\}\right] \\ & \times \prod_{i=1}^{k} \exp\left\{-\frac{1}{2}\operatorname{tr}(\boldsymbol{\Psi}_{ii}\boldsymbol{S}_{ii\cdot 1}\boldsymbol{\Psi}_{ii}')\right\} \prod_{j=1}^{p_i} \psi_{ijj}^{n+\alpha_{ij}} \exp(-\beta_{ij}\psi_{ijj}^2).\end{aligned}$$

Thus we proves parts (b), (c), and (d). For part (a), it is easy to show that $p(\boldsymbol{\Psi} \mid \boldsymbol{S})$ is proper if and only if the marginal posterior $p(\boldsymbol{\Psi}_{ii} \mid \boldsymbol{S})$ is proper, $i = 1, \ldots, k$. By taking the transformation $\boldsymbol{\Psi}_{ii} \to \boldsymbol{\Theta}_{ii} = \boldsymbol{\Psi}_{ii}\boldsymbol{T}_{ii}$, we will readily get that $p(\boldsymbol{\Psi}_{ii} \mid \boldsymbol{S})$ is proper if and only if $n + \alpha_{ij} + 1 > 0$, $j = 1, \ldots, p_i$, $i = 1, \ldots, k$. Hence (a) holds.

From Theorem 4.1, each of the posteriors under the left Haar invariant measure $\nu_{\mathcal{G}}^{l}(\boldsymbol{\Psi})$, the right Haar invariant measure $\nu_{\mathcal{G}}^{r}(\boldsymbol{\Psi})$ (the Jeffreys prior $\pi_J(\boldsymbol{\Psi})$), and the reference prior $\pi_R(\boldsymbol{\Psi})$ will be proper. Specifically, the posterior under the left Haar invariant measure $\nu_{\mathcal{G}}^{l}(\boldsymbol{\Psi})$ is related to Wishart distribution as shown below.

COROLLARY 4.1. *If we take the left invariant Haar measure of the group $\mathcal{G}$, $\nu_{\mathcal{G}}^{l}(d\boldsymbol{\Psi})$ as a prior, then the posterior distribution of $\boldsymbol{\Psi}$ has the following properties:*

(a) $\boldsymbol{\Psi}_{11}, (\boldsymbol{\Psi}_{21}, \boldsymbol{\Psi}_{22}), \ldots, (\boldsymbol{\Psi}_{k1}, \boldsymbol{\Psi}_{kk})$ *are mutually independent*;

(b) $\boldsymbol{\Psi}_{11}\boldsymbol{S}_{11}\boldsymbol{\Psi}_{11}' \sim W_{p_1}(n, \boldsymbol{I}_{p_1})$;

(c) *For* $i = 2, \ldots, k$, $\boldsymbol{\Psi}_{ii}\boldsymbol{S}_{ii\cdot 1}\boldsymbol{\Psi}_{ii}' \sim W_{p_i}(n - p_1, \boldsymbol{I}_{p_i})$;

(d) *For* $i = 2, \ldots, k$, *conditional distribution of* $\boldsymbol{\Psi}_{i1}$ *given* $\boldsymbol{\Psi}_{ii}$ *is* $N_{p_i,p_1}(-\boldsymbol{\Psi}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1}, \boldsymbol{I}_{p_i} \otimes \boldsymbol{S}_{11}^{-1})$.

5. Bayesian estimators of $\boldsymbol{\Omega}$ under the entropy loss

To find the Bayesian estimate of $\boldsymbol{\Omega}$ with respect to the prior $p(\boldsymbol{\Psi})$ under the entropy loss L_1, we need the following two lemmas.

LEMMA 5.1. *Let $\boldsymbol{A}$ be a constant positive definite matrix and $\boldsymbol{B} = (b_{ij})_{m\times m}$ be its Cholesky decomposition. Assume that $\boldsymbol{Z} = (z_{ij})_{m\times m}$ is a randomly lower-triangular with positive diagonal elements whose distribution follows*

$$(5.1) \qquad \boldsymbol{Z} \sim \exp\left\{-\frac{1}{2}\operatorname{tr}(\boldsymbol{Z}\boldsymbol{A}\boldsymbol{Z}')\right\} \prod_{i=1}^{m} z_{ii}^{\alpha_i} \exp(-\beta_i z_{ii}^2).$$

(a) *If* $\alpha_i > 0$, $\beta_i \geq 0$, $i = 1, \ldots, m$, *then*

$$E(\boldsymbol{Z}'\boldsymbol{Z}) = (\boldsymbol{B}')^{-1} \operatorname{diag}(\delta_1, \ldots, \delta_m)\boldsymbol{B}^{-1}, \tag{5.2}$$

where $\delta_i = (\alpha_i + 1)/(1 + 2\beta_i b_{ii}^{-2}) + m - i$, $i = 1, \ldots, m$.

(b) *If* $\alpha_i > 1$, $\beta_i \geq 0$, $i = 1, \ldots, m$, *then*

$$E(\boldsymbol{Z}'\boldsymbol{Z})^{-1} = \boldsymbol{B} \operatorname{diag}(\eta_1, \ldots, \eta_m)\boldsymbol{B}', \tag{5.3}$$

where $\eta_1 = u_1$, $\eta_j = u_j \prod_{i=1}^{j-1}(1 + u_i)$, $j = 2, \ldots, m$ *with* $u_i = (1 + 2\beta_i b_{ii}^{-2})/(\alpha_i - 1)$, $i = 1, \ldots, m$.

PROOF. Letting $\boldsymbol{Y} = \boldsymbol{Z}\boldsymbol{B}$, then $\boldsymbol{Y} = (y_{ij})$ is still lower-triangular and

$$\boldsymbol{Y} \sim \exp\left\{-\frac{1}{2}\operatorname{tr}(\boldsymbol{Y}\boldsymbol{Y}')\right\} \prod_{i=1}^{m} y_{ii}^{\alpha_i} \exp(-\beta_i b_{ii}^{-2} y_{ii}^2). \tag{5.4}$$

From above, we know that all y_{ij}, $1 \leq j \leq i \leq m$ are independent and

$$y_{ij} \sim N(0,1), \qquad 1 \leq j < i \leq m;$$
$$y_{ii} \sim y_{ii}^{\alpha_i} \exp\left(-\frac{1 + 2\beta_i b_{ii}^{-2}}{2} y_{ii}^2\right), \qquad 1 \leq i \leq m.$$

If $\alpha_i > 0$, $\beta_i \geq 0$, $i = 1, \ldots, m$, then $y_{ii}^2 \sim \Gamma((\alpha_i + 1)/2, (1 + 2\beta_i b_{ii}^{-2})/2)$ and $E(y_{ii}^2)$ exists, $i = 1, \ldots, m$. Thus it is straightforward to get (5.2). For (5.3), we just need to show $E(\boldsymbol{Y}'\boldsymbol{Y})^{-1} = \operatorname{diag}(\eta_1, \ldots, \eta_m)$. Under the condition $\alpha_i > 1$, $\beta_i \geq 0$, $E(y_{ii}^{-2})$ exists and is equal to u_i, $i = 1, \ldots, m$. Thus we can get the result by using the same procedure in the Appendix on p. 1648 of Eaton and Olkin (1987).

For $A = \boldsymbol{I}$, $\alpha_i = n$, $\beta_i = 0$, $i = 1, \ldots, m$, $\boldsymbol{Z}\boldsymbol{Z}'$ follows Wishart distribution with parameters n and $\boldsymbol{I}_m$, and in this special case (5.3) was first obtained by Eaton and Olkin (1987). We also note that δ_i, η_i in Lemma 5.1 are independent of $\boldsymbol{A}$ if and only if $\beta_i = 0$, $i = 1, \ldots, m$.

LEMMA 5.2. *Let* $\mathcal{A} = \{\boldsymbol{B} \in \boldsymbol{R}^{p\times p} \mid \boldsymbol{B}$ *is lower-triangular with positive diagonal elements*$\}$. *If* $\boldsymbol{\Delta} > 0$, *then*

$$\min_{B\in\mathcal{A}}\{\operatorname{tr}(\boldsymbol{B}\boldsymbol{\Delta}\boldsymbol{B}') - \log|\boldsymbol{B}\boldsymbol{B}'|\} = p + \log|\boldsymbol{\Delta}|$$

is achieved at $\boldsymbol{B} = \boldsymbol{\Xi}^{-1}$, *where* $\boldsymbol{\Xi}$ *is Cholesky decomposition of* $\boldsymbol{\Delta}$.

Note that any positive definite matrix has a unique Cholesky decomposition, and the proof of the above lemma is directly obtained by applying Lemma 2.1 in Eaton and Olkin (1987).

THEOREM 5.1. *Suppose that* $n + \alpha_{ij} - 1 > 0$, $i = 1, \ldots, k$, $j = 1, \ldots, p_i$. *Then under the entropy loss* L_1, *the Bayesian estimator of* $\boldsymbol{\Omega}$ *with respect to the prior* $p(\boldsymbol{\Psi})$ *in* (4.1) *is given by*

$$\hat{\boldsymbol{\Omega}}_1 = \boldsymbol{R}'\boldsymbol{B}^{-1}\boldsymbol{R}, \tag{5.5}$$

where $\boldsymbol{R}$ *is given by* (3.10), $\boldsymbol{B} = \mathrm{diag}(\boldsymbol{B}_1, \boldsymbol{B}_2, \ldots, \boldsymbol{B}_k)$ *and* $\boldsymbol{B}_i = \mathrm{diag}(b_{i1}, \ldots, b_{ip_i})$ *with*

$$b_{11} = u_{11}, \qquad b_{1j} = u_{1j} \prod_{t=1}^{j-1}(1 + u_{1t}), \qquad j = 2, \ldots, p_1 \tag{5.6}$$

and

$$b_{i1} = \{1 + \mathrm{tr}(\boldsymbol{B}_1)\}u_{i1}, \tag{5.7}$$

$$b_{ij} = \{1 + \mathrm{tr}(\boldsymbol{B}_1)\}u_{ij} \prod_{t=1}^{j-1}(1 + u_{it}), \qquad j = 2, \ldots, p_i, \;\; i = 2, \ldots, k. \tag{5.8}$$

Here $u_{ij} = (1 + 2\beta_{ij}t_{ijj}^{-2})/(n + \alpha_{ij} - 1)$, $j = 1, \ldots, p_i$, $i = 1, \ldots, k$ *with* t_{ijj} *being the* j*-th diagonal element of* $\boldsymbol{T}_{ii}$.

PROOF. The Bayesian estimator of $\boldsymbol{\Omega}$ under the entropy loss L_1 will be produced by minimizing the posterior risk

$$b_1(\hat{\boldsymbol{\Omega}}) = \int [\mathrm{tr}\{\hat{\boldsymbol{\Omega}}(\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\} - \log|\hat{\boldsymbol{\Omega}}(\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}| - p]p(\boldsymbol{\Psi} \mid \boldsymbol{S})d\boldsymbol{\Psi},$$

where $p(\boldsymbol{\Psi} \mid \boldsymbol{S})$ is described in Theorem 4.1. Let $\hat{\boldsymbol{\Omega}} = \hat{\boldsymbol{\Psi}}'\hat{\boldsymbol{\Psi}}$, where $\hat{\boldsymbol{\Psi}} \in \mathcal{G}$ and has the similar block partition as in (3.2). The question is then how to minimize

$$g_1(\hat{\boldsymbol{\Psi}}) = \int \mathrm{tr}\{(\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1})(\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1})'\}p(\boldsymbol{\Psi} \mid \boldsymbol{S})d\boldsymbol{\Psi} - \log|\hat{\boldsymbol{\Psi}}'\hat{\boldsymbol{\Psi}}|.$$

So we need to calculate the posterior expectation of $\mathrm{tr}\{(\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1})(\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1})'\}$. By (3.4), it follows

$$\begin{aligned} &\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1} \\ &= \begin{pmatrix} \hat{\boldsymbol{\Psi}}_{11}\boldsymbol{\Psi}_{11}^{-1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ (\hat{\boldsymbol{\Psi}}_{21} - \hat{\boldsymbol{\Psi}}_{22}\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{\Psi}_{21})\boldsymbol{\Psi}_{11}^{-1} & \hat{\boldsymbol{\Psi}}_{22}\boldsymbol{\Psi}_{22}^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ (\hat{\boldsymbol{\Psi}}_{31} - \hat{\boldsymbol{\Psi}}_{33}\boldsymbol{\Psi}_{33}^{-1}\boldsymbol{\Psi}_{31})\boldsymbol{\Psi}_{11}^{-1} & \mathbf{0} & \hat{\boldsymbol{\Psi}}_{33}\boldsymbol{\Psi}_{33}^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (\hat{\boldsymbol{\Psi}}_{k1} - \hat{\boldsymbol{\Psi}}_{kk}\boldsymbol{\Psi}_{kk}^{-1}\boldsymbol{\Psi}_{k1})\boldsymbol{\Psi}_{11}^{-1} & \mathbf{0} & \mathbf{0} & \cdots & \hat{\boldsymbol{\Psi}}_{kk}\boldsymbol{\Psi}_{kk}^{-1} \end{pmatrix}, \end{aligned} \tag{5.9}$$

and therefore

$$\begin{aligned} &\mathrm{tr}\{(\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1})(\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1})'\} \\ &\quad = \sum_{i=1}^{k} \mathrm{tr}\{\hat{\boldsymbol{\Psi}}_{ii}(\boldsymbol{\Psi}_{ii}'\boldsymbol{\Psi}_{ii})^{-1}\hat{\boldsymbol{\Psi}}_{ii}'\} \\ &\quad\quad + \sum_{i=2}^{k} \mathrm{tr}\{(\hat{\boldsymbol{\Psi}}_{i1} - \hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{\Psi}_{ii}^{-1}\boldsymbol{\Psi}_{i1})(\boldsymbol{\Psi}_{11}'\boldsymbol{\Psi}_{11})^{-1}(\hat{\boldsymbol{\Psi}}_{i1} - \hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{\Psi}_{ii}^{-1}\boldsymbol{\Psi}_{i1})'\}. \end{aligned} \tag{5.10}$$

From Theorem 4.1(d) and (5.3) in Lemma 5.1, it follows that

$$\begin{aligned} E\{(\boldsymbol{\Psi}_{11}'\boldsymbol{\Psi}_{11})^{-1} \mid \boldsymbol{S}\} &= \boldsymbol{T}_{11}\boldsymbol{B}_1\boldsymbol{T}_{11}', \\ E\{(\boldsymbol{\Psi}_{ii}'\boldsymbol{\Psi}_{ii})^{-1} \mid \boldsymbol{S}\} &= \boldsymbol{T}_{ii}\boldsymbol{B}_i\boldsymbol{T}_{ii}'/\{1 + \mathrm{tr}(\boldsymbol{B}_1)\}, \qquad i = 2, \ldots, k \end{aligned}$$

because of the condition $n+\alpha_{ij}-1>0$, $i=1,\ldots,k$, $j=1,\ldots,p_i$. In addition, by Theorem 4.1(a)(d) and Theorem 2.3.5 of Gupta and Nagar (2000),

$$\begin{aligned}
&E[\mathrm{tr}\{(\hat{\boldsymbol{\Psi}}_{i1}-\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{\Psi}_{ii}^{-1}\boldsymbol{\Psi}_{i1})(\boldsymbol{\Psi}_{11}'\boldsymbol{\Psi}_{11})^{-1}(\hat{\boldsymbol{\Psi}}_{i1}-\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{\Psi}_{ii}^{-1}\boldsymbol{\Psi}_{i1})'\}\mid \boldsymbol{S}]\\
&\quad=\mathrm{tr}[E\{(\hat{\boldsymbol{\Psi}}_{i1}-\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{\Psi}_{ii}^{-1}\boldsymbol{\Psi}_{i1})'(\hat{\boldsymbol{\Psi}}_{i1}-\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{\Psi}_{ii}^{-1}\boldsymbol{\Psi}_{i1})\mid \boldsymbol{S}\}E\{(\boldsymbol{\Psi}_{11}'\boldsymbol{\Psi}_{11})^{-1}\mid \boldsymbol{S}\}]\\
&\quad=\mathrm{tr}\{(\hat{\boldsymbol{\Psi}}_{i1}+\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1})'(\hat{\boldsymbol{\Psi}}_{i1}+\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1})\boldsymbol{T}_{11}\boldsymbol{B}_1\boldsymbol{T}_{11}'\}\\
&\qquad+\mathrm{tr}(\boldsymbol{B}_1)\,\mathrm{tr}(\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{T}_{ii}\boldsymbol{B}_i\boldsymbol{T}_{ii}'\hat{\boldsymbol{\Psi}}_{ii}{}')/\{1+\mathrm{tr}(\boldsymbol{B}_1)\}.
\end{aligned}$$

Thus we have

$$\begin{aligned}
(5.11)\qquad &E[\mathrm{tr}\{(\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1})(\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1})'\}\mid \boldsymbol{S}]\\
&\quad=\sum_{i=1}^{k}\mathrm{tr}(\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{T}_{ii}\boldsymbol{B}_i\boldsymbol{T}_{ii}'\hat{\boldsymbol{\Psi}}_{ii}')\\
&\qquad+\sum_{i=2}^{k}\mathrm{tr}\{(\hat{\boldsymbol{\Psi}}_{i1}+\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1})'(\hat{\boldsymbol{\Psi}}_{i1}+\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1})\boldsymbol{T}_{11}\boldsymbol{B}_1\boldsymbol{T}_{11}'\}.
\end{aligned}$$

Hence,

$$\begin{aligned}
g_1(\hat{\boldsymbol{\Psi}})&=\sum_{i=1}^{k}\{\mathrm{tr}(\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{T}_{ii}\boldsymbol{B}_i\boldsymbol{T}_{ii}'\hat{\boldsymbol{\Psi}}_{ii}')-\log|\hat{\boldsymbol{\Psi}}_{ii}'\hat{\boldsymbol{\Psi}}_{ii}|\}\\
&\quad+\sum_{i=2}^{k}\mathrm{tr}\{(\hat{\boldsymbol{\Psi}}_{i1}+\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1})'(\hat{\boldsymbol{\Psi}}_{i1}+\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1})\boldsymbol{T}_{11}\boldsymbol{B}_1\boldsymbol{T}_{11}'\}.
\end{aligned}$$

Because $\boldsymbol{T}_{11}\boldsymbol{B}_1\boldsymbol{T}_{11}'>0$, then by Lemma 5.2, we can readily see that $g_1(\hat{\boldsymbol{\Psi}})$ is minimized at $\hat{\boldsymbol{\Psi}}_{ii}=\boldsymbol{B}_i^{-1/2}\boldsymbol{T}_{ii}^{-1}$ for $i=1,2,\ldots,k$ and $\hat{\boldsymbol{\Psi}}_{j1}=-\hat{\boldsymbol{\Psi}}_{jj}\boldsymbol{S}_{j1}\boldsymbol{S}_{11}^{-1}$ for $j=2,3,\ldots,k$. Thus the proof is completed.

From Theorem 5.1, the Bayesian estimator $\hat{\boldsymbol{\Omega}}_1$ is equivariant with respect to the group $\mathcal{G}$ if and only if all $\beta_{ij}=0$, $j=1,\ldots,p_i$, $i=1,\ldots,k$. In this case, $\hat{\boldsymbol{\Omega}}_1$ will have the form (3.15), which includes the MLE $\hat{\boldsymbol{\Omega}}_M$ and the unbiased estimator $\hat{\boldsymbol{\Omega}}_U$. Conversely, we can show that any estimator having the form (3.15) will be the Bayesian estimator of $\boldsymbol{\Omega}$ with respect to the prior $p(\boldsymbol{\Psi})$ by taking $\beta_{ij}=0$ and some appropriate a_is. A similar result will hold for the symmetric loss L_2 discussed in the following section.

As a corollary of Theorem 5.1, we get the Bayesian estimator of $\boldsymbol{\Omega}$ under the entropy loss L_1 with respect to the left Haar invariant measure $\nu_{\mathcal{G}}^l(\boldsymbol{\Psi})$. This estimator will be shown the best equivariant estimator under the group $\mathcal{G}$. Bayesian estimators of $\boldsymbol{\Omega}$ with respect to the right Haar invariant measure $\nu_{\mathcal{G}}^r(\boldsymbol{\Psi})$ (the Jeffreys prior $\pi_J(\boldsymbol{\Psi})$) and the reference prior $\pi_R(\boldsymbol{\Psi})$ will be given in Corollary 5.2.

COROLLARY 5.1. *Under the entropy loss L_1, the best $\mathcal{G}$-equivariant estimator of $\boldsymbol{\Omega}$ is the same as the Bayesian estimator with respect to the left Haar invariant measure $\nu_{\mathcal{G}}^l(\boldsymbol{\Psi})$ and is given by*

$$\hat{\boldsymbol{\Omega}}_{1B}=\boldsymbol{R}'\boldsymbol{B}_B^{-1}\boldsymbol{R}, \tag{5.12}$$

where $\boldsymbol{R}$ *is given by* (3.10), $\boldsymbol{B}_B = \mathrm{diag}(\boldsymbol{B}_{1B}, \boldsymbol{B}_{2B}, \dots, \boldsymbol{B}_{kB})$ *and* $\boldsymbol{B}_{iB} = \mathrm{diag}(b_{i1B}, \dots, b_{ip_iB})$ *with*

$$b_{ijB} = \begin{cases} \frac{n-1}{(n-j-1)(n-j)}, & \text{if } i = 1,\ j = 1, \dots, p_1, \\ \frac{n-1}{(n-p_1-j-1)(n-p_1-j)}, & \text{if } i = 2, \dots, k,\ j = 1, \dots, p_i. \end{cases} \tag{5.13}$$

PROOF. Suppose $\boldsymbol{\Sigma} = \boldsymbol{\Delta}\boldsymbol{\Delta}'$, where $\boldsymbol{\Delta} \in \mathcal{G}$. By Theorem 6.5 in Eaton (1989), the best equivariant estimator of $\boldsymbol{\Sigma}$ with respect to the group $\mathcal{G}$ will be the Bayesian estimator if we take a right invariant Haar measure $\nu_{\mathcal{G}}^r(d\boldsymbol{\Delta})$ on the group $\mathcal{G}$ as a prior. Because $\boldsymbol{\Omega} = \boldsymbol{\Psi}'\boldsymbol{\Psi} = (\boldsymbol{\Delta}')^{-1}\boldsymbol{\Delta}^{-1}$ and $\nu_{\mathcal{G}}^r(d\boldsymbol{\Delta}) = \nu_{\mathcal{G}}^l(d\boldsymbol{\Psi})$, thus the best equivariant estimator of $\boldsymbol{\Omega}$ with respect to the group $\mathcal{G}$ will be the Bayesian estimator if we take the left invariant Haar measure $\nu_{\mathcal{G}}^l(d\boldsymbol{\Psi})$ on the group $\mathcal{G}$ as a prior. So this completes the proof by taking $\alpha_{1j} = -j$ if $1 \le j \le p_1$, $\alpha_{ij} = -p_1 - j$ if $1 \le j \le p_i$, $2 \le i \le k$ and $\beta_{ij} = 0$, $1 \le j \le p_i$, $1 \le i \le k$ in Theorem 5.1.

Remark 4. It is well-known that the group of lower-triangular matrices is solvable and thus its subgroup $\mathcal{G}$ is also solvable (see Bondar and Milnes (1981) for a survey). By Kiefer (1957), the best $\mathcal{G}$-equivariant estimator $\hat{\boldsymbol{\Omega}}_{1B}$ is also minimax with respect to the entropy loss L_1.

COROLLARY 5.2. *Under the entropy loss* L_1, *the Bayesian estimator* $\hat{\boldsymbol{\Omega}}_{1J}$ *of* $\boldsymbol{\Omega}$ *with respect to the Jeffreys prior* $\pi_J(\boldsymbol{\Psi})$ *is*

$$\hat{\boldsymbol{\Omega}}_{1J} = \boldsymbol{R}'\boldsymbol{B}_J^{-1}\boldsymbol{R}, \tag{5.14}$$

where $\boldsymbol{B}_J$ *has the form*

$$\mathrm{diag}\left(\frac{1}{n-p-1}\boldsymbol{I}_{p_1}, \frac{n-p+p_1-1}{(n-p-1)(n-p_2-1)}\boldsymbol{I}_{p_2}, \dots, \frac{n-p+p_1-1}{(n-p-1)(n-p_k-1)}\boldsymbol{I}_{p_k}\right).$$

The Bayesian estimator $\hat{\boldsymbol{\Omega}}_{1R}$ *with respect to the reference prior* $\pi_R(\boldsymbol{\Psi})$ *under the entropy loss* L_1 *is*

$$\hat{\boldsymbol{\Omega}}_{1R} = \boldsymbol{R}'\boldsymbol{B}_R^{-1}\boldsymbol{R}, \tag{5.15}$$

where $\boldsymbol{B}_R = \mathrm{diag}(\boldsymbol{B}_{1R}, \boldsymbol{B}_{2R}, \dots, \boldsymbol{B}_{kR})$ *and* $\boldsymbol{B}_{iR} = \mathrm{diag}(b_{i1R}, b_{i2R}, \dots, b_{ip_iR})$ *with*

$$b_{1jR} = \frac{(n-1)^{j-1}}{(n-2)^j}, \quad j = 1, 2, \dots, p_1; \tag{5.16}$$

$$b_{ijR} = \{1 + \mathrm{tr}(\boldsymbol{B}_{1R})\}\frac{(n-1)^{j-1}}{(n-2)^j}, \quad j = 1, 2, \dots, p_i,\ i = 2, \dots, k. \tag{5.17}$$

6. Bayesian estimators of $\boldsymbol{\Omega}$ under the symmetric loss

Similarly to Lemma 5.2, we need the following lemma, which is a direct corollary of Lemma 2.2 in Eaton and Olkin (1987).

LEMMA 6.1. *Suppose that $\mathcal{A}$ is defined in Lemma 5.2. If $\boldsymbol{\Delta}$ and $\boldsymbol{\Lambda}$ are both diagonal with known positive diagonal elements, then*

$$\min_{B\in\mathcal{A}}[\operatorname{tr}(\boldsymbol{B}\boldsymbol{\Delta}\boldsymbol{B}') + \operatorname{tr}\{(\boldsymbol{B}')^{-1}\boldsymbol{\Lambda}\boldsymbol{B}^{-1}\}] = 2\operatorname{tr}(\boldsymbol{\Delta}^{1/2}\boldsymbol{\Lambda}^{1/2})$$

is achieved at $\boldsymbol{B} = (\boldsymbol{\Delta}^{-1/2}\boldsymbol{\Lambda}^{1/2})^{1/2} = \boldsymbol{\Delta}^{-1/4}\boldsymbol{\Lambda}^{1/4}$.

THEOREM 6.1. *Suppose that $n+\alpha_{ij}-1>0$, $i=1,\ldots,k$, $j=1,\ldots,p_i$. Under the symmetric loss L_2, the Bayesian estimator of $\boldsymbol{\Omega}$ with respect to the prior $p(\boldsymbol{\Psi})$ in (4.1) is given by*

$$\hat{\boldsymbol{\Omega}}_2 = \boldsymbol{R}'\boldsymbol{H}^{-1}\boldsymbol{R}, \tag{6.1}$$

where $\boldsymbol{R}$ is given by (3.10), $\boldsymbol{H} = \boldsymbol{B}^{1/2}\boldsymbol{C}^{-1/2}$ with $\boldsymbol{B}$ being defined in Theorem 5.1 and $\boldsymbol{C} = \operatorname{diag}(\boldsymbol{C}_1,\ldots,\boldsymbol{C}_k)$. Here $\boldsymbol{C}_i = \operatorname{diag}(c_{i1},\ldots,c_{ip_i})$,

$$\begin{aligned}
c_{1j} &= \frac{n+\alpha_{1j}+1}{1+2\beta_{1j}t_{1jj}^{-2}} + p - j, \qquad j=1,\ldots,p_1;\\
c_{ij} &= \frac{n+\alpha_{ij}+1}{1+2\beta_{ij}t_{ijj}^{-2}} + p_i - j, \qquad j=1,\ldots,p_i,\ \ i=2,\ldots,k.
\end{aligned}$$

and t_{ijj} is the j-th diagonal element of $\boldsymbol{T}_{ii}$.

PROOF. Under the symmetric loss L_2, the Bayesian estimator of $\boldsymbol{\Omega}$ with respect to the prior $p(\boldsymbol{\Psi})$ will be produced by minimizing the posterior risk

$$b_2(\hat{\boldsymbol{\Omega}}) = \int[\operatorname{tr}\{\hat{\boldsymbol{\Omega}}(\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\} + \operatorname{tr}\{\hat{\boldsymbol{\Omega}}^{-1}(\boldsymbol{\Psi}'\boldsymbol{\Psi})\} - 2p]p(\boldsymbol{\Psi}\mid\boldsymbol{S})d\boldsymbol{\Psi}.$$

Similarly to the proof of Theorem 5.1, by setting $\hat{\boldsymbol{\Omega}} = \hat{\boldsymbol{\Psi}}'\hat{\boldsymbol{\Psi}}$, we just need to minimize

$$g_2(\hat{\boldsymbol{\Psi}}) = \int[\operatorname{tr}\{(\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1})(\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1})'\} + \operatorname{tr}\{(\boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{-1})(\boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{-1})'\}]p(\boldsymbol{\Psi}\mid\boldsymbol{S})d\boldsymbol{\Psi}$$

in terms of $\hat{\boldsymbol{\Psi}}$. With the condition $n+\alpha_{ij}-1>0$, $i=1,\ldots,k$, $j=1,\ldots,p_i$, the posterior expectation of $\operatorname{tr}\{(\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1})(\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}^{-1})'\}$ is shown by (5.11). We now calculate the posterior expectation of $\operatorname{tr}\{(\boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{-1})(\boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{-1})'\}$. Similar to (5.10), it follows

$$\begin{aligned}
&\operatorname{tr}\{(\boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{-1})(\boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{-1})'\}\\
&\quad= \sum_{i=1}^{k}\operatorname{tr}(\hat{\boldsymbol{\Psi}}_{ii}^{-1\prime}\boldsymbol{\Psi}_{ii}'\boldsymbol{\Psi}_{ii}\hat{\boldsymbol{\Psi}}_{ii}^{-1})\\
&\quad\quad+\sum_{i=2}^{k}\operatorname{tr}\{(\boldsymbol{\Psi}_{i1}-\boldsymbol{\Psi}_{ii}\hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})\hat{\boldsymbol{\Psi}}_{11}^{-1}\hat{\boldsymbol{\Psi}}_{11}^{-1\prime}(\boldsymbol{\Psi}_{i1}-\boldsymbol{\Psi}_{ii}\hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})'\}.
\end{aligned} \tag{6.2}$$

By (5.2) in Lemma 5.1, we have

$$E(\boldsymbol{\Psi}_{ii}'\boldsymbol{\Psi}_{ii}\mid\boldsymbol{S}) = (\boldsymbol{T}_{ii}')^{-1}\boldsymbol{D}_i\boldsymbol{T}_{ii}^{-1}, \qquad i=1,\ldots,k,$$

where $\boldsymbol{D}_i$ is $p_i \times p_i$ diagonal with the j-th diagonal element

$$d_{ij} = \frac{n + \alpha_{ij} + 1}{1 + 2\beta_{ij} t_{ijj}^{-1}} + p_i - j, \qquad j = 1, 2, \ldots, p_i, \; i = 1, \ldots, k.$$

Moreover, by Theorem 4.1(d) and applying Theorem 2.3.5 in Gupta and Nagar (2000), we have

$$\begin{aligned} &E\{(\boldsymbol{\Psi}_{i1} - \boldsymbol{\Psi}_{ii}\hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})(\hat{\boldsymbol{\Psi}}_{11}'\hat{\boldsymbol{\Psi}}_{11})^{-1}(\boldsymbol{\Psi}_{i1} - \boldsymbol{\Psi}_{ii}\hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})' \mid \boldsymbol{S}\} \\ &\quad = \operatorname{tr}\{(\hat{\boldsymbol{\Psi}}_{11}'\hat{\boldsymbol{\Psi}}_{11})^{-1}\boldsymbol{S}_{11}^{-1}\}\boldsymbol{I}_{p_i} \\ &\qquad + E\{\boldsymbol{\Psi}_{ii}(\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1} + \hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})(\hat{\boldsymbol{\Psi}}_{11}'\hat{\boldsymbol{\Psi}}_{11})^{-1}(\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1} + \hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})'\boldsymbol{\Psi}_{ii}' \mid \boldsymbol{S}\}. \end{aligned}$$

Thus,

$$\begin{aligned} (6.3) \quad & E[\operatorname{tr}\{(\boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{-1})(\boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{-1})'\} \mid \boldsymbol{S}] \\ &= \sum_{i=1}^{k} \operatorname{tr}\{(\hat{\boldsymbol{\Psi}}_{ii}')^{-1}(\boldsymbol{T}_{ii}')^{-1}\boldsymbol{D}_i\boldsymbol{T}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{ii}^{-1}\} + \sum_{i=2}^{k} \operatorname{tr}\{(\hat{\boldsymbol{\Psi}}_{11}'\hat{\boldsymbol{\Psi}}_{11})^{-1}\boldsymbol{S}_{11}^{-1}\}\operatorname{tr}(\boldsymbol{I}_{p_i}) \\ &\quad + \sum_{i=2}^{k} \operatorname{tr}\{(\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1} + \hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})\hat{\boldsymbol{\Psi}}_{11}^{-1}\hat{\boldsymbol{\Psi}}_{11}^{-1\prime} \\ &\qquad\qquad \times (\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1} + \hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})'(\boldsymbol{T}_{ii}')^{-1}\boldsymbol{D}_i\boldsymbol{T}_{ii}^{-1}\} \\ &= \sum_{i=1}^{k} \operatorname{tr}\{(\hat{\boldsymbol{\Psi}}_{ii}')^{-1}(\boldsymbol{T}_{ii}')^{-1}\boldsymbol{C}_i\boldsymbol{T}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{ii}^{-1}\} \\ &\quad + \sum_{i=2}^{k} \operatorname{tr}\{(\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1} + \hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})\hat{\boldsymbol{\Psi}}_{11}^{-1}\hat{\boldsymbol{\Psi}}_{11}^{-1\prime} \\ &\qquad\qquad \times (\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1} + \hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})'(\boldsymbol{T}_{ii}')^{-1}\boldsymbol{D}_i\boldsymbol{T}_{ii}^{-1}\}. \end{aligned}$$

Combining (5.11) and (6.3), we get

$$\begin{aligned} (6.4) \quad g_2(\hat{\boldsymbol{\Psi}}) &= \sum_{i=1}^{k} \operatorname{tr}(\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{T}_{ii}\boldsymbol{B}_i\boldsymbol{T}_{ii}'\hat{\boldsymbol{\Psi}}_{ii}') + \sum_{i=1}^{k} \operatorname{tr}\{(\hat{\boldsymbol{\Psi}}_{ii}')^{-1}(\boldsymbol{T}_{ii}')^{-1}\boldsymbol{C}_i\boldsymbol{T}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{ii}^{-1}\} \\ &\quad + \frac{1}{1 + \operatorname{tr}(\boldsymbol{B}_1)} \sum_{i=2}^{k} \operatorname{tr}\{(\hat{\boldsymbol{\Psi}}_{i1} + \hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1})' \\ &\qquad\qquad \times (\hat{\boldsymbol{\Psi}}_{i1} + \hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1})\boldsymbol{T}_{ii}\boldsymbol{B}_i\boldsymbol{T}_{ii}'\} \\ &\quad + \sum_{i=2}^{k} \operatorname{tr}\{(\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1} + \hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})\hat{\boldsymbol{\Psi}}_{11}^{-1}\hat{\boldsymbol{\Psi}}_{11}^{-1\prime} \\ &\qquad\qquad \times (\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1} + \hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})'\boldsymbol{T}_{ii}^{\prime -1}\boldsymbol{D}_i\boldsymbol{T}_{ii}^{-1}\} \\ &\geq \sum_{i=1}^{k} \operatorname{tr}(\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{T}_{ii}\boldsymbol{B}_i\boldsymbol{T}_{ii}'\hat{\boldsymbol{\Psi}}_{ii}') + \sum_{i=1}^{k} \operatorname{tr}\{(\hat{\boldsymbol{\Psi}}_{ii}')^{-1}(\boldsymbol{T}_{ii}')^{-1}\boldsymbol{C}_i\boldsymbol{T}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{ii}^{-1}\}, \end{aligned}$$

and the equality holds if we take

$$\hat{\boldsymbol{\Psi}}_{i1} = -\hat{\boldsymbol{\Psi}}_{ii}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1}, \qquad i = 2, 3, \ldots, k.$$

Also, by Lemma 6.1, we can easily see that the right hand of (6.4) attaches minimum at

$$\hat{\boldsymbol{\Psi}}_{ii} = (\boldsymbol{B}_i^{1/2}\boldsymbol{C}_i^{-1/2})^{-1/2}\boldsymbol{T}_{ii}^{-1} = \boldsymbol{H}_i^{-1/2}\boldsymbol{T}_{ii}^{-1}, \qquad i = 1, 2, \ldots, k,$$

which completes the proof.

Similar to Corollaries 5.1 and 5.2, we give the corresponding results for the symmetric loss L_2 without proofs.

COROLLARY 6.1. *Under the symmetric loss L_2, the best $\mathcal{G}$-equivariant estimator of $\boldsymbol{\Omega}$ is the same as the Bayesian estimator with respect to the left Haar invariant measure $\nu_{\mathcal{G}}^l(d\boldsymbol{\Psi})$ and is given by*

$$\hat{\boldsymbol{\Omega}}_{2B} = \boldsymbol{R}'\boldsymbol{H}_B^{-1}\boldsymbol{R}, \tag{6.5}$$

where $\boldsymbol{R}$ is given by (3.10), $\boldsymbol{H}_B = \mathrm{diag}(\boldsymbol{H}_{1B}, \boldsymbol{H}_{2B}, \ldots, \boldsymbol{H}_{kB})$. *Here* $\boldsymbol{H}_{iB} = \mathrm{diag}(h_{i1B}, \ldots, h_{ip_iB})$ *and*

$$h_{1jB} = \left\{\frac{n-1}{(n-j-1)(n-j)(n+p-2j+1)}\right\}^{1/2}, \qquad j = 1, 2, \ldots, p_1;$$

$$h_{ijB} = \left\{\frac{n-1}{(n-p_1-j-1)(n-p_1-j)(n-p_1+p_i-2j+1)}\right\}^{1/2}, \qquad j = 1, 2, \ldots, p_i;\ i = 2, \ldots, k.$$

Remark 5. Similar to Remark 1, both the MLE $\hat{\boldsymbol{\Omega}}_M$ and the unbiased estimator $\hat{\boldsymbol{\Omega}}_U$ are also inadmissible under the symmetric loss L_2.

Remark 6. Similar to Remark 4, the best $\mathcal{G}$-equivariant estimators $\hat{\boldsymbol{\Omega}}_2$ is also minimax with respect to the symmetric loss L_2.

COROLLARY 6.2. *Under the symmetric loss L_2, the Bayesian estimator $\hat{\boldsymbol{\Omega}}_{2J}$ of $\boldsymbol{\Omega}$ with respect to the Jeffreys prior $\pi_J(\boldsymbol{\Psi})$ is*

$$\hat{\boldsymbol{\Omega}}_{2J} = \boldsymbol{R}'\boldsymbol{H}_J^{-1}\boldsymbol{R}, \tag{6.6}$$

where

$$\boldsymbol{H}_J = \mathrm{diag}\Bigg(\left\{\frac{1}{n(n-p-1)}\right\}^{1/2}\boldsymbol{I}_{p_1}, \left\{\frac{n-p+p_1-1}{n(n-p-1)(n-p_2-1)}\right\}^{1/2}\boldsymbol{I}_{p_2}, \ldots, \left\{\frac{n-p+p_1-1}{n(n-p-1)(n-p_k-1)}\right\}^{1/2}\boldsymbol{I}_{p_k}\Bigg).$$

The Bayesian estimator $\hat{\boldsymbol{\Omega}}_{2R}$ with respect to the reference prior $\pi_R(\boldsymbol{\Psi})$ under the symmetric loss L_2 is

$$\hat{\boldsymbol{\Omega}}_{2R} = \boldsymbol{R}'\boldsymbol{H}_R^{-1}\boldsymbol{R}, \tag{6.7}$$

where $\boldsymbol{H}_R = \mathrm{diag}(\boldsymbol{H}_{1R}, \boldsymbol{H}_{2R}, \ldots, \boldsymbol{H}_{kR})$ *and* $\boldsymbol{H}_{iR} = \mathrm{diag}(h_{i1R}, h_{i2R}, \ldots, h_{ip_iR})$ *with*

$$h_{ijR} = \begin{cases} \{\frac{(n-1)^{j-1}}{(n+p-j)(n-2)^j}\}^{1/2}, & \text{if } i = 1,\ j = 1, \ldots, p_1, \\ \{1 + \sum_{r=1}^{p_1} \frac{(n-1)^{j-1}}{(n-2)^j}\}^{1/2}\{\frac{(n-1)^{j-1}}{(n+p_i-j)(n-2)^j}\}^{1/2}, & \\ & \text{if } i = 2, \ldots, k,\ j = 1, \ldots, p_i. \end{cases}$$

7. Risks of equivariant estimators of $\boldsymbol{\Omega}$

In this section, we will calculate the risks of equivariant estimators defined by (3.11) under the entropy loss L_1 and the symmetric loss L_2.

THEOREM 7.1. *Suppose that* $\boldsymbol{Q}$ *is in the group* $\mathcal{G}$ *and has a similar block partition of* $\boldsymbol{\Psi}$ *as in* (3.2). *Then under the entropy loss* L_1, *the risk of the equivariant estimator* $\hat{\boldsymbol{\Omega}} = \boldsymbol{R}'\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{R}$ *is given by*

$$\begin{aligned}(7.1)\qquad R_1(\hat{\boldsymbol{\Omega}}, \boldsymbol{\Omega}) &= \sum_{i=1}^{k}\{\mathrm{tr}(\boldsymbol{Q}_{ii}\boldsymbol{B}_{iB}\boldsymbol{Q}_{ii}') - \log|\boldsymbol{Q}_{ii}\boldsymbol{Q}_{ii}'| - p_i\} + \sum_{i=2}^{k}\mathrm{tr}(\boldsymbol{Q}_{i1}\boldsymbol{B}_{1B}\boldsymbol{Q}_{i1}') \\ &\quad + \sum_{j=1}^{p_1} E(\log\chi^2_{n-j+1}) + \sum_{i=2}^{k}\sum_{j=1}^{p_i} E(\log\chi^2_{n-p_1-j+1}),\end{aligned}$$

where $\boldsymbol{B}_{iB}$, $i = 1, \ldots, k$ *were defined in Corollary* 5.1 *and* χ^2_m *stands for the central Chi-square distribution with* m *degrees of freedom.*

PROOF. Because the risk of any $\mathcal{G}$-equivariant estimator will not depend on $\boldsymbol{\Omega}$, without loss of generality, we assume that $\boldsymbol{\Omega} = \boldsymbol{I}_p$. For $\boldsymbol{R}$ defined in (3.10), we have

$$\begin{aligned}(7.2)\qquad \mathrm{tr}(\boldsymbol{R}'\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{R}) &= \sum_{i=1}^{k}\mathrm{tr}\{\boldsymbol{Q}_{ii}(\boldsymbol{T}_{ii}'\boldsymbol{T}_{ii})^{-1}\boldsymbol{Q}_{ii}'\} \\ &\quad + \sum_{i=2}^{k}\mathrm{tr}\{\boldsymbol{Q}_{ii}\boldsymbol{T}_{ii}^{-1}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1i}(\boldsymbol{T}_{ii}')^{-1}\boldsymbol{Q}_{ii}\} \\ &\quad + \sum_{i=2}^{k}\mathrm{tr}\{\boldsymbol{Q}_{i1}(\boldsymbol{T}_{11}'\boldsymbol{T}_{11})^{-1}\boldsymbol{Q}_{i1}'\} \\ &\quad - \sum_{i=2}^{k}\mathrm{tr}\{\boldsymbol{Q}_{i1}\boldsymbol{T}_{11}^{-1}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1i}(\boldsymbol{T}_{ii}')^{-1}\boldsymbol{Q}_{ii}'\}.\end{aligned}$$

From (2.8), it follows

$$\begin{aligned} E(\boldsymbol{T}_{11}'\boldsymbol{T}_{11})^{-1} &= \boldsymbol{B}_{1B}, \\ E(\boldsymbol{T}_{ii}'\boldsymbol{T}_{ii})^{-1} &= \frac{n-p_1-1}{n-1}\boldsymbol{B}_{iB}, \qquad i = 2, \ldots, k.\end{aligned}$$

In addition, by (2.10) and the independence between $(\boldsymbol{S}_{i1}, \boldsymbol{S}_{11})$ and $\boldsymbol{S}_{ii\cdot 1}$, we get

$$\begin{aligned} E\{\boldsymbol{T}_{11}^{-1}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1i}(\boldsymbol{T}'_{ii})^{-1}\} &= E\{\boldsymbol{T}_{11}^{-1}\boldsymbol{S}_{11}^{-1}E(\boldsymbol{S}_{1i} \mid \boldsymbol{S}_{11})\}E(\boldsymbol{T}'_{ii})^{-1} = \boldsymbol{0}, \\ E\{\boldsymbol{T}_{ii}^{-1}\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1i}(\boldsymbol{T}'_{ii})^{-1}\} &= E\{\boldsymbol{T}_{ii}^{-1}E(\boldsymbol{S}_{i1}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{1i} \mid \boldsymbol{S}_{11})(\boldsymbol{T}'_{ii})^{-1}\} \\ &= E\{\mathrm{tr}(\boldsymbol{S}_{11}^{-1})\}E(\boldsymbol{T}'_{ii}\boldsymbol{T}_{ii})^{-1} = p_1\boldsymbol{B}_{iB}/(n-1). \end{aligned}$$

The last equality holds because $E(\boldsymbol{S}_{11}^{-1}) = \boldsymbol{I}_{p_1}/(n-p_1-1)$. Therefore,

$$E\{\mathrm{tr}(\boldsymbol{R}'\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{R})\} = \sum_{i=1}^{k}\mathrm{tr}(\boldsymbol{Q}_{ii}\boldsymbol{B}_{iB}\boldsymbol{Q}'_{ii}) + \sum_{i=2}^{k}\mathrm{tr}(\boldsymbol{Q}_{i1}\boldsymbol{B}_{1B}\boldsymbol{Q}'_{i1}).$$

Moreover,

$$\begin{aligned} &E(\log|\boldsymbol{R}'\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{R}|) \\ &\quad= \log|\boldsymbol{Q}'\boldsymbol{Q}| + E\left(\log\prod_{i=1}^{k}|\boldsymbol{S}_{ii\cdot 1}^{-1}|\right) = \sum_{i=1}^{k}\log|\boldsymbol{Q}_i| - \sum_{i=1}^{k}E(\log|\boldsymbol{S}_{ii\cdot 1}|) \\ &\quad= \sum_{i=1}^{k}\log|\boldsymbol{Q}_{ii}\boldsymbol{Q}'_{ii}| - \sum_{j=1}^{p_1}E(\log\chi^2_{n-j+1}) - \sum_{i=2}^{k}\sum_{j=1}^{p_i}E(\log\chi^2_{n-p_1-j+1}). \end{aligned}$$

This implies (7.1).

If $\boldsymbol{Q}$ is diagonal, we have the following corollary.

COROLLARY 7.1. *Define* $\boldsymbol{W} = \mathrm{diag}(\boldsymbol{W}_1, \boldsymbol{W}_2, \ldots, \boldsymbol{W}_k)$, *where* $\boldsymbol{W}_i = \mathrm{diag}(w_{i1}, \ldots, w_{ip_i})$, $i = 1, \ldots k$. *Under the entropy loss* L_1, *the risk of the equivariant estimator* $\hat{\boldsymbol{\Omega}} = \boldsymbol{R}'\boldsymbol{W}\boldsymbol{R}$ *is*

$$\begin{aligned} R_1(\hat{\boldsymbol{\Omega}}, \boldsymbol{\Omega}) &= \sum_{i=1}^{k}\sum_{j=1}^{p_i}(b_{ijB}w_{ij} - \log w_{ij} - 1) \\ &\quad + \sum_{j=1}^{p_1}E(\log\chi^2_{n-j+1}) + \sum_{i=2}^{k}\sum_{j=1}^{p_i}E(\log\chi^2_{n-p_1-j+1}), \end{aligned} \tag{7.3}$$

where b_{ijB} *is defined by* (5.13) *in Corollary* 5.1.

By (7.1), the risk $R_1(\hat{\boldsymbol{\Omega}}, \boldsymbol{\Omega})$ consists of two parts, namely,

$$\begin{aligned} R_{11} &= \sum_{i=1}^{k}\{\mathrm{tr}(\boldsymbol{Q}_{ii}\boldsymbol{B}_{iB}\boldsymbol{Q}'_{ii}) - \log|\boldsymbol{Q}_{ii}\boldsymbol{Q}'_{ii}| - p_i\} + \sum_{i=2}^{k}\mathrm{tr}(\boldsymbol{Q}_{i1}\boldsymbol{B}_{1B}\boldsymbol{Q}'_{i1}), \\ R_{12} &= \sum_{j=1}^{p_1}E(\log\chi^2_{n-j+1}) + \sum_{i=2}^{k}\sum_{j=1}^{p_i}E(\log\chi^2_{n-p_1-j+1}), \end{aligned}$$

where the second part R_{12} is independent of $\boldsymbol{Q}$ and just depends on n and $(p_1, p_2, \ldots, p_k)$. The best equivariant estimator $\hat{\boldsymbol{\Omega}}_{1B}$ also can be derived by minimizing (7.1) or R_{11}.

From Remark 4, the best $\mathcal{G}$-equivariant estimator $\hat{\boldsymbol{\Omega}}_{1B}$ is minimax and thus we get the minimax risk in the following:

Remark 7. Under the entropy loss L_1, the minimax risk is given by

$$\sum_{i=1}^{k}\sum_{j=1}^{p_i}\log b_{ijB} + \sum_{j=1}^{p_1} E(\log\chi^2_{n-j+1}) + \sum_{i=2}^{k}\sum_{j=1}^{p_i} E(\log\chi^2_{n-p_1-j+1}), \tag{7.4}$$

where b_{ijB} is defined by (5.13) in Corollary 5.1.

Based on (7.3), we can easily get the risk expressions for the MLE $\hat{\boldsymbol{\Omega}}_M$, the unbiased estimator $\hat{\boldsymbol{\Omega}}_U$, and the Bayesian estimators $\hat{\boldsymbol{\Omega}}_{1J}$ and $\hat{\boldsymbol{\Omega}}_{1R}$ under the entropy loss L_1. Some numerical results will be given in next section.

We now give the frequentist risks of a general class of equivariant estimators under the symmetric loss L_2. The derivation of the risks is similar to that under the entropy loss and is omitted.

THEOREM 7.2. *Let* $\boldsymbol{G} = \mathrm{diag}(\boldsymbol{G}_1, \ldots, \boldsymbol{G}_k)$ *and* $\boldsymbol{G}_i = \mathrm{diag}(g_{i1}, \ldots, g_{ip_i})$ *with*

$$g_{ij} = \begin{cases} n + p - 2j + 1, & \text{if } i = 1,\ j = 1, \ldots, p_1, \\ n - p_1 + p_i - 2i + 1, & \text{if } i = 2, \ldots, k,\ j = 1, \ldots, p_i. \end{cases} \tag{7.5}$$

Then for any $\boldsymbol{Q} \in \mathcal{G}$ *in Theorem* 7.1, *the risk of the* $\mathcal{G}$*-equivariant estimator* $\hat{\boldsymbol{\Omega}} = \boldsymbol{R}'\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{R}$ *under the symmetric loss* L_2 *is*

$$\begin{aligned} R_2(\hat{\boldsymbol{\Omega}}, \boldsymbol{\Omega}) &= \sum_{i=1}^{k}[\mathrm{tr}(\boldsymbol{Q}_{ii}\boldsymbol{B}_{iB}\boldsymbol{Q}'_{ii}) + \mathrm{tr}\{(\boldsymbol{Q}'_{ii})^{-1}\boldsymbol{G}_i\boldsymbol{Q}_{ii}^{-1}\}] - 2p \\ &\quad + \sum_{i=2}^{k}[\mathrm{tr}(\boldsymbol{Q}_{i1}\boldsymbol{B}_{1B}\boldsymbol{Q}'_{i1}) \\ &\qquad + \mathrm{tr}\{\boldsymbol{Q}_{ii}^{-1}\boldsymbol{Q}_{i1}(\boldsymbol{Q}'_{11}\boldsymbol{Q}_{11})^{-1}\boldsymbol{Q}'_{i1}(\boldsymbol{Q}'_{ii})^{-1}\boldsymbol{G}_i\}], \end{aligned} \tag{7.6}$$

where $\boldsymbol{B}_{iB}$ *is defined by* (5.13) *in Corollary* 5.1, $i = 1, \ldots, k$.

COROLLARY 7.2. *Suppose that* $\boldsymbol{W}$ *is the same as in Corollary* 7.1. *Under the symmetric loss* L_2, *the risk of the equivariant estimator* $\hat{\boldsymbol{\Omega}} = \boldsymbol{R}'\boldsymbol{W}\boldsymbol{R}$ *is*

$$R_2(\hat{\boldsymbol{\Omega}}, \boldsymbol{\Omega}) = \sum_{i=1}^{k}\sum_{j=1}^{p_i}(b_{ijB}w_{ij} + w_{ij}^{-1}g_{ij} - 1), \tag{7.7}$$

where b_{ijB} *and* g_{ij} *are defined by* (5.13) *and* (7.5), *respectively.*

Remark 8. Under the symmetric loss L_2, the minimax risk is given by

$$\begin{aligned} &2\sum_{j=1}^{p_1}\left\{\frac{(n-1)(n+p-2j+1)}{(n-j-1)(n-j)}\right\}^{1/2} \\ &\quad + 2\sum_{i=2}^{k}\sum_{j=1}^{p_i}\left\{\frac{(n-1)(n-p_1+p-2j+1)}{(n-p_1-j-1)(n-p_1-j)}\right\}^{1/2} - 2p. \end{aligned} \tag{7.8}$$

8. Estimating the covariance matrix $\boldsymbol{\Sigma}$

As immediate corollaries of our results on estimating the precision matrix, we now list the results for estimating covariance matrix under the star-shape model.

For estimating the covariance matrix, the entropy loss and the symmetric loss are

$$L_1^*(\hat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) = \operatorname{tr}(\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}) - \log|\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}| - p \tag{8.1}$$

and

$$L_2^*(\hat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) = L_0^*(\hat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) + L_1^*(\hat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) = \operatorname{tr}(\hat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}) + \operatorname{tr}(\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}) - 2p. \tag{8.2}$$

THEOREM 8.1. *Under the entropy loss L_1^*, the Bayesian estimator of $\boldsymbol{\Sigma}$ with respect to the prior $p(\boldsymbol{\Psi})$ in* (4.1) *is given by*

$$\hat{\boldsymbol{\Sigma}}_1 = \boldsymbol{T}\boldsymbol{B}\boldsymbol{T}', \tag{8.3}$$

where $\boldsymbol{T}$ is given by (3.8) *and $\boldsymbol{B}$ is defined in Theorem* 5.1.

THEOREM 8.2. *Suppose that $\boldsymbol{P}$ is in the group $\mathcal{G}$ and has a similar block partition of $\boldsymbol{\Psi}$ as in* (3.2). *Then, under the entropy loss L_1^*, the risk of the equivariant estimator $\hat{\boldsymbol{\Sigma}} = \boldsymbol{T}\boldsymbol{P}\boldsymbol{P}'\boldsymbol{T}'$ is given by*

$$\begin{aligned} R_1^*(\hat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) = & \sum_{i=1}^{k}[\operatorname{tr}\{\boldsymbol{P}_{ii}^{-1}\boldsymbol{B}_{iB}(\boldsymbol{P}_{ii}')^{-1}\} - \log|\boldsymbol{P}_{ii}^{-1}(\boldsymbol{P}_{ii}')^{-1}| - p_i] \\ & + \sum_{i=2}^{k}\operatorname{tr}\{\boldsymbol{P}_{i1}^{-1}\boldsymbol{P}_{i1}\boldsymbol{P}_{11}^{-1}\boldsymbol{B}_{1B}(\boldsymbol{P}_{i1}^{-1}\boldsymbol{P}_{i1}\boldsymbol{P}_{11}^{-1})'\} \\ & + \sum_{j=1}^{p_1}E(\log\chi^2_{n-j+1}) \\ & + \sum_{i=2}^{k}\sum_{j=1}^{p_i}E(\log\chi^2_{n-p_1-j+1}), \end{aligned} \tag{8.4}$$

where $\boldsymbol{B}_{iB}$, $i = 1, \ldots, k$ were defined in Corollary 5.1 *and χ_m^2 stands for the central Chi-square distribution with m degrees of freedom.*

COROLLARY 8.1. *Under the entropy loss L_1^*, the best $\mathcal{G}$-equivariant estimate of $\boldsymbol{\Sigma}$ is given by*

$$\hat{\boldsymbol{\Sigma}}_{1B} = \boldsymbol{T}\boldsymbol{B}_B\boldsymbol{T}', \tag{8.5}$$

where $\boldsymbol{T}$ is given by (3.8) *and $\boldsymbol{B}_B$ is shown in Corollary* 5.1. *Furthermore, $\hat{\boldsymbol{\Sigma}}_{1B}$ is minimax and its minimax risk is given by* (7.4).

THEOREM 8.3. *Under the symmetric loss L_2^*, the Bayesian estimator of $\boldsymbol{\Sigma}$ with respect to the prior $p(\boldsymbol{\Psi})$ in* (4.1) *is given by*

$$\hat{\boldsymbol{\Sigma}}_2 = \boldsymbol{T}\boldsymbol{H}\boldsymbol{T}', \tag{8.6}$$

where $\boldsymbol{T}$ *is given by* (3.8) *and* $\boldsymbol{H}$ *is defined in Theorem* 6.1.

THEOREM 8.4. *Under the symmetric loss* L_2^*, *then, for any* $\boldsymbol{P} \in \mathcal{G}$, *the risk of the* $\mathcal{G}$*-equivariant estimator* $\hat{\boldsymbol{\Sigma}} = \boldsymbol{T P P' T'}$ *is*

$$(8.7) \quad R_2^*(\hat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) = \sum_{i=1}^{k} [\mathrm{tr}\{\boldsymbol{P}_{ii}^{-1} \boldsymbol{B}_{iB} (\boldsymbol{P}_{ii}')^{-1}\} + \mathrm{tr}(\boldsymbol{P}_{ii}' \boldsymbol{G}_i \boldsymbol{P}_{ii})] + \sum_{i=2}^{k} [\mathrm{tr}\{\boldsymbol{P}_{ii}^{-1} \boldsymbol{P}_{i1} \boldsymbol{P}_{11}^{-1} \boldsymbol{B}_{1B} (\boldsymbol{P}_{ii}^{-1} \boldsymbol{P}_{i1} \boldsymbol{P}_{11}^{-1})'\} + \mathrm{tr}(\boldsymbol{P}_{i1} \boldsymbol{P}_{i1}' \boldsymbol{G}_i)] - 2p,$$

where $\boldsymbol{B}_{iB}$ *is defined by* (5.13) *in Corollary* 5.1, $i = 1, \dots, k$ *and* $\boldsymbol{G} = \mathrm{diag}(\boldsymbol{G}_1, \dots, \boldsymbol{G}_k)$ *is given in Theorem* 7.2.

COROLLARY 8.2. *Under the symmetric loss* L_2^*, *the best* $\mathcal{G}$*-equivariant estimator of* $\boldsymbol{\Sigma}$ *is given by*

$$(8.8) \qquad \hat{\boldsymbol{\Sigma}}_{2B} = \boldsymbol{T H}_B \boldsymbol{T}',$$

where $\boldsymbol{T}$ *is given by* (3.8) *and* $\boldsymbol{H}_B$ *is shown in Corollary* 6.1. *Also,* $\hat{\boldsymbol{\Sigma}}_{2B}$ *is minimax and its minimax risk is given by* (7.8).

9. Numerical results

9.1 *Numerical computation*

In this subsection, we will compare the risks of MLE $\hat{\boldsymbol{\Omega}}_M$, the unbiased estimator $\hat{\boldsymbol{\Omega}}_U$, the best equivariant estimator $\hat{\boldsymbol{\Omega}}_{1B}$, the Bayesian estimator $\hat{\boldsymbol{\Omega}}_{1J}$, and the Bayesian estimator $\hat{\boldsymbol{\Omega}}_{1R}$ under the entropy loss L_1. Each risk will be denoted as $R_{1M}, R_{1U}, R_{1B}, R_{1J}$ and R_{1R}, respectively. We also will compare the risks of MLE $\hat{\boldsymbol{\Omega}}_M$, the unbiased estimator $\hat{\boldsymbol{\Omega}}_U$, the best equivariant estimator $\hat{\boldsymbol{\Omega}}_{2B}$, the Bayesian estimator $\hat{\boldsymbol{\Omega}}_{2J}$, and the Bayesian estimator $\hat{\boldsymbol{\Omega}}_{2R}$ under the symmetric loss L_2, denoting each risk as $R_{2M}, R_{2U}, R_{2B}, R_{2J}$ and R_{2R}.

For the entropy loss L_1, we will denote each first part on the right hand of (7.3) as R_{11M}, R_{11U}, R_{11B}, R_{11J}, R_{11R} respectively and the common second term as R_{12}. Note that R_{12} will just depend on n and $(p_1, p_2, \dots, p_k)$. Because there is no explicit form for the expectation of natural logarithm of chi-square distribution, we use Monte Carlo method to get the value for the common second part R_{12}. Some simulation results are given in Table 1. From the simulation study, we found that the improvements over the risk of $\hat{\boldsymbol{\Omega}}_M$ by $\hat{\boldsymbol{\Omega}}_{1B}$ are significant. Of course, the best equivariant estimator $\hat{\boldsymbol{\Omega}}_{1B}$ is the best among five estimators. Simulation study also shows that these five estimators have the following relationship, $\hat{\boldsymbol{\Omega}}_{1B} \prec \hat{\boldsymbol{\Omega}}_{1R} \prec \hat{\boldsymbol{\Omega}}_{1J} \prec \hat{\boldsymbol{\Omega}}_U \prec \hat{\boldsymbol{\Omega}}_M$, where "$\prec$" stands for "better than". Another interesting thing is that except the best equivariant estimator $\hat{\boldsymbol{\Omega}}_{1B}$, the Bayesian estimator $\hat{\boldsymbol{\Omega}}_{1R}$ with respect to the reference prior will be the best one because the power of each ψ_{ii} is always one.

For the symmetric loss L_2, we can compare their risks by (7.7) directly. The improvements over the risk of $\hat{\boldsymbol{\Omega}}_M$ by $\hat{\boldsymbol{\Omega}}_{2B}$ are also significant. Some simulation results are given in Table 2. The relationship among $\hat{\boldsymbol{\Omega}}_M$, $\hat{\boldsymbol{\Omega}}_U$, $\hat{\boldsymbol{\Omega}}_{2B}$, $\hat{\boldsymbol{\Omega}}_{2J}$, $\hat{\boldsymbol{\Omega}}_{2R}$ will be $\hat{\boldsymbol{\Omega}}_{2B} \prec \hat{\boldsymbol{\Omega}}_{2R} \prec \hat{\boldsymbol{\Omega}}_{2J} \prec \hat{\boldsymbol{\Omega}}_M \prec \hat{\boldsymbol{\Omega}}_U$. The Bayesian estimator $\hat{\boldsymbol{\Omega}}_{2R}$ with respect to the reference prior is still the best one except the best equivariant estimator $\hat{\boldsymbol{\Omega}}_{2B}$.

Table 1. Risks of $\hat{\boldsymbol{\Omega}}_M$, $\hat{\boldsymbol{\Omega}}_U$, $\hat{\boldsymbol{\Omega}}_{1B}$, $\hat{\boldsymbol{\Omega}}_{1J}$ and $\hat{\boldsymbol{\Omega}}_{1R}$ under L_1.

p	$p_i's$	n	R_{1M}	R_{1U}	R_{1B}	R_{1J}	R_{1R}
4	$(2,1,1)$	7	4.9099	2.6100	1.9938	2.4746	2.3896
		12	1.4941	1.0497	0.9384	1.0381	1.0030
		17	0.8502	0.6658	0.6201	0.6628	0.6454
		22	0.5825	0.4819	0.4571	0.4807	0.4705
	$(1,2,1)$	7	3.7287	2.1356	1.7018	2.0869	1.9399
		12	1.2322	0.9032	0.8200	0.8979	0.8609
		17	0.7197	0.5803	0.5457	0.5789	0.5620
		22	0.4998	0.4231	0.4042	0.4225	0.4129
	$(1,1,1,1)$	7	2.5475	1.7791	1.4097	1.7260	1.4903
		12	0.9704	0.7724	0.7017	0.7653	0.7189
		17	0.5891	0.5008	0.4713	0.4988	0.4786
5	$(2,2,1)$	8	6.4323	3.3930	2.4771	3.2286	3.1270
		13	2.0997	1.4432	1.2561	1.4242	1.3747
		18	1.2126	0.9301	0.8503	0.9248	0.8983
	$(1,2,2)$	8	4.0314	2.5236	1.9599	2.4751	2.2137
		13	1.5492	1.1543	1.0303	1.1458	1.0865
		18	0.9387	0.7602	0.7061	0.7577	0.7301
	$(3,1,1)$	8	8.0801	3.8673	2.7691	3.6136	3.7236
		13	2.4282	1.5897	1.3744	1.5661	1.5366
		18	1.3681	1.0156	0.9247	1.0092	0.9888
	$(1,3,1)$	8	5.6792	2.9447	2.2519	2.8989	2.8103
		13	1.8777	1.2975	1.1486	1.2909	1.2484
		18	1.0942	0.8448	0.7805	0.8429	0.8207
	$(2,1,1,1)$	8	4.7844	3.0364	2.1851	2.8495	2.5304
		13	1.7712	1.3124	1.1377	1.2866	1.2127
		18	1.0572	0.8506	0.7759	0.8431	0.8077
	$(1,2,1,1)$	8	3.2783	2.2576	1.7348	2.2130	1.8970
		13	1.3273	1.0368	0.9228	1.0271	0.9603
		18	0.8203	0.6859	0.6362	0.6829	0.6525

One thing should be mentioned if there is $\beta_{ij} \neq 0$, then the risk of the Bayesian estimator with respect to the prior $p(\boldsymbol{\Psi})$ under either L_1 or L_2 is very complicated. It is not clear whether there is such a Bayesian estimator that will be better than the maximal likelihood estimator or the best equivariant estimator in theoretical view. We will explore this point in the future.

9.2 *Analysis of a real example*

We now analyze a data set from Mardia *et al.* (1979). It consists of the examination marks of 88 students in five subjects: algebra, mechanics, vectors, analysis, and statistics. Mechanics and vectors were closed book examinations and the reminders were open book. Whittaker (1990) shows that given algebra, mechanics, and vectors are conditionally independent with analysis and statistics. Because the mean of the population distribution is unknown, we will have $\boldsymbol{S} = \sum_{i=1}^{88}(\boldsymbol{Y}_i - \bar{Y})(\boldsymbol{Y}_i - \bar{Y})'$, which follows

Table 2. Risks of $\hat{\mathbf{\Omega}}_M$, $\hat{\mathbf{\Omega}}_U$, $\hat{\mathbf{\Omega}}_{2B}$, $\hat{\mathbf{\Omega}}_{2J}$ and $\hat{\mathbf{\Omega}}_{2R}$ under L_2.

p	$p_i's$	n	R_{2M}	R_{2U}	R_{2B}	R_{2J}	R_{2R}
4	$(2,1,1)$	7	6.5000	8.3333	4.2514	5.2553	4.9304
		12	2.3333	2.5000	1.9146	2.1139	2.0319
		17	1.4176	1.4744	1.2449	1.3296	1.2921
		22	1.0175	1.0458	0.9236	0.9704	0.9490
	$(1,2,1)$	7	5.1000	6.5000	3.5723	4.3461	3.9833
		12	1.9667	2.1151	1.6627	1.8155	1.7374
		17	1.2183	1.2711	1.0905	1.1556	1.1210
		22	0.8825	0.9092	0.8122	0.8483	0.8288
	$(1,1,1,1)$	7	3.7000	5.5000	2.8783	3.4947	3.0219
		12	1.6000	1.8095	1.4065	1.5308	1.4387
		17	1.0190	1.0952	0.9340	0.9879	0.9479
5	$(2,2,1)$	8	8.4667	11.5000	5.3279	6.8967	6.4342
		13	3.2333	3.5437	2.5787	2.9161	2.7916
		18	1.9956	2.1044	1.7154	1.8636	1.8040
	$(1,2,2)$	8	5.6667	8.0000	4.0900	5.1135	4.5328
		13	2.4848	2.7619	2.0941	2.3209	2.1970
		18	1.5893	1.6905	1.4142	1.5153	1.4592
	$(3,1,1)$	8	10.3333	13.3333	6.0614	7.8943	7.6748
		13	3.6667	3.9286	2.8418	3.2311	3.1309
		18	2.2198	2.3077	1.8746	2.0439	1.9924
	$(1,3,1)$	8	7.5333	9.4000	4.8289	6.1593	5.7772
		13	2.9182	3.1286	2.3581	2.6396	2.5371
		18	1.8135	1.8897	1.5737	1.6965	1.6478
	$(2,1,1,1)$	8	6.6000	10.5000	4.5783	5.9182	5.1794
		13	2.8000	3.2381	2.3111	2.6100	2.4481
		18	1.7714	1.9286	1.5542	1.6881	1.6137
	$(1,2,1,1)$	8	4.7333	7.3000	3.5713	4.5048	3.8566
		13	2.1697	2.4952	1.8668	2.0712	1.9359
		18	1.4071	1.5286	1.2691	1.3605	1.2997

$W_5(87, \mathbf{\Sigma})$ with

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} & \mathbf{\Sigma}_{13} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} & \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{13} \\ \mathbf{\Sigma}_{31} & \mathbf{\Sigma}_{31}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12} & \mathbf{\Sigma}_{33} \end{pmatrix},$$

where $\mathbf{\Sigma}_{11}, \mathbf{\Sigma}_{22}, \mathbf{\Sigma}_{33}$ are 1 by 1, 2 by 2 and 2 by 2, respectively. From the data, we get the maximum likelihood estimates of the covariance matrix $\mathbf{\Sigma}$ and the precision matrix $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$ as follows,

$$\hat{\mathbf{\Sigma}}_M = \begin{pmatrix} 112.8860 & 101.5794 & 85.1573 & 112.1134 & 121.8706 \\ 101.5794 & 305.7680 & 127.2226 & 100.8842 & 109.6641 \\ 85.1573 & 127.2226 & 172.8422 & 84.5744 & 91.9349 \\ 112.1134 & 100.8842 & 84.5744 & 220.3804 & 155.5355 \\ 121.8706 & 109.6641 & 91.9349 & 155.5355 & 297.7554 \end{pmatrix}$$

and

$$\hat{\mathbf{\Omega}}_M = \begin{pmatrix} 0.0285 & -0.0029 & -0.0056 & -0.0075 & -0.0049 \\ -0.0029 & 0.0052 & -0.0024 & 0 & 0 \\ -0.0056 & -0.0024 & 0.0103 & 0 & 0 \\ -0.0075 & 0 & 0 & 0.0098 & -0.0020 \\ -0.0049 & 0 & 0 & -0.0020 & 0.0064 \end{pmatrix}.$$

According to Corollary 5.1 and Corollary 8.1, the best equivariant estimates of $\mathbf{\Sigma}$ and $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$ under the entropy loss are

$$\hat{\mathbf{\Sigma}}_{1B} = \begin{pmatrix} 115.5421 & 103.9695 & 87.1610 & 114.7513 & 124.7381 \\ 103.9695 & 318.1865 & 131.4490 & 103.2579 & 112.2444 \\ 87.1610 & 131.4490 & 181.9965 & 86.5644 & 94.0981 \\ 114.7513 & 103.2579 & 86.5644 & 228.2229 & 160.0359 \\ 124.7381 & 112.2444 & 94.0981 & 160.0359 & 312.7319 \end{pmatrix}$$

and

$$\hat{\mathbf{\Omega}}_{1B} = \begin{pmatrix} 0.0272 & -0.0028 & -0.0052 & -0.0072 & -0.0046 \\ -0.0028 & 0.0050 & -0.0023 & 0 & 0 \\ -0.0052 & -0.0023 & 0.0096 & 0 & 0 \\ -0.0072 & 0 & 0 & 0.0094 & -0.0019 \\ -0.0046 & 0 & 0 & -0.0019 & 0.0060 \end{pmatrix}.$$

Similarly, we can get the best equivariant estimates of $\mathbf{\Sigma}$ and $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$ under the symmetric loss

$$\hat{\mathbf{\Sigma}}_{2B} = \begin{pmatrix} 111.6681 & 100.4835 & 84.2385 & 110.9038 & 120.5557 \\ 100.4835 & 309.8557 & 127.5934 & 99.7958 & 108.4810 \\ 84.2385 & 127.5934 & 177.0760 & 83.6620 & 90.9431 \\ 110.9038 & 99.7958 & 83.6620 & 221.7599 & 155.0463 \\ 120.5557 & 108.4810 & 90.9431 & 155.0463 & 304.0543 \end{pmatrix}$$

and

$$\hat{\mathbf{\Omega}}_{2B} = \begin{pmatrix} 0.0280 & -0.0028 & -0.0053 & -0.0074 & -0.0047 \\ -0.0028 & 0.0051 & -0.0023 & 0 & 0 \\ -0.0053 & -0.0023 & 0.0099 & 0 & 0 \\ -0.0074 & 0 & 0 & 0.0096 & -0.0019 \\ -0.0047 & 0 & 0 & -0.0019 & 0.0061 \end{pmatrix}.$$

For the entropy loss L_1, the posterior risks of $\hat{\mathbf{\Omega}}_M, \hat{\mathbf{\Omega}}_{1B}, \hat{\mathbf{\Omega}}_{1J}, \hat{\mathbf{\Omega}}_{1R}$ are given in Table 3.

By comparing the posterior risks, we find that the maximum likelihood estimator of $\mathbf{\Omega}$ is always the worst one among the given four estimators under any of three usual priors,

Table 3. Posterior risks of $\hat{\mathbf{\Omega}}_M$, $\hat{\mathbf{\Omega}}_{1B}$, $\hat{\mathbf{\Omega}}_{1J}$, $\hat{\mathbf{\Omega}}_{1R}$ under L_1 and L_2.

	L_1				L_2		
	$\nu_G^l(\mathbf{\Psi})$	$\nu_G^r(\mathbf{\Psi})$	$\pi_R(\mathbf{\Psi})$		$\nu_G^l(\mathbf{\Psi})$	$\nu_G^r(\mathbf{\Psi})$	$\pi_R(\mathbf{\Psi})$
$\hat{\mathbf{\Omega}}_{1B}$	0.1299	0.1339	0.1295	$\hat{\mathbf{\Omega}}_{2B}$	0.2599	0.2678	0.2590
$\hat{\mathbf{\Omega}}_{1J}$	0.1316	0.1322	0.1301	$\hat{\mathbf{\Omega}}_{2J}$	0.2632	0.2645	0.2602
$\hat{\mathbf{\Omega}}_{1R}$	0.1306	0.1335	0.1288	$\hat{\mathbf{\Omega}}_{2R}$	0.2613	0.2670	0.2576
$\hat{\mathbf{\Omega}}_M$	0.1375	0.1394	0.1394	$\hat{\mathbf{\Omega}}_M$	0.2656	0.2681	0.2591

$\nu_G^l(\Psi), \nu_G^r(\Psi), \pi_R(\Psi)$. For example, under the left Haar invariant measure $\nu_G^l(\Psi)$, the MLE $\hat{\Omega}_M$ will be improved by the best equivariant estimator $\hat{\Omega}_{1B}$ for about 5.5 percent. The posterior risks of $\hat{\Omega}_M, \hat{\Omega}_{2B}, \hat{\Omega}_{2J}, \hat{\Omega}_{2R}$ under the symmetric loss L_2 are also given in the above table and the similar comparing results can be obtained.

Acknowledgements

The authors would like to thank Professors James O. Berger, Michael Perlman, and Paul Speckman for helpful comments and discussions. The author is grateful to Dr. Karon F. Speckman for her carefully proof-reading this paper.

References

Andersson, S. A. and Perlman, M. D. (1993). Lattice models for conditional independence in a multivariate normal distribution, *The Annals of Statistics*, **21**, 1318–1358.

Berger, J. O. and Bernardo, J. M. (1992). On the development of reference priors, *Proceedings of the Fourth Valencia International Meeting*, *Bayesian Statistics*, **4**, 35–49.

Bondar, J. V. and Milnes, P. (1981). Amenability: A survey for statistical applications of Hunt-Stein and related conditions on groups, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **57**, 103–128.

Dempster, A. M. (1972). Covariance selection, *Biometrics*, **28**, 157–175.

Eaton, M. L. (1970). Some problems in covariance estimation (prelimary report), Tech. Report, No. 49, Department of Statistics, Stanford University, California.

Eaton, M. L. (1989). *Group Invariance Applications in Statistics*, Institute of Mathematical Statistics, Hayward, California.

Eaton, M. L. and Olkin, I. (1987). Best equivariant estimators of a Cholesky decomposition, *The Annals of Statistics*, **15**, 1639–1650.

Gupta, A. K. and Nagar, D. K. (2000). *Matrix Variate Distributions*, Chapman & Hall Ltd., New York.

Gupta, A. K. and Ofori-Nyarko, S. (1995). Improved minimax estimators of normal covariance and precision matrices, *Statistics*, **26**, 19–25.

Haff, L. R. (1980). Empirical Bayes estimation of the multivariate normal covariance matrix, *The Annals of Statistics*, **8**, 586–597.

James, W. and Stein, C. (1961). Estimation with quadratic loss, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, 361–379, University of California Press, Berkeley.

Kiefer, J. (1957). Invariance, minimax sequential estimation, and continuous time process, *The Annals of Mathematical Statistics*, **28**, 573–601.

Konno, Y. (2001). Inadmissibility of the maximum likelihood estimator of normal covariance matrices with the lattice conditional independence, *Journal of Multivariate Analysis*, **79**, 33–51.

Krishnamoorthy, K. and Gupta, A. K. (1989). Improved minimax estimation of a normal precision matrix, *The Canadian Journal of Statistics*, **17**, 91–102.

Kubokawa, T. and Konno, Y. (1990). Estimating the covariance matrix and the generalized variance under a symmetric loss, *Annals of the Institute of Statistical Mathematics*, **42**, 331–343.

Lauritzen, S. L. (1996). *Graphical Models*, Oxford University Press.

Mardia, K. V., Kent, J. T. and Bibby, J. M. (1979). *Multivariate Analysis*, Academic Press, New York.

Olkin, I. and Selliah, J. B. (1977). Estimating covariances in a multivariate distribution, *Statistical Decision Theory and Related Topics*, *Proceedings of a Symposium at Purdue University*, *May, 1976*, *in two volumes*, Vol. 2, 313–326, Academic Press, New York.

Pal, N. (1993). Estimating the normal dispersion matrix and the precision matrix from a decision-theoretic point of view: A review, *Statistical Papers*, **34**, 1–26.

Robert, C. P. (1994). *The Bayesian Choice: A Decision-Theoretic Motivation*, Springer-Verlag Inc., New York.

Sharma, D. and Krishnamoorthy, K. (1983). Orthogonal equivariant minimax estimators of bivariate normal covariance matrix and precision matrix, *Calcutta Statistical Association Bulletin*, **32**, 23–45.

Sinha, B. K. and Ghosh, M. (1987). Inadmissibility of the best equivariant estimators of the variance-covariance matrix, the precision matrix, and the generalized variance under entropy loss, *Statistics & Decisions*, **5**, 201–227.

Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*, John Wiley & Sons, New York.

Wong, F., Carter, C. K. and Kohn, R. (2003). Efficient estimation of covariance selection models, *Biometrika*, **90**, 809–830.

Yang, R. and Berger, J. O. (1994). Estimation of a covariance matrix using the reference prior, *The Annals of Statistics*, **22**, 1195–1211.