ELSEVIER
Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Multivariate Analysis 97 (2006) 698–719
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Estimation of multivariate normal covariance and precision matrices in a star-shape model with missing data

Dongchu Sun, Xiaoqian Sun*

Department of Statistics, University of Missouri, Columbia, MO 65211, USA

Received 22 September 2003
Available online 7 November 2005

Abstract

In this paper, we study the problem of estimating the covariance matrix $\boldsymbol{\Sigma}$ and the precision matrix $\boldsymbol{\Omega}$ (the inverse of the covariance matrix) in a star-shape model with missing data. By considering a type of Cholesky decomposition of the precision matrix $\boldsymbol{\Omega}=\boldsymbol{\Psi}'\boldsymbol{\Psi}$, where $\boldsymbol{\Psi}$ is a lower triangular matrix with positive diagonal elements, we get the MLEs of the covariance matrix and precision matrix and prove that both of them are biased. Based on the MLEs, unbiased estimators of the covariance matrix and precision matrix are obtained. A special group $\mathcal{G}$, which is a subgroup of the group consisting all lower triangular matrices, is introduced. By choosing the left invariant Haar measure on $\mathcal{G}$ as a prior, we obtain the closed forms of the best equivariant estimates of $\boldsymbol{\Omega}$ under any of the Stein loss, the entropy loss, and the symmetric loss. Consequently, the MLE of the precision matrix (covariance matrix) is inadmissible under any of the above three loss functions. Some simulation results are given for illustration.

AMS 1991 subject classification: 62H12

Keywords: Star-shape model; Missing data; Maximum likelihood estimator; Covariance matrix; Precision matrix; Invariant Haar measure; Stein loss; Entropy loss; Symmetric loss; Inadmissibility

1. Introduction

Multivariate normal distribution plays a key role in multivariate statistical analysis. There is a large literature on estimating the covariance matrix and precision matrix in the saturated multivariate normal population, where no additional restriction other than being positive definite is

* Corresponding author.
E-mail address: xsun@stat.missouri.edu (X. Sun).

0047-259X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2005.09.006

required. See, for example, Haff [9], Sinha and Ghosh [22], Krishnamoorthy and Gupta [14], Yang and Berger [25], 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 case uses several pair variables that are conditionally independent, giving other remaining variables. For multivariate normal distribution, this will correspond to some zeros among the entries of the precision matrix. See Dempster [4], Whittaker [24], or Lauritzen [16].

Assume that $\mathbf{X} \sim N_p(0, \mathbf{\Sigma})$. The vector $\mathbf{X}$ is partitioned into k groups, that is, $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2', \ldots, \mathbf{X}_k')'$, where $\mathbf{X}_i$ is p_i-dimensional, and $\sum_{i=1}^k p_i = p$. We assume that for giving $\mathbf{X}_1$, the other subvectors $\mathbf{X}_2, \ldots, \mathbf{X}_k$ are mutually conditionally independent. From Whittaker [24] and Lauritzen [16], the precision matrix $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$ has the following special structure:

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

In fact, we can easily show that (1) is equivalent to

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} & \mathbf{\Sigma}_{13} & \cdots & \mathbf{\Sigma}_{1k} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} & \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{13} & \cdots & \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{1k} \\ \mathbf{\Sigma}_{31} & \mathbf{\Sigma}_{31}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12} & \mathbf{\Sigma}_{33} & \cdots & \mathbf{\Sigma}_{31}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{1k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Sigma}_{k1} & \mathbf{\Sigma}_{k1}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12} & \mathbf{\Sigma}_{k1}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{13} & \cdots & \mathbf{\Sigma}_{kk} \end{pmatrix}. \tag{2}$$

The case of $k = 3$ is considered in detail by Whittaker [24] and is called a "butterfly model". For general k, we called the model a *star-shape model* in [23] because the graphical shape of the relationship among the variables described by Whittaker [24] or Lauritzen [16] is like a star.

The above model is very popular in most areas, especially in economics. For example, let X_1 be the federal interest rate, which is a global variable, and $X_2, \ldots, X_{51}$ be the house price in each state, which are local variables. Then $X_2, \ldots, X_{51}$ are conditionally independent given X_1 because each house price $X_i,\ i = 2, \ldots, k$ will normally depend on its local situation if the federal interest rate is fixed.

The above star-shape model is the special case of the lattice conditional independence model introduced by Andersson and Perlman [2]. 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 [2] gave the form of the maximum likelihood estimator (MLE) of the covariance matrix $\mathbf{\Sigma}$. Konno [13] considered the estimation of the covariance matrix under the Stein loss

$$L_1(\hat{\mathbf{\Sigma}}, \mathbf{\Sigma}) \;=\; \mathrm{tr}(\hat{\mathbf{\Sigma}}\mathbf{\Sigma}^{-1}) - \log|\hat{\mathbf{\Sigma}}\mathbf{\Sigma}^{-1}| - p \tag{3}$$

and proved that the MLE of $\mathbf{\Sigma}$ is inadmissible. In fact, the Stein loss for estimating the covariance matrix is equivalent to the following loss for estimating the precision matrix $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$,

$$L_1^*(\hat{\mathbf{\Omega}}, \mathbf{\Omega}) \;=\; \mathrm{tr}(\hat{\mathbf{\Omega}}^{-1}\mathbf{\Omega}) - \log|\hat{\mathbf{\Omega}}^{-1}\mathbf{\Omega}| - p. \tag{4}$$

Of course, the Stein loss is related to the commonly used entropy loss. See [20]. Let $f(\mathbf{x} \mid \mathbf{\Sigma})$ be the density of $\mathbf{X}$ under $\mathbf{\Sigma}$. The entropy loss is obtained as follows,

$$\begin{aligned} L_2(\hat{\mathbf{\Sigma}}, \mathbf{\Sigma}) &= 2 \int \log \left\{ \frac{f(\mathbf{X} \mid \Sigma)}{f(\mathbf{X} \mid \hat{\mathbf{\Sigma}})} \right\} f(\mathbf{X} \mid \mathbf{\Sigma})\, d\mathbf{X} \\ &= \mathrm{tr}(\hat{\mathbf{\Sigma}}^{-1}\mathbf{\Sigma}) - \log |\hat{\mathbf{\Sigma}}^{-1}\mathbf{\Sigma}| - p. \end{aligned} \tag{5}$$

The Stein loss is obtained from the entropy loss by switching the role of two arguments, $\hat{\mathbf{\Omega}}$ and $\mathbf{\Omega}$. The loss function L_2 is typical entropy loss and has been studied by many authors such as Sinha and Ghosh [22], Krishnamoorthy and Gupta [14], and others.

Note that because neither L_1 nor L_2 is symmetric, we could consider a symmetric version by adding the Stein loss and entropy loss

$$L_3(\hat{\mathbf{\Sigma}}, \mathbf{\Sigma}) \;=\; L_1(\hat{\mathbf{\Sigma}}, \mathbf{\Sigma}) + L_2(\hat{\mathbf{\Sigma}}, \mathbf{\Sigma}) \;=\; \mathrm{tr}(\hat{\mathbf{\Sigma}}\mathbf{\Sigma}^{-1}) + \mathrm{tr}(\hat{\mathbf{\Sigma}}^{-1}\mathbf{\Sigma}) - 2p. \tag{6}$$

The symmetric loss L_3 was introduced by Kubokawa and Konno [15] and Gupta and Ofori-Nyarko [8]. It can be seen as estimating the covariance matrix and the precision matrix simultaneously.

For estimating the precision matrix $\mathbf{\Omega}$, the entropy loss and the symmetric loss will be

$$L_2^*(\hat{\mathbf{\Omega}}, \mathbf{\Omega}) \;=\; L_2(\hat{\mathbf{\Sigma}}, \mathbf{\Sigma}) \;=\; \mathrm{tr}(\hat{\mathbf{\Omega}}\mathbf{\Omega}^{-1}) - \log |\hat{\mathbf{\Omega}}\mathbf{\Omega}^{-1}| - p \tag{7}$$

and

$$L_3^*(\hat{\mathbf{\Omega}}, \mathbf{\Omega}) = L_3(\hat{\mathbf{\Sigma}}, \mathbf{\Sigma}) \;=\; \mathrm{tr}(\hat{\mathbf{\Omega}}\mathbf{\Omega}^{-1}) + \mathrm{tr}(\hat{\mathbf{\Omega}}^{-1}\mathbf{\Omega}) - 2p. \tag{8}$$

For convenience, we will still name L_1^*, L_2^*, L_3^* as the Stein loss, the entropy loss, and the symmetric loss for estimating the precision matrix $\mathbf{\Omega}$.

Sun and Sun [23] considered the estimating problems of the precision matrix under the entropy loss L_2^* and the symmetric loss L_3^* in the star-shape with complete observations. They obtained the closed forms of Bayesian estimators with respect to a class of priors of $\mathbf{\Psi}$. Consequently, the MLE of the precision matrix is inadmissible under either the entropy loss L_2^* or the symmetric loss L_3^*.

Considering that missing data problems occur frequently in practice and their analysis can be challenging, we will study the problem of estimating the covariance matrix and the precision matrix in a star-shape model with missing data. For estimating the covariance matrix without restriction, Anderson [1] listed several general cases, where the MLEs of the parameters can be obtained in closed form. Among these cases, the monotone missing-data pattern is most important. One also can see the related references by Little and Rubin [17, §1.3], Konno [12], Liu [18], Domonici et al. [5] and so on. However, because there are some restrictions on the covariance matrix in our model, we will see a lot of differences.

In this paper, we will consider the estimation of the covariance matrix and the precision matrix in a star-shape model with missing data, which generalizes some results in [23]. In Section 2, we first introduce the sample observations. By introducing a type of Cholesky decomposition of the precision matrix $\mathbf{\Omega} = \mathbf{\Psi}'\mathbf{\Psi}$, where $\mathbf{\Psi}$ is a lower triangular matrix with positive diagonal elements, the MLEs of the covariance matrix and the precision matrix are obtained, and it is proved that both of them are not unbiased. Based on the MLEs, unbiased estimates of the covariance matrix and the precision matrix are given. Considering that the parameter $\mathbf{\Psi}$ plays an important role in estimating the covariance matrix and the precision matrix, the special group $\mathcal{G}$, which is related to the decomposition, is introduced in Section 3. The invariant Haar measures of this group are given and the posterior properties of $\mathbf{\Psi}$ are discussed when choosing the left Haar measure as a prior. In Section 4, the closed form of the best equivariant estimator of the precision matrix is obtained under the Stein loss by using Bayesian method introduced by Eaton [6]. Consequently, the MLE of $\mathbf{\Omega}$ is inadmissible under the Stein loss. Results on the entropy loss and symmetric loss are shown in Sections 5 and 6. The results on estimating covariance matrix are given in Section 7. Some simulation results are given in Section 8. Finally, we give some concluding remarks.

2. MLEs and unbiased estimators

2.1. Sample observations

Now suppose that $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$ has the structure (1) and we got the following observations:

$$\begin{aligned} &\mathbf{Z}_{01}, \mathbf{Z}_{02}, \ldots, \mathbf{Z}_{0n} \sim N_p(\mathbf{0}, \mathbf{\Sigma}), \\ &\mathbf{Z}_{11}, \mathbf{Z}_{12}, \ldots, \mathbf{Z}_{1n_1} \sim N_{p_1}(\mathbf{0}, \mathbf{\Sigma}_{11}), \\ &\mathbf{Z}_{i1}, \mathbf{Z}_{i2}, \ldots, \mathbf{Z}_{in_i} \sim N_{p_1+p_i}\left(\mathbf{0}, \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{1i} \\ \mathbf{\Sigma}_{i1} & \mathbf{\Sigma}_{ii} \end{pmatrix}\right), \quad i = 2, \ldots, k. \end{aligned} \tag{9}$$

All $\mathbf{Z}_{ij}$s are independent. Let

$$\mathbf{V}_0 = \sum_{i=1}^{n} \mathbf{Z}_{0i}\mathbf{Z}_{0i}' \quad \text{and} \quad \mathbf{V}_i = \sum_{j=1}^{n_i} \mathbf{Z}_{ij}\mathbf{Z}_{ij}', \quad i = 1, \ldots, k. \tag{10}$$

Then $\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k$ are mutually independent and are sufficient statistics of $\mathbf{\Sigma}$ or $\mathbf{\Omega}$. Now write $\mathbf{V}_0 = (\mathbf{V}_{0ij})$, where $\mathbf{V}_{0ij}$ is a $p_i \times p_j$ submatrix and

$$\mathbf{V}_i = \begin{pmatrix} \mathbf{V}_{i11} & \mathbf{V}_{i12} \\ \mathbf{V}_{i21} & \mathbf{V}_{i22} \end{pmatrix}, \quad i = 2, \ldots, k, \tag{11}$$

where V_{i11} is a $p_1 \times p_1$ submatrix. Also let $\mathbf{V}_1 = \mathbf{V}_{111}$ for convenience. Assume that $n > p$, $n_1 > p_1$ and $n_i > p_1 + p_i, i = 2, \ldots, k$. Then, $\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k$ are all positive definite with probability one and

$$\begin{aligned} &\mathbf{V}_0 \sim W_p(n, \mathbf{\Sigma}), \quad \mathbf{V}_1 \sim W_{p_1}(n_1, \mathbf{\Sigma}_{11}), \\ &\mathbf{V}_i \sim W_{p_1+p_i}\left(n_i, \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{1i} \\ \mathbf{\Sigma}_{i1} & \mathbf{\Sigma}_{ii} \end{pmatrix}\right), \quad i = 2, \ldots, k, \end{aligned} \tag{12}$$

where $W_q(k, \mathbf{A})$ denotes a Wishart distribution with scale matrix $\mathbf{A}$ and degrees of freedom parameter k. We will estimate $\mathbf{\Sigma}$ and $\mathbf{\Omega}$ based on the sufficient statistics $\mathbf{V}_0, \mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_k$.

2.2. Cholesky decomposition

Usually, it is difficult to get appropriate estimators of the covariance matrix or the precision matrix with some restrictions. For example, if you want to estimate $\mathbf{\Omega}$ in (1) directly, you have to estimate $\mathbf{\Omega}_{11}, \mathbf{\Omega}_{22}, \ldots, \mathbf{\Omega}_{kk}$, $\mathbf{\Omega}_{12}, \ldots, \mathbf{\Omega}_{1k}$ first. However, this will not guarantee that the estimate of $\mathbf{\Omega}$ obtained in this way is positive definite. Now we will introduce the following Cholesky decomposition method to get MLEs of $\mathbf{\Sigma}$ and $\mathbf{\Omega}$ in a star-shape model. This method will guarantee the estimate of $\mathbf{\Omega}$ obtained is positive definite. In addition, we will see that this decomposition is still useful in getting the best equivariant estimates of $\mathbf{\Omega}$ or $\mathbf{\Sigma}$ under different loss functions later.

Let

$$\mathbf{\Omega} = \mathbf{\Psi}'\mathbf{\Psi} \quad \text{or} \quad \mathbf{\Sigma} = \mathbf{\Delta}\mathbf{\Delta}', \tag{13}$$

where both $\mathbf{\Psi}$ and $\mathbf{\Delta}$ are p by p lower-triangular matrices with positive diagonal entries. Thus, $\mathbf{\Psi} = \mathbf{\Delta}^{-1}$. For convenience, $\mathbf{\Delta}$ will be viewed as Cholesky decomposition of $\mathbf{\Sigma}$. From the structure of $\mathbf{\Omega}$ given by (1), it is easy to show that $\mathbf{\Psi}$ has the following block structure:

$$\mathbf{\Psi} = \begin{pmatrix} \mathbf{\Psi}_{11} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{\Psi}_{21} & \mathbf{\Psi}_{22} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{\Psi}_{31} & \mathbf{0} & \mathbf{\Psi}_{33} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Psi}_{k1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Psi}_{kk} \end{pmatrix}, \tag{14}$$

and thus

$$\mathbf{\Delta} = \mathbf{\Psi}^{-1} = \begin{pmatrix} \mathbf{\Psi}_{11}^{-1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ -\mathbf{\Psi}_{22}^{-1}\mathbf{\Psi}_{21}\mathbf{\Psi}_{11}^{-1} & \mathbf{\Psi}_{22}^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ -\mathbf{\Psi}_{33}^{-1}\mathbf{\Psi}_{31}\mathbf{\Psi}_{11}^{-1} & \mathbf{0} & \mathbf{\Psi}_{33}^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\mathbf{\Psi}_{kk}^{-1}\mathbf{\Psi}_{k1}\mathbf{\Psi}_{11}^{-1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Psi}_{kk}^{-1} \end{pmatrix}, \tag{15}$$

with $\mathbf{\Psi}_{ii}$ being p_i by p_i lower-triangular matrix, $i = 1, \ldots, k$. Note that there is no restriction on $\mathbf{\Psi}_{ij} (i \geqslant j)$ except requiring that all diagonal elements of $\mathbf{\Psi}_{ii}$ are positive. This good property enables us to estimate $\mathbf{\Psi}_{ij}$ first; then we can get the estimates of $\mathbf{\Sigma}$ and $\mathbf{\Omega}$ directly from the relationship between $\mathbf{\Sigma}$ (or $\mathbf{\Omega}$) and $\mathbf{\Psi}$. This method will ensure that the estimate of $\mathbf{\Sigma}$ (or $\mathbf{\Omega}$) obtained is positive definite if the estimates of the diagonal elements in each $\mathbf{\Psi}_{ii}$ are positive. Other properties of this decomposition will be discussed in the next section.

2.3. The maximum likelihood estimates

Whittaker [24] gives the expression of MLE of the covariance matrix $\mathbf{\Sigma}$ for $k = 3$ with complete observations. Sun and Sun [23] get the corresponding result for general k. We will generalize them to the star-shape model with missing observations. Let

$$\begin{aligned} \mathbf{W}_{11} &= \sum_{i=0}^{k} \mathbf{V}_{i11}, \\ \mathbf{W}_{i11} &= \mathbf{V}_{011} + \mathbf{V}_{i11}, \\ \mathbf{W}_{i1} &= \mathbf{W}_{1i}' = \mathbf{V}_{0i1} + \mathbf{V}_{i21}, \\ \mathbf{W}_{i22} &= \mathbf{V}_{0ii} + \mathbf{V}_{i22}, \\ \mathbf{W}_{ii\cdot 1} &= \mathbf{W}_{i22} - \mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i}, \quad i = 2, \ldots, k, \end{aligned} \tag{16}$$

and let $m_1 = n + \sum_{t=1}^{k} n_t$ and $m_i = n + n_i$, $i = 2, \ldots, k$ throughout this paper. Also, let $\mathbf{W}_{11\cdot 1} = \mathbf{W}_{11}$ for convenience.

Proposition 1. *Based on the incomplete data* $(\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)$ *in the star-shape model, the MLE* $\hat{\mathbf{\Sigma}}_M$ *of* $\mathbf{\Sigma}$ *is given as follows*:

$$\begin{aligned}
\hat{\mathbf{\Sigma}}_{11}^M &= \frac{\mathbf{W}_{11}}{m_1},\\
\hat{\mathbf{\Sigma}}_{i1}^M &= (\hat{\mathbf{\Sigma}}_{1i}^M)' \;=\; \frac{\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}\mathbf{W}_{11}}{m_1},\\
\hat{\mathbf{\Sigma}}_{ii}^M &= \frac{1}{m_i}\mathbf{W}_{ii\cdot 1} + \frac{1}{m_1}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}\mathbf{W}_{11}\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i}, \quad i = 2, \ldots, k,\\
\hat{\mathbf{\Sigma}}_{ij}^M &= \frac{1}{m_1}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}\mathbf{W}_{11}\mathbf{W}_{j11}^{-1}\mathbf{W}_{1j}, \quad 1 < i < j \leqslant k.
\end{aligned} \tag{17}$$

Proof. The likelihood function $f(\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k \mid \mathbf{\Psi})$ is proportional to

$$\begin{aligned}
&|\mathbf{V}_0|^{\frac{n-p-1}{2}}|\mathbf{\Sigma}|^{-\frac{n}{2}} etr\left\{-\frac{1}{2}\mathbf{\Sigma}^{-1}\mathbf{V}_0\right\}|\mathbf{V}_1|^{\frac{n_1-p_1-1}{2}}|\mathbf{\Sigma}_{11}|^{-\frac{n_1}{2}} etr\left\{-\frac{1}{2}\mathbf{\Sigma}_{11}^{-1}\mathbf{V}_1\right\}\\
&\times\prod_{i=2}^{k}|\mathbf{V}_i|^{\frac{n_i-p_1-p_i-1}{2}}\begin{vmatrix}\mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{1i}\\ \mathbf{\Sigma}_{i1} & \mathbf{\Sigma}_{ii}\end{vmatrix}^{-\frac{n_i}{2}} etr\left\{-\frac{1}{2}\begin{pmatrix}\mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{1i}\\ \mathbf{\Sigma}_{i1} & \mathbf{\Sigma}_{ii}\end{pmatrix}^{-1}\mathbf{V}_i\right\}\\
&\propto |\mathbf{\Psi}|^n|\mathbf{\Psi}_{11}|^{n_1}\prod_{i=2}^{k}\begin{vmatrix}\mathbf{\Psi}_{11} & \mathbf{0}\\ \mathbf{\Psi}_{i1} & \mathbf{\Psi}_{ii}\end{vmatrix}^{n_i}\cdot etr\left\{-\frac{1}{2}\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}'\right\} etr\left\{-\frac{1}{2}\mathbf{\Psi}_{11}\mathbf{V}_1\mathbf{\Psi}_{11}'\right\}\\
&\times\prod_{i=2}^{k} etr\left\{-\frac{1}{2}\begin{pmatrix}\mathbf{\Psi}_{11} & \mathbf{0}\\ \mathbf{\Psi}_{i1} & \mathbf{\Psi}_{ii}\end{pmatrix}\begin{pmatrix}\mathbf{V}_{i11} & \mathbf{V}_{i12}\\ \mathbf{V}_{i21} & \mathbf{V}_{i22}\end{pmatrix}\begin{pmatrix}\mathbf{\Psi}_{11} & \mathbf{0}\\ \mathbf{\Psi}_{i1} & \mathbf{\Psi}_{ii}\end{pmatrix}'\right\}\\
&= \prod_{i=1}^{k}|\mathbf{\Psi}_{ii}|^{m_i}\cdot\prod_{i=1}^{k} etr\left\{-\frac{1}{2}\mathbf{\Psi}_{ii}\mathbf{W}_{ii\cdot 1}\mathbf{\Psi}_{ii}'\right\}\\
&\times\prod_{i=2}^{k} etr\left\{-\frac{1}{2}(\mathbf{\Psi}_{i1} + \mathbf{\Psi}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1})\mathbf{W}_{i11}(\mathbf{\Psi}_{i1} + \mathbf{\Psi}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1})'\right\}\\
&\leqslant \prod_{i=1}^{k}|\mathbf{\Psi}_{ii}|^{m_i}\cdot\prod_{i=1}^{k} etr\left\{-\frac{1}{2}\mathbf{\Psi}_{ii}\mathbf{W}_{ii\cdot 1}\mathbf{\Psi}_{ii}'\right\}
\end{aligned} \tag{18}$$

where $etr(\mathbf{A}) = \exp(trace(\mathbf{A}))$. Hence, the MLE $\hat{\mathbf{\Psi}}$ of $\mathbf{\Psi}$ will be determined by

$$\begin{aligned}
\hat{\mathbf{\Psi}}_{ii}'\hat{\mathbf{\Psi}}_{ii} &= m_i\mathbf{W}_{ii\cdot 1}^{-1}, \quad i = 1, \ldots, k,\\
\hat{\mathbf{\Psi}}_{i1} &= -\hat{\mathbf{\Psi}}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}, \quad i = 2, \ldots, k,
\end{aligned} \tag{19}$$

and thus by (13) and (15) the MLE of $\mathbf{\Sigma}$ is obtained as described in the proposition. □

Under the conditions $n > p, n_1 > p_1$ and $n_i > p_1 + p_i,\ i = 2, \ldots, k$, the MLE $\hat{\mathbf{\Sigma}}_M$ is positive definite with probability one. In addition, by (19), the MLE $\hat{\mathbf{\Omega}}_M$ of the precision matrix $\mathbf{\Omega}$ can be straightforwardly obtained as follows:

Proposition 2. *Based on the incomplete data* $(\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)$ *in the star-shape model, the MLE* $\hat{\mathbf{\Omega}}_M$ *of* $\mathbf{\Omega}$ *is given by*

$$
\begin{aligned}
\hat{\mathbf{\Omega}}_{11}^M &= m_1 \mathbf{W}_{11}^{-1} + \sum_{i=2}^{k} m_i \mathbf{W}_{i11}^{-1} \mathbf{W}_{1i} \mathbf{W}_{ii\cdot 1}^{-1} \mathbf{W}_{i1} \mathbf{W}_{i11}^{-1}, \\
\hat{\mathbf{\Omega}}_{1i}^M &= (\hat{\mathbf{\Omega}}_{i1}^M)' \ = \ - m_i \mathbf{W}_{i11}^{-1} \mathbf{W}_{1i} \mathbf{W}_{ii\cdot 1}^{-1}, \\
\hat{\mathbf{\Omega}}_{ii}^M &= m_i \mathbf{W}_{ii\cdot 1}^{-1}, \quad i = 2, \ldots, k.
\end{aligned} \tag{20}
$$

The MLE $\hat{\mathbf{\Omega}}_M$ of the precision matrix $\mathbf{\Omega}$ also can be obtained by the following relationships between $\mathbf{\Omega}$ and $\mathbf{\Sigma}$,

$$
\begin{aligned}
\mathbf{\Omega}_{11} &= \mathbf{\Sigma}_{11}^{-1} + \sum_{i=2}^{k} \mathbf{\Sigma}_{11}^{-1} \mathbf{\Sigma}_{1i} \mathbf{\Sigma}_{ii\cdot 1}^{-1} \mathbf{\Sigma}_{i1} \mathbf{\Sigma}_{11}^{-1}, \\
\mathbf{\Omega}_{1i} &= -\mathbf{\Sigma}_{11}^{-1} \mathbf{\Sigma}_{1i} \mathbf{\Sigma}_{ii\cdot 1}^{-1}, \\
\mathbf{\Omega}_{ii} &= \mathbf{\Sigma}_{ii\cdot 1}^{-1}, \quad i = 2, \ldots, k,
\end{aligned} \tag{21}
$$

where

$$
\mathbf{\Sigma}_{ii\cdot 1} = \mathbf{\Sigma}_{ii} - \mathbf{\Sigma}_{i1} \mathbf{\Sigma}_{11}^{-1} \mathbf{\Sigma}_{1i}, \quad i = 2, \ldots, k.
$$

Remark 1. For a star-shape model with missing data, the MLE $\hat{\mathbf{\Sigma}}_M$ is no longer a minimal sufficient statistic for $\mathbf{\Sigma}$, which is different from the case with complete observations in [23]. In fact, $\mathbf{W}_{11}, \mathbf{W}_{211}, \ldots, \mathbf{W}_{k11}, \mathbf{W}_{21}, \ldots, \mathbf{W}_{k1}, \mathbf{W}_{22\cdot 1}, \ldots, \mathbf{W}_{kk\cdot 1}$ are minimal statistics of $\mathbf{\Sigma}$, which can be shown by the likelihood function in (18).

Sun and Sun [23] showed that for a star-shape model with complete observations, the MLE $\hat{\mathbf{\Sigma}}_M$ of the covariance matrix $\mathbf{\Sigma}$ is unbiased while $\hat{\mathbf{\Omega}}_M$ is biased. However, the following proposition shows that for a missing case, neither $\hat{\mathbf{\Sigma}}_M$ is unbiased for $\mathbf{\Sigma}$, nor $\hat{\mathbf{\Omega}}_M$ is unbiased for $\mathbf{\Omega}$.

Proposition 3. *Consider a star-shape model with missing data.*

(a) *The MLE* $\hat{\mathbf{\Sigma}}_M$ *in* (17) *is not an unbiased estimate of* $\mathbf{\Sigma}$.
(b) *The MLE* $\hat{\mathbf{\Omega}}_M$ in (20) *is not an unbiased estimator of* $\mathbf{\Omega}$.

Proof. (a) In fact, we will show $\mathbb{E}(\hat{\mathbf{\Sigma}}_{ii}^M) \neq \mathbf{\Sigma}_{ii}$, $i = 2, \ldots, k$. By Proposition 1,

$$
\begin{aligned}
\hat{\mathbf{\Sigma}}_{ii}^M &= \frac{\mathbf{W}_{i22} - \mathbf{W}_{i1} \mathbf{W}_{i11}^{-1} \mathbf{W}_{1i}}{m_i} + \frac{\mathbf{W}_{i1} \mathbf{W}_{i11}^{-1} \mathbf{W}_{11} \mathbf{W}_{i11}^{-1} \mathbf{W}_{1i}}{m_1} \\
&= \frac{\mathbf{W}_{i22}}{m_i} + \Big(\frac{1}{m_1} - \frac{1}{m_i}\Big) \mathbf{W}_{i1} \mathbf{W}_{i11}^{-1} \mathbf{W}_{1i} + \frac{\mathbf{W}_{i1} \mathbf{W}_{i11}^{-1} (\sum_{t=1}^{k} \mathbf{V}_{t11} - \mathbf{V}_{i11}) \mathbf{W}_{i11}^{-1} \mathbf{W}_{1i}}{m_1}.
\end{aligned}
$$

Obviously, from (12), $\mathbb{E}(\mathbf{W}_{i22}) = \mathbb{E}(\mathbf{V}_{0ii} + \mathbf{V}_{i22}) = m_i \mathbf{\Sigma}_{ii}$, $i = 2, \ldots, k$ and

$$
\mathbb{E}\left\{ \mathbf{W}_{i1} \mathbf{W}_{i11}^{-1} \left(\sum_{t=1}^{k} \mathbf{V}_{t11} - \mathbf{V}_{i11} \right) \mathbf{W}_{i11}^{-1} \mathbf{W}_{1i} \right\} = (m_1 - m_i) \mathbb{E}(\mathbf{W}_{i1} \mathbf{W}_{i11}^{-1} \mathbf{\Sigma}_{11} \mathbf{W}_{i11}^{-1} \mathbf{W}_{1i}).
$$

Because

$$
\begin{pmatrix} \mathbf{V}_{011} & \mathbf{V}_{01i} \\ \mathbf{V}_{0i1} & \mathbf{V}_{0ii} \end{pmatrix} \sim W_{p_1+p_i}\left(n, \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{1i} \\ \mathbf{\Sigma}_{i1} & \mathbf{\Sigma}_{ii} \end{pmatrix}\right),
$$

it follows

$$\mathbf{V}_{0i1} \mid \mathbf{V}_{011} \sim N_{p_i, p_1}(\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{V}_{011}, \boldsymbol{\Sigma}_{ii\cdot 1} \otimes \mathbf{V}_{011}). \quad (22)$$

Similarly,

$$\mathbf{V}_{i12} \mid \mathbf{V}_{i11} \sim N_{p_i, p_1}(\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{V}_{i11}, \boldsymbol{\Sigma}_{ii\cdot 1} \otimes \mathbf{V}_{i11}), \quad (23)$$

and thus we have

$$(\mathbf{V}_{0i1} + \mathbf{V}_{i12}) \mid (\mathbf{V}_{011}, \mathbf{V}_{i11}) \sim N_{p_i, p_1}(\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{V}_{011} + \mathbf{V}_{i11}), \boldsymbol{\Sigma}_{ii\cdot 1} \otimes (\mathbf{V}_{011} + \mathbf{V}_{i11})). \quad (24)$$

So,

$$\begin{aligned}\mathbb{E}(\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i}) &= \mathbb{E}\{\mathbb{E}(\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i} \mid (\mathbf{V}_{011}, \mathbf{V}_{i11}))\} \\ &= \mathbb{E}\{\mathrm{tr}(\mathbf{W}_{i11}^{-1}\mathbf{W}_{i11})\boldsymbol{\Sigma}_{ii\cdot 1}\} + \mathbb{E}(\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{W}_{i11}\mathbf{W}_{i11}^{-1}\mathbf{W}_{i11}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}) \\ &= \mathrm{tr}(\mathbf{I}_{p_1})\boldsymbol{\Sigma}_{ii\cdot 1} + \boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}\mathbb{E}(\mathbf{W}_{i11})\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i} \\ &= p_1\boldsymbol{\Sigma}_{ii\cdot 1} + m_i\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}(\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}\boldsymbol{\Sigma}_{11}\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i}) \\ &\quad = \mathbb{E}\{\mathbb{E}(\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}\boldsymbol{\Sigma}_{11}\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i} \mid (\mathbf{V}_{011}, \mathbf{V}_{i11}))\} \\ &\quad = \mathbb{E}\{\mathrm{tr}(\mathbf{W}_{i11}^{-1}\boldsymbol{\Sigma}_{11}\mathbf{W}_{i11}^{-1}\mathbf{W}_{i11})\boldsymbol{\Sigma}_{ii\cdot 1}\} + \mathbb{E}(\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{W}_{i11}\mathbf{W}_{i11}^{-1}\boldsymbol{\Sigma}_{11}\mathbf{W}_{i11}^{-1}\mathbf{W}_{i11}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}) \\ &\quad = \mathrm{tr}\{\mathbb{E}(\mathbf{V}_{011} + \mathbf{V}_{i11})^{-1}\boldsymbol{\Sigma}_{11}\}\boldsymbol{\Sigma}_{ii\cdot 1} + \boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i} \\ &\quad = \frac{p_1}{m_i - p_1 - 1}\boldsymbol{\Sigma}_{ii\cdot 1} + \boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}.\end{aligned}$$

Let

$$r_i = 1 + p_1\left\{\frac{1}{m_1} - \frac{1}{m_i} + \frac{m_1 - m_i}{m_1(m_i - p_1 - 1)}\right\},$$

then

$$\begin{aligned}\mathbb{E}(\hat{\boldsymbol{\Sigma}}_{ii}^M) &= \frac{\mathbb{E}(\mathbf{W}_{i22})}{m_i} + \left(\frac{1}{m_1} - \frac{1}{m_i}\right)\mathbb{E}(\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i}) \\ &\quad + \frac{m_1 - m_i}{m_1}\,\mathbb{E}(\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}\boldsymbol{\Sigma}_{11}\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i}) \\ &= r_i\boldsymbol{\Sigma}_{ii} + r_i\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i},\end{aligned}$$

which is not equal to $\boldsymbol{\Sigma}_{ii}$.

(b) Because

$$\begin{pmatrix}\mathbf{V}_{011} + \mathbf{V}_{i11} & \mathbf{V}_{01i} + \mathbf{V}_{i12} \\ \mathbf{V}_{0i1} + \mathbf{V}_{i21} & \mathbf{V}_{0ii} + \mathbf{V}_{i22}\end{pmatrix} \sim W_{p_1+p_i}\left(n + n_i, \begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{1i} \\ \boldsymbol{\Sigma}_{i1} & \boldsymbol{\Sigma}_{ii}\end{pmatrix}\right),$$

we have $\mathbf{W}_{ii\cdot 1} \sim W_{p_i}(m_i - p_1, \boldsymbol{\Sigma}_{ii\cdot 1})$, and thus

$$\mathbb{E}(\hat{\boldsymbol{\Omega}}_{ii}^M) = \frac{m_i}{m_i - p_1 - p_i - 1}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1} \neq \boldsymbol{\Omega}_{ii}, \quad i = 2, \ldots, k,$$

which proves the second part. □

2.4. Unbiased estimators

Based on $\hat{\mathbf{\Sigma}}_M, \hat{\mathbf{\Omega}}_M$, we create unbiased estimates of $\mathbf{\Sigma}$ and $\mathbf{\Omega}$, respectively.

Proposition 4. *Consider a star-shape model with missing data.*

(a) *An unbiased estimate $\hat{\mathbf{\Sigma}}_U$ of $\mathbf{\Sigma}$ is given by*

$$\hat{\mathbf{\Sigma}}_{ii}^U = \left[1 - \frac{p_1(m_1 - p_1 - 1)}{m_1(m_i - p_1 - 1)}\right]\frac{\mathbf{W}_{ii\cdot 1}}{m_i - p_1} + \frac{\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}\mathbf{W}_{11}\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i}}{m_1}, \quad i = 2, \dots, k$$

and $\hat{\mathbf{\Sigma}}_{ij}^U = \hat{\mathbf{\Sigma}}_{ij}^M$ for other i, j, where $\hat{\mathbf{\Sigma}}_{ij}^M$ is shown by (17) *in Proposition* 1.

(b) *An unbiased estimate $\hat{\mathbf{\Omega}}_U$ of $\mathbf{\Omega}$ is given by*

$$\hat{\mathbf{\Omega}}_{11}^U = (m_1 - p_1 - 1)\Big(1 - \sum_{i=2}^{k} \frac{p_i}{m_i - p_1 - 1}\Big)\mathbf{W}_{11}^{-1} + \sum_{i=2}^{k}(m_i - p_1 - p_i - 1)\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i}\mathbf{W}_{ii\cdot 1}^{-1}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1},$$

and for $i = 2, \dots, k$,

$$\hat{\mathbf{\Omega}}_{1i}^U = -(m_i - p_1 - p_i - 1)\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i}\mathbf{W}_{ii\cdot 1}^{-1},$$
$$\hat{\mathbf{\Omega}}_{ii}^U = (m_i - p_1 - p_i - 1)\mathbf{W}_{ii\cdot 1}^{-1}.$$

Proof. (a) Obviously, $\mathbb{E}(\hat{\mathbf{\Sigma}}_{11}^M) = \mathbb{E}(\mathbf{V}_{011} + \sum_{i=1}^k \mathbf{V}_{i11})/m_1 = \mathbf{\Sigma}_{11}$. By (24),

$$\mathbb{E}(\mathbf{V}_{0i1} + \mathbf{V}_{i21} \mid \mathbf{V}_{011}, \mathbf{V}_{i11}) = \mathbf{\Sigma}_{i1}\mathbf{\Sigma}_{11}^{-1}(\mathbf{V}_{011} + \mathbf{V}_{i11})$$

and thus $\mathbb{E}(\hat{\mathbf{\Sigma}}_{i1}^M) = \mathbf{\Sigma}_{i1}$ because of $\mathbb{E}(\mathbf{V}_{011} + \mathbf{V}_{i11}) = m_i\mathbf{\Sigma}_{11}$, $i = 2, \dots, k$. In addition, for any $1 < i < j \leqslant k$,

$$\begin{aligned}
\begin{pmatrix}\mathbf{V}_{0i1}\\ \mathbf{V}_{0j1}\end{pmatrix} \mid \mathbf{V}_{011} &\sim N_{p_i+p_j,p_1}\left(\begin{pmatrix}\mathbf{\Sigma}_{i1}\\ \mathbf{\Sigma}_{j1}\end{pmatrix}\mathbf{\Sigma}_{11}^{-1}\mathbf{V}_{011}, \begin{pmatrix}\mathbf{\Sigma}_{ii\cdot 1} & \mathbf{0}\\ \mathbf{0} & \mathbf{\Sigma}_{jj\cdot}\end{pmatrix}\bigotimes \mathbf{V}_{011}\right),\\
\mathbf{V}_{i21} \mid \mathbf{V}_{i11} &\sim N_{p_i,p_1}(\mathbf{\Sigma}_{i1}\mathbf{\Sigma}_{11}^{-1}\mathbf{V}_{i11}, \mathbf{\Sigma}_{ii\cdot 1} \otimes \mathbf{V}_{i11}),\\
\mathbf{V}_{j21} \mid \mathbf{V}_{j11} &\sim N_{p_j,p_1}(\mathbf{\Sigma}_{j1}\mathbf{\Sigma}_{11}^{-1}\mathbf{V}_{j11}, \mathbf{\Sigma}_{jj\cdot} \otimes \mathbf{V}_{j11}),
\end{aligned} \tag{25}$$

then we get

$$\begin{aligned}
&\begin{pmatrix}\mathbf{V}_{0i1} + \mathbf{V}_{i21}\\ \mathbf{V}_{0j1} + \mathbf{V}_{j21}\end{pmatrix} \mid (\mathbf{V}_{011}, \mathbf{V}_{i11}, \mathbf{V}_{j11})\\
&\quad \sim N_{p_i+p_j,p_1}\left(\begin{pmatrix}\mathbf{\Sigma}_{i1}\mathbf{\Sigma}_{11}^{-1}(\mathbf{V}_{011} + \mathbf{V}_{i11})\\ \mathbf{\Sigma}_{j1}\mathbf{\Sigma}_{11}^{-1}(\mathbf{V}_{011} + \mathbf{V}_{j11})\end{pmatrix},\right.\\
&\qquad \left.\begin{pmatrix}\mathbf{\Sigma}_{ii\cdot 1}\bigotimes(\mathbf{V}_{011} + \mathbf{V}_{i11}) & \mathbf{0}\\ \mathbf{0} & \mathbf{\Sigma}_{jj\cdot}\bigotimes(\mathbf{V}_{011} + \mathbf{V}_{j11})\end{pmatrix}\right).
\end{aligned} \tag{26}$$

Also, by $\mathbb{E}(\mathbf{W}_{11}) = m_1\mathbf{\Sigma}_{11}$, we can easily obtain that $\mathbb{E}(\hat{\mathbf{\Sigma}}_{ij}^M) = \mathbf{\Sigma}_{i1}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{1j} = \mathbf{\Sigma}_{ij}$. In addition, similar to (25), we can easily see that $\mathbb{E}(\hat{\mathbf{\Sigma}}_{ii}^U) = \mathbf{\Sigma}_{ii}$, $i = 2, \dots, k$ and part (a) is proved.

(b) By (24),

$$\mathbf{W}_{i1} \mid \mathbf{W}_{i11} \sim N_{p_i,p_1}(\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}\mathbf{W}_{i11}, \boldsymbol{\Sigma}_{ii\cdot 1} \otimes \mathbf{W}_{i11}), \tag{27}$$

and $(\mathbf{W}_{i1}, \mathbf{W}_{i11})$ is independent of $\mathbf{W}_{ii\cdot 1}$. Therefore,

$$\begin{aligned}\mathbb{E}(\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i}\mathbf{W}_{ii\cdot 1}^{-1} \mid \mathbf{W}_{i11}) &= \mathbf{W}_{i11}^{-1}\mathbb{E}(\mathbf{W}_{1i} \mid \mathbf{W}_{i11})\mathbb{E}(\mathbf{W}_{ii\cdot 1}^{-1}) \\ &= \frac{1}{m_i - p_1 - p_i - 1}\mathbf{W}_{i11}^{-1} \cdot \mathbf{W}_{i11}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i} \cdot \boldsymbol{\Sigma}_{ii\cdot 1}^{-1} \\ &= \frac{1}{m_i - p_1 - p_i - 1}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1}\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}(\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i}\mathbf{W}_{ii\cdot 1}^{-1}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1} \mid \mathbf{W}_{i11}) \\ &\quad = \frac{1}{m_i - p_1 - p_i - 1}\mathbf{W}_{i11}^{-1}\mathbb{E}(\mathbf{W}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1}\mathbf{W}_{i1} \mid \mathbf{W}_{i11})\mathbf{W}_{i11}^{-1} \\ &\quad = \frac{p_i}{m_i - p_1 - p_i - 1}\mathbf{W}_{i11}^{-1} + \frac{1}{m_i - p_1 - p_i - 1}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1}\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1}.\end{aligned}$$

So we get

$$\begin{aligned}\mathbb{E}(\hat{\boldsymbol{\Omega}}_{11}^{U}) &= (m_1 - p_1 - 1)\left(1 - \sum_{i=2}^{k}\frac{p_i}{m_i - p_1 - 1}\right)\mathbb{E}(\mathbf{W}_{11}^{-1}) \\ &\quad + \sum_{i=2}^{k}\frac{p_i}{m_i - p_1 - 1}\boldsymbol{\Sigma}_{11}^{-1} + \sum_{i=2}^{k}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1}\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1} \\ &= \boldsymbol{\Sigma}_{11}^{-1} + \sum_{i=2}^{k}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1}\boldsymbol{\Sigma}_{i1}\boldsymbol{\Sigma}_{11}^{-1} = \boldsymbol{\Omega}_{11}, \\ \mathbb{E}(\hat{\boldsymbol{\Omega}}_{1i}^{U}) &= -(m_i - p_1 - p_i - 1)\mathbb{E}(\mathbf{W}_{i11}^{-1}\mathbf{W}_{1i}\mathbf{W}_{ii\cdot 1}^{-1}) = -\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{1i}\boldsymbol{\Sigma}_{ii\cdot 1}^{-1} = \boldsymbol{\Omega}_{1i}, \\ \mathbb{E}(\hat{\boldsymbol{\Omega}}_{ii}^{U}) &= (m_i - p_1 - p_i - 1)\mathbb{E}(\mathbf{W}_{ii\cdot 1}^{-1}) = \boldsymbol{\Sigma}_{ii\cdot 1}^{-1} = \boldsymbol{\Omega}_{ii}, \quad i = 2, \ldots, k.\end{aligned}$$

The proof is completed. □

3. The invariant Haar measures

Group invariance plays an important role in finding better estimates of the covariance and precision matrices in multivariate normal distribution. See for example, James and Stein [10], Olkin and Selliah [19], Sharma and Krishnamoorthy [21], Konno [13], Sun and Sun [23] etc. Define

$$\mathcal{G} = \left\{\mathbf{A} \in R^{p\times p} \mid \mathbf{A} \text{ has a structure as } (14)\right\}. \tag{28}$$

Sun and Sun [23] showed that $\mathcal{G}$ is a group with respect to matrix multiplication. For any $i = 1, 2, \ldots, k$, let

$$\boldsymbol{\Psi}_{ii} = \begin{pmatrix} \psi_{i11} & 0 & \cdots & 0 \\ \psi_{i21} & \psi_{i22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{ip_i1} & \psi_{ip_i1} & \cdots & \psi_{ip_ip_i} \end{pmatrix}. \tag{29}$$

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

$$\mathbf{\Psi}_{i1} = \begin{pmatrix} \phi_{i11} & \phi_{i12} & \cdots & \phi_{i1p_1} \\ \phi_{i21} & \phi_{i22} & \cdots & \phi_{i2p_1} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{ip_i1} & \phi_{ip_i2} & \cdots & \phi_{ip_ip_1} \end{pmatrix}. \tag{30}$$

Similar to Example 1.14 of [7], the left invariant Haar measure and the right invariant Haar measure of $\mathcal{G}$ are given by

$$\nu^l_{\mathcal{G}}(d\mathbf{\Psi}) \propto \frac{d\mathbf{\Psi}}{\prod_{j=1}^{p_1} \psi_{1jj}^{j} \cdot \prod_{i=2}^{k}\prod_{j=1}^{p_i} \psi_{ijj}^{p_1+j}}, \tag{31}$$

$$\nu^r_{\mathcal{G}}(d\mathbf{\Psi}) \propto \frac{d\mathbf{\Psi}}{\prod_{j=1}^{p_1} \psi_{1jj}^{p-j+1} \cdot \prod_{i=2}^{k}\prod_{j=1}^{p_i} \psi_{ijj}^{p_i-j+1}}, \tag{32}$$

respectively. In addition, we can readily verify that $\nu^r_{\mathcal{G}}(d\mathbf{\Psi}) = \nu^l_{\mathcal{G}}(d\mathbf{\Delta})$ and $\nu^l_{\mathcal{G}}(d\mathbf{\Psi}) = \nu^r_{\mathcal{G}}(d\mathbf{\Delta})$ because $\mathbf{\Delta} = \mathbf{\Psi}^{-1}$.

However, unlike the case with complete data in [23], it seems impossible to get the closed form of equivariant estimators of $\mathbf{\Sigma}$ or $\mathbf{\Omega}$ with respect to $\mathcal{G}$ in a star-shape model with missing data. So it is impossible to obtain the best equivariant estimate of $\mathbf{\Omega}$ (or $\mathbf{\Sigma}$) under an invariant loss. Like [6], the Bayesian method will be applied to get the best equivariant estimates of $\mathbf{\Omega}$ (or $\mathbf{\Sigma}$) in the next few sections.

Let $\mathbf{T}_{ii}$ be the Cholesky decomposition of $\mathbf{W}_{ii\cdot 1}$, $i = 1, 2, \ldots, k$.

Theorem 1. *If we take the left invariant Haar measure of the group $\mathcal{G}$, $\nu^l_{\mathcal{G}}(d\mathbf{\Psi})$ as a prior, then the posterior distribution of $\mathbf{\Psi}$ in a star-shape model with missing data satisfies*

(a) $\mathbf{\Psi}_{11}, (\mathbf{\Psi}_{21}, \mathbf{\Psi}_{22}), \ldots, (\mathbf{\Psi}_{k1}, \mathbf{\Psi}_{kk})$ *are mutually independent*;
(b) *For* $i = 2, \ldots, k$, $\mathbf{\Psi}_{i1} \mid (\mathbf{\Psi}_{ii}, \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k) \sim N_{p_i,p_1}(-\mathbf{\Psi}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}, \mathbf{I}_{p_i} \otimes \mathbf{W}_{i11}^{-1})$;
(c) $\mathbf{\Psi}_{11}\mathbf{W}_{11}\mathbf{\Psi}'_{11} \mid (\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k) \sim W_{p_1}(m_1, \mathbf{I}_{p_1})$ *and*

$$\mathbb{E}[\mathbf{\Psi}'_{11}\mathbf{\Psi}_{11} \mid (\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)] = (\mathbf{T}'_{11})^{-1} diag(\delta_{11}, \ldots, \delta_{1p_1})\mathbf{T}_{11}^{-1}, \tag{33}$$

$$\mathbb{E}[(\mathbf{\Psi}'_{11}\mathbf{\Psi}_{11})^{-1} \mid (\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)] = \mathbf{T}_{11} diag(\eta_{11}, \ldots, \eta_{1p_1})\mathbf{T}'_{11}, \tag{34}$$

where

$$\delta_{1j} = m_1 + p_1 - 2j + 1, \quad j = 1, \ldots, p_1, \tag{35}$$

$$\eta_{1j} = \frac{m_1 - 1}{(m_1 - j - 1)(m_1 - j)}, \quad j = 2, \ldots, p_1; \tag{36}$$

(d) *For* $i = 2, \ldots, k$, $\mathbf{\Psi}_{ii}\mathbf{W}_{ii\cdot 1}\mathbf{\Psi}'_{ii} \mid (\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k) \sim W_{p_i}(m_i - p_1, \mathbf{I}_{p_i})$ *and*

$$\mathbb{E}[\mathbf{\Psi}'_{ii}\mathbf{\Psi}_{ii} \mid (\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)] = (\mathbf{T}'_{ii})^{-1} diag(\delta_{i1}, \ldots, \delta_{ip_i})\mathbf{T}_{ii}^{-1}, \tag{37}$$

$$\mathbb{E}[(\mathbf{\Psi}'_{ii}\mathbf{\Psi}_{ii})^{-1} \mid (\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)] = \mathbf{T}_{ii} diag(\eta_{i1}, \ldots, \eta_{ip_i})\mathbf{T}'_{ii}, \tag{38}$$

where

$$\delta_{ij} = m_i - p_1 + p_i - 2j + 1, \quad j = 1, \ldots, p_i, \tag{39}$$

$$\eta_{ij} = \frac{m_i - 1}{(m_i - p_1 - j - 1)(m_i - p_1 - j)}, \quad j = 1, \ldots, p_i. \tag{40}$$

Proof. Combining the likelihood function of $\boldsymbol{\Psi}$ in (18) with the prior of $\boldsymbol{\Psi}$ in (31), we have the posterior of $\boldsymbol{\Psi}$,

$$\begin{aligned} &p(\boldsymbol{\Psi} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k) \\ &\quad\propto \prod_{i=2}^{k} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{(\boldsymbol{\Psi}_{i1} + \boldsymbol{\Psi}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1})\mathbf{W}_{i11}(\boldsymbol{\Psi}_{i1} + \boldsymbol{\Psi}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1})'\right\}\right] \\ &\qquad\times \exp\left\{-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Psi}_{11}\mathbf{W}_{11\cdot 1}\boldsymbol{\Psi}_{11}')\right\} \prod_{j=1}^{p_1} \psi_{1jj}^{m_1-j} \\ &\qquad\times \prod_{i=2}^{k} \exp\left\{-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Psi}_{ii}\mathbf{W}_{ii\cdot 1}\boldsymbol{\Psi}_{ii}')\right\} \prod_{j=1}^{p_i} \psi_{ijj}^{m_i-p_1-j}. \end{aligned} \tag{41}$$

Thus, we prove parts (a) and (b). For part (c), we have the posterior marginal of $\boldsymbol{\Psi}_{11}$

$$p(\boldsymbol{\Psi}_{11} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k) \propto \exp\left\{-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Psi}_{11}\mathbf{W}_{11\cdot 1}\boldsymbol{\Psi}_{11}')\right\} \prod_{j=1}^{p_i} \psi_{1jj}^{m_1-j}$$

and thus $\boldsymbol{\Psi}_{11}\mathbf{W}_{11}\boldsymbol{\Psi}_{11}' \mid (\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k) \sim W_{p_1}(m_1, \mathbf{I}_{p_1})$. Eqs. (33) and (34) are the special cases of Lemma 5.1 in [23]. Similarly, we can prove part (d). □

4. Best equivariant estimator of $\boldsymbol{\Omega}$ under the Stein loss

For estimating the covariance matrix $\boldsymbol{\Sigma}$, Eaton [7] showed that under some conditions, the best equivariant estimate of $\boldsymbol{\Sigma}$ under the group $\mathcal{G}$ will be a Bayesian estimate if we take the right Haar measure $v_{\mathcal{G}}^r(d\boldsymbol{\Delta})$ as a prior. So for estimating the precision matrix $\boldsymbol{\Omega}$, the best equivariant estimate of $\boldsymbol{\Omega}$ under the group $\mathcal{G}$ will be a Bayesian estimate if we take the left Haar measure $v_{\mathcal{G}}^l(d\boldsymbol{\Psi})$ as a prior because $v_{\mathcal{G}}^l(d\boldsymbol{\Psi}) = v_{\mathcal{G}}^r(d\boldsymbol{\Delta})$.

Now define

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_{11} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{W}_{21}\mathbf{W}_{211}^{-1}\mathbf{T}_{11} & \mathbf{T}_{22} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{W}_{31}\mathbf{W}_{311}^{-1}\mathbf{T}_{11} & \mathbf{0} & \mathbf{T}_{33} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{W}_{k1}\mathbf{W}_{k11}^{-1}\mathbf{T}_{11} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{T}_{kk} \end{pmatrix}. \tag{42}$$

Then we have

$$\mathbf{R} = \mathbf{T}^{-1} = \begin{pmatrix} \mathbf{T}_{11}^{-1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ -\mathbf{T}_{22}^{-1}\mathbf{W}_{21}\mathbf{W}_{211}^{-1} & \mathbf{T}_{22}^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ -\mathbf{T}_{33}^{-1}\mathbf{W}_{31}\mathbf{W}_{311}^{-1} & \mathbf{0} & \mathbf{T}_{33}^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\mathbf{T}_{kk}^{-1}\mathbf{W}_{k1}\mathbf{W}_{k11}^{-1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{T}_{kk}^{-1} \end{pmatrix}. \tag{43}$$

Theorem 2. *Under the Stein loss L_1^*, the best $\mathcal{G}$-equivariant estimator of $\boldsymbol{\Omega}$ is given by*

$$\hat{\boldsymbol{\Omega}}_{1B} = \mathbf{R}'\mathbf{B}_{1B}\mathbf{R}, \tag{44}$$

where $\mathbf{R}$ *is given by* (43), $\mathbf{B}_{1B} = diag(\mathbf{B}_{11B}, \mathbf{B}_{12B}, \ldots, \mathbf{B}_{1kB})$, $\mathbf{B}_{11B} = \mathbf{D}_1 + \sum_{i=2}^k p_i \mathbf{T}'_{11} \mathbf{W}_{i11}^{-1}\mathbf{T}_{11}$ *and* $\mathbf{D}_1 = diag(d_{11}, \ldots, d_{1p_1})$ *with* $d_{1j} = m_1 + p_1 - 2j + 1, j = 1, \ldots, p_1$; $\mathbf{B}_{1iB} = diag(b_{i1B}, \ldots, b_{ip_iB})$ *with* $b_{ijB} = m_i - p_1 + p_i - 2j + 1, j = 1, \ldots, p_i, i = 2, \ldots, k$.

Proof. The best equivariant estimate of $\mathbf{\Omega}$ under the Stein loss L_1^* will be produced by minimizing the posterior risk

$$b_1(\hat{\mathbf{\Omega}}) = \int \left[\text{tr}\left\{\hat{\mathbf{\Omega}}^{-1}(\mathbf{\Psi}'\mathbf{\Psi})\right\} - \log|\hat{\mathbf{\Omega}}^{-1}(\mathbf{\Psi}'\mathbf{\Psi})| - p\right] f(\mathbf{\Psi} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)\, d\mathbf{\Psi},$$

where $f(\mathbf{\Psi} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)$ is described in Theorem 1. Letting $\hat{\mathbf{\Omega}} = \hat{\mathbf{\Psi}}'\hat{\mathbf{\Psi}}$, where $\hat{\mathbf{\Psi}} \in \mathcal{G}$ and has the similar block partition as in (14). Thus, the question becomes to minimize

$$g_1(\hat{\mathbf{\Psi}}) = \int \text{tr}\{(\mathbf{\Psi}\hat{\mathbf{\Psi}}^{-1})(\mathbf{\Psi}\hat{\mathbf{\Psi}}^{-1})'\} f(\mathbf{\Psi} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)\, d\mathbf{\Psi} - \log|(\hat{\mathbf{\Psi}}'\hat{\mathbf{\Psi}})^{-1}|. \quad (45)$$

So we need to calculate the posterior expectation of tr $\left\{(\mathbf{\Psi}\hat{\mathbf{\Psi}}^{-1})(\mathbf{\Psi}\hat{\mathbf{\Psi}}^{-1})'\right\}$. Because

$$\mathbf{\Psi}\hat{\mathbf{\Psi}}^{-1} = \begin{pmatrix} \mathbf{\Psi}_{11}\hat{\mathbf{\Psi}}_{11}^{-1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ (\mathbf{\Psi}_{21} - \mathbf{\Psi}_{22}\hat{\mathbf{\Psi}}_{22}^{-1}\hat{\mathbf{\Psi}}_{21})\hat{\mathbf{\Psi}}_{11}^{-1} & \mathbf{\Psi}_{22}\hat{\mathbf{\Psi}}_{22}^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ (\mathbf{\Psi}_{31} - \mathbf{\Psi}_{33}\hat{\mathbf{\Psi}}_{33}^{-1}\hat{\mathbf{\Psi}}_{31})\hat{\mathbf{\Psi}}_{11}^{-1} & \mathbf{0} & \hat{\mathbf{\Psi}}_{33}\mathbf{\Psi}_{33}^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (\mathbf{\Psi}_{k1} - \mathbf{\Psi}_{kk}\hat{\mathbf{\Psi}}_{kk}^{-1}\hat{\mathbf{\Psi}}_{k1})\hat{\mathbf{\Psi}}_{11}^{-1} & \mathbf{0} & \mathbf{0} & \cdots & \hat{\mathbf{\Psi}}_{kk}\mathbf{\Psi}_{kk}^{-1} \end{pmatrix}, \quad (46)$$

we have

$$\begin{aligned} &\text{tr}\{(\mathbf{\Psi}\hat{\mathbf{\Psi}}^{-1})(\mathbf{\Psi}\hat{\mathbf{\Psi}}^{-1})'\} \\ &\quad = \sum_{i=1}^k \text{tr}\{(\hat{\mathbf{\Psi}}'_{ii})^{-1}\mathbf{\Psi}'_{ii}\mathbf{\Psi}_{ii}\hat{\mathbf{\Psi}}_{ii}^{-1}\} \\ &\qquad + \sum_{i=2}^k \text{tr}\{(\mathbf{\Psi}_{i1} - \mathbf{\Psi}_{ii}\hat{\mathbf{\Psi}}_{ii}^{-1}\hat{\mathbf{\Psi}}_{i1})(\hat{\mathbf{\Psi}}'_{11}\hat{\mathbf{\Psi}}_{11})^{-1}(\mathbf{\Psi}_{i1} - \mathbf{\Psi}_{ii}\hat{\mathbf{\Psi}}_{ii}^{-1}\hat{\mathbf{\Psi}}_{i1})'\}. \end{aligned} \quad (47)$$

From (33) and (37) in Theorem 1, it follows

$$\mathbb{E}(\mathbf{\Psi}'_{11}\mathbf{\Psi}_{11} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k) = (\mathbf{T}'_{11})^{-1}\mathbf{D}_1\mathbf{T}_{11}^{-1} \quad (48)$$

and

$$\mathbb{E}(\mathbf{\Psi}'_{ii}\mathbf{\Psi}_{ii} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k) = (\mathbf{T}'_{ii})^{-1}\mathbf{B}_{1iB}\mathbf{T}_{ii}^{-1}, \quad i = 2, \ldots, k. \quad (49)$$

Moreover, for any $2 \leqslant i \leqslant k$, by Theorem 1(b) and applying Theorem 2.3.5 in [8], we have

$$\begin{aligned} &\mathbb{E}\{(\mathbf{\Psi}_{i1} - \mathbf{\Psi}_{ii}\hat{\mathbf{\Psi}}_{ii}^{-1}\hat{\mathbf{\Psi}}_{i1})(\hat{\mathbf{\Psi}}'_{11}\hat{\mathbf{\Psi}}_{11})^{-1}(\mathbf{\Psi}_{i1} - \mathbf{\Psi}_{ii}\hat{\mathbf{\Psi}}_{ii}^{-1}\hat{\mathbf{\Psi}}_{i1})' \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k\} \\ &\quad = \text{tr}\{(\hat{\mathbf{\Psi}}'_{11}\hat{\mathbf{\Psi}}_{11})^{-1}\mathbf{W}_{i11}^{-1}\}\mathbf{I}_{p_i} \\ &\qquad + \mathbb{E}\{\mathbf{\Psi}_{ii}(\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1} + \hat{\mathbf{\Psi}}_{ii}^{-1}\hat{\mathbf{\Psi}}_{i1})(\hat{\mathbf{\Psi}}'_{11}\hat{\mathbf{\Psi}}_{11})^{-1} \\ &\qquad \times (\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1} + \hat{\mathbf{\Psi}}_{ii}^{-1}\hat{\mathbf{\Psi}}_{i1})'\mathbf{\Psi}'_{ii} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k\} \\ &\quad = \text{tr}\{(\hat{\mathbf{\Psi}}'_{11})^{-1}\mathbf{W}_{i11}^{-1}\hat{\mathbf{\Psi}}_{11}^{-1})\mathbf{I}_{p_i} \\ &\qquad + (\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1} + \hat{\mathbf{\Psi}}_{ii}^{-1}\hat{\mathbf{\Psi}}_{i1})(\hat{\mathbf{\Psi}}'_{11}\hat{\mathbf{\Psi}}_{11})^{-1}(\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1} + \hat{\mathbf{\Psi}}_{ii}^{-1}\hat{\mathbf{\Psi}}_{i1})'(\mathbf{T}'_{ii})^{-1}\mathbf{B}_{1i}\mathbf{T}_{ii}^{-1}. \end{aligned} \quad (50)$$

Combining (47)–(50), it yields

$$\begin{aligned}
&\mathbb{E}\left[\operatorname{tr}\{(\boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{-1})(\boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{-1})'\} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k\right] \\
&\quad = \sum_{i=1}^{k} \operatorname{tr}\{(\hat{\boldsymbol{\Psi}}'_{ii})^{-1}(\mathbf{T}'_{ii})^{-1}\mathbf{B}_{1iB}\mathbf{T}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{ii}^{-1}\} \\
&\qquad + \sum_{i=2}^{k} \operatorname{tr}\left\{(\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1} + \hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})(\hat{\boldsymbol{\Psi}}'_{11}\hat{\boldsymbol{\Psi}}_{11})^{-1}\right. \\
&\qquad \left.\times (\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1} + \hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})'(\mathbf{T}'_{ii})^{-1}\mathbf{B}_{1iB}\mathbf{T}_{ii}^{-1}\right\} \qquad (51)
\end{aligned}$$

and thus we have

$$\begin{aligned}
g_1(\hat{\boldsymbol{\Psi}}) &= \sum_{i=1}^{k} \operatorname{tr}\left\{(\hat{\boldsymbol{\Psi}}'_{ii})^{-1}(\mathbf{T}'_{ii})^{-1}\mathbf{B}_{1iB}\mathbf{T}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{ii}^{-1}\right\} - \sum_{i=1}^{k} \log|(\hat{\boldsymbol{\Psi}}'_{ii})^{-1}\hat{\boldsymbol{\Psi}}_{ii}^{-1}| \\
&\quad + \sum_{i=2}^{k} \operatorname{tr}\left\{(\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1} + \hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})\right. \\
&\quad \left.\times (\hat{\boldsymbol{\Psi}}'_{11}\hat{\boldsymbol{\Psi}}_{11})^{-1}(\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1} + \hat{\boldsymbol{\Psi}}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{i1})'(\mathbf{T}'_{ii})^{-1}\mathbf{B}_{1iB}\mathbf{T}_{ii}^{-1}\right\} \\
&\geqslant \sum_{i=1}^{k} \operatorname{tr}\left\{(\hat{\boldsymbol{\Psi}}'_{ii})^{-1}(\mathbf{T}'_{ii})^{-1}\mathbf{B}_{1iB}\mathbf{T}_{ii}^{-1}\hat{\boldsymbol{\Psi}}_{ii}^{-1}\right\} - \sum_{i=1}^{k} \log|(\hat{\boldsymbol{\Psi}}'_{ii})^{-1}\hat{\boldsymbol{\Psi}}_{ii}^{-1}|,
\end{aligned}$$

and the equality holds if we take $\hat{\boldsymbol{\Psi}}_{i1} = -\hat{\boldsymbol{\Psi}}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}, i = 2, 3, \ldots, k$. Consequently, $g_1(\hat{\boldsymbol{\Psi}})$ attaches minimum at $\hat{\boldsymbol{\Psi}}_{11} = \mathbf{G}_{11}\mathbf{T}_{11}^{-1}, \hat{\boldsymbol{\Psi}}_{ii} = \mathbf{B}_{1iB}^{1/2}\mathbf{T}_{ii}^{-1}$ and $\hat{\boldsymbol{\Psi}}_{i1} = -\hat{\boldsymbol{\Psi}}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1}, i = 2, 3, \ldots, k$, where $\mathbf{G}_{11}$ is the inverse of the Cholesky decomposition of $\mathbf{B}_{11}^{-1}$. This completes the proof. □

Remark 2. It is well-known that the group of lower-triangular matrices is solvable and thus its subgroup $\mathcal{G}$ is also solvable (see [3] for a survey). By Kiefer [11], the best $\mathcal{G}$-equivariant estimator $\hat{\boldsymbol{\Omega}}_{1B}$ is also minimax with respect to the Stein loss L_1^*.

The MLE of $\boldsymbol{\Omega}$ given by (20) can be expressed as

$$\hat{\boldsymbol{\Omega}}_M = \mathbf{R}' \, diag(m_1\mathbf{I}_{p_1}, m_2\mathbf{I}_{p_2}, \ldots, m_k\mathbf{I}_{p_k})\mathbf{R},$$

and the unbiased estimate $\hat{\boldsymbol{\Omega}}_U$ of $\boldsymbol{\Omega}$ given by Proposition 4(b) can be expressed as $\hat{\boldsymbol{\Omega}}_U = \mathbf{R}'\mathbf{U}\mathbf{R}$, where

$$\begin{aligned}
\mathbf{U} = diag &\left\{(m_1 - p_1 - 1)\left(1 - \sum_{i=2}^{k} \frac{p_i}{m_i - p_1 - 1}\right)\mathbf{I}_{p_1}, (m_1 - p_1 - p_2 - 1)\mathbf{I}_{p_2}, \ldots,\right. \\
&\left. (m_k - p_1 - p_k - 1)\mathbf{I}_{p_k}\right\}.
\end{aligned}$$

Remark 3. By Corollary 2, both $\hat{\boldsymbol{\Omega}}_M, \hat{\boldsymbol{\Omega}}_U$ are inadmissible under the Stein loss L_1^* because they are equivariant under the group $\mathcal{G}$.

Note that any estimate of $\mathbf{\Omega}$ having the form of $\mathbf{R}'\, diag(a_1, \ldots, a_p)\mathbf{R}$ will be equivariant with respect to the group $\mathcal{G}$, where a_i is a constant, $i = 1, \ldots, p$. Thus, by Theorem 2, any estimator having the form like $\mathbf{R}'\, diag(a_1, \ldots, a_p)\mathbf{R}$ will be inadmissible under the Stein loss. However, it is unclear whether $\mathbf{R}'\, diag(a_1, \ldots, a_p)\mathbf{R}$ is a Bayesian estimate, which also is different from the case with complete data in [23].

5. Best equivariant estimator of $\mathbf{\Omega}$ under the entropy loss

Theorem 3. *Under the entropy loss L_2^*, the best $\mathcal{G}$-equivariant estimator of $\mathbf{\Omega}$ is given by*

$$\hat{\mathbf{\Omega}}_{2B} = \mathbf{R}'\mathbf{B}_{2B}\mathbf{R}, \tag{52}$$

where $\mathbf{R}$ *is given by* (43), $\mathbf{B}_{2B} = diag(\mathbf{B}_{21B}, \mathbf{B}_{22B}, \ldots, \mathbf{B}_{2kB})$, *and* $\mathbf{B}_{2iB} = diag(d_{i1B}, \ldots, d_{ip_iB})$ *with*

$$d_{ijB} = \begin{cases} \dfrac{(m_1 - j - 1)(m_1 - j)}{m_1 - 1} & \text{if } i = 1,\ j = 1, \ldots, p_1, \\ \dfrac{(m_i - p_1 - j - 1)(m_i - p_1 - j)}{(m_i - 1)\{1 + \mathrm{tr}(\mathbf{B}_{21B}^{-1}\mathbf{T}_{11}'\mathbf{W}_{i11}^{-1}\mathbf{T}_{11})\}} & \text{if } i = 2, \ldots, k,\ j = 1, \ldots, p_i. \end{cases} \tag{53}$$

Proof. Similarly to the proof of Theorem 2, best $\mathcal{G}$-equivariant estimator of $\mathbf{\Omega}$ under the entropy loss L_2^* will be produced by minimizing the posterior risk

$$b_2(\hat{\mathbf{\Omega}}) = \int \Big[\mathrm{tr}\big\{\hat{\mathbf{\Omega}}(\mathbf{\Psi}'\mathbf{\Psi})^{-1}\big\} - \log|\hat{\mathbf{\Omega}}(\mathbf{\Psi}'\mathbf{\Psi})^{-1}| - p\Big]\, f(\mathbf{\Psi} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)\, d\mathbf{\Psi},$$

which is equivalent to minimize

$$g_2(\hat{\mathbf{\Psi}}) = \int \mathrm{tr}\{(\hat{\mathbf{\Psi}}\mathbf{\Psi}^{-1})(\hat{\mathbf{\Psi}}\mathbf{\Psi}^{-1})'\} f(\mathbf{\Psi} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)\, d\mathbf{\Psi} - \log|\hat{\mathbf{\Psi}}'\hat{\mathbf{\Psi}}|.$$

Similar to (47), we have

$$\begin{aligned} \mathrm{tr}\{(\hat{\mathbf{\Psi}}\mathbf{\Psi}^{-1})(\hat{\mathbf{\Psi}}\mathbf{\Psi}^{-1})'\} &= \sum_{i=1}^{k} \mathrm{tr}\{\hat{\mathbf{\Psi}}_{ii}(\mathbf{\Psi}_{ii}'\mathbf{\Psi}_{ii})^{-1}\hat{\mathbf{\Psi}}_{ii}'\} \\ &\quad + \sum_{i=2}^{k} \mathrm{tr}\{(\hat{\mathbf{\Psi}}_{i1} - \hat{\mathbf{\Psi}}_{ii}\mathbf{\Psi}_{ii}^{-1}\mathbf{\Psi}_{i1})(\mathbf{\Psi}_{11}'\mathbf{\Psi}_{11})^{-1} \\ &\quad \times(\hat{\mathbf{\Psi}}_{i1} - \hat{\mathbf{\Psi}}_{ii}\mathbf{\Psi}_{ii}^{-1}\mathbf{\Psi}_{i1})'\}. \end{aligned} \tag{54}$$

From (34) and (38),

$$\begin{aligned} &\mathbb{E}\{(\mathbf{\Psi}_{11}'\mathbf{\Psi}_{11})^{-1}|\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k\} = \mathbf{T}_{11}\mathbf{B}_{21B}^{-1}\mathbf{T}_{11}', \\ &\mathbb{E}\{(\mathbf{\Psi}_{ii}'\mathbf{\Psi}_{ii})^{-1}|\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k\} = \mathbf{T}_{ii}\mathbf{B}_{2iB}^{-1}\mathbf{T}_{ii}'/\{1 + \mathrm{tr}(\mathbf{T}_{11}\mathbf{B}_{21B}^{-1}\mathbf{T}_{11}'\mathbf{W}_{i11}^{-1})\}, \\ &\quad i = 2, \ldots, k. \end{aligned}$$

In addition, similar to (50), for $i = 2, \ldots, k$,

$$\begin{aligned} &\mathbb{E}\Big[\mathrm{tr}\big\{(\hat{\mathbf{\Psi}}_{i1} - \hat{\mathbf{\Psi}}_{ii}\mathbf{\Psi}_{ii}^{-1}\mathbf{\Psi}_{i1})(\mathbf{\Psi}_{11}'\mathbf{\Psi}_{11})^{-1}(\hat{\mathbf{\Psi}}_{i1} - \hat{\mathbf{\Psi}}_{ii}\mathbf{\Psi}_{ii}^{-1}\mathbf{\Psi}_{i1})'\big\} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k\Big] \\ &\quad = \mathrm{tr}\Big[\mathbb{E}\big\{(\hat{\mathbf{\Psi}}_{i1} - \hat{\mathbf{\Psi}}_{ii}\mathbf{\Psi}_{ii}^{-1}\mathbf{\Psi}_{i1})'(\hat{\mathbf{\Psi}}_{i1} - \hat{\mathbf{\Psi}}_{ii}\mathbf{\Psi}_{ii}^{-1}\mathbf{\Psi}_{i1}) \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k\big\} \\ &\qquad \times \mathbb{E}\{(\mathbf{\Psi}_{11}'\mathbf{\Psi}_{11})^{-1} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k\}\Big] \end{aligned}$$

$$= \mathrm{tr}\left\{(\hat{\mathbf{\Psi}}_{i1} + \hat{\mathbf{\Psi}}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1})'(\hat{\mathbf{\Psi}}_{i1} + \hat{\mathbf{\Psi}}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1})\mathbf{T}_{11}\mathbf{B}_{21B}^{-1}\mathbf{T}_{11}'\right\}$$
$$+\mathrm{tr}(\mathbf{T}_{11}\mathbf{B}_{21B}^{-1}\mathbf{T}_{11}'\mathbf{W}_{i11}^{-1})\,\mathrm{tr}(\hat{\mathbf{\Psi}}_{ii}\mathbf{T}_{ii}\mathbf{B}_{2iB}\mathbf{T}_{ii}'\hat{\mathbf{\Psi}}_{ii}')/\{1 + \mathrm{tr}(\mathbf{T}_{11}\mathbf{B}_{21B}^{-1}\mathbf{T}_{11}'\mathbf{W}_{i11}^{-1})\}.$$

So we have

$$\mathbb{E}[\mathrm{tr}\{(\hat{\mathbf{\Psi}}\mathbf{\Psi}^{-1})(\hat{\mathbf{\Psi}}\mathbf{\Psi}^{-1})'\} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k]$$
$$= \sum_{i=1}^{k} \mathrm{tr}(\hat{\mathbf{\Psi}}_{ii}\mathbf{T}_{ii}\mathbf{B}_{2iB}^{-1}\mathbf{T}_{ii}'\hat{\mathbf{\Psi}}_{ii}')$$
$$+ \sum_{i=2}^{k} \mathrm{tr}\left\{(\hat{\mathbf{\Psi}}_{i1} + \hat{\mathbf{\Psi}}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1})'(\hat{\mathbf{\Psi}}_{i1} + \hat{\mathbf{\Psi}}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1})\mathbf{T}_{11}\mathbf{B}_{21B}^{-1}\mathbf{T}_{11}'\right\} \tag{55}$$

and thus

$$g_2(\hat{\mathbf{\Psi}}) = \sum_{i=1}^{k} \left\{\mathrm{tr}(\hat{\mathbf{\Psi}}_{ii}\mathbf{T}_{ii}\mathbf{B}_{2iB}^{-1}\mathbf{T}_{ii}'\hat{\mathbf{\Psi}}_{ii}') - \log|\hat{\mathbf{\Psi}}_{ii}'\hat{\mathbf{\Psi}}_{ii}|\right\}$$
$$+ \sum_{i=2}^{k} \mathrm{tr}\left\{(\hat{\mathbf{\Psi}}_{i1} + \hat{\mathbf{\Psi}}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1})'(\hat{\mathbf{\Psi}}_{i1} + \hat{\mathbf{\Psi}}_{ii}\mathbf{W}_{i1}\mathbf{W}_{i11}^{-1})\mathbf{T}_{11}\mathbf{B}_{21B}^{-1}\mathbf{T}_{11}'\right\}.$$

Hence, we can readily see that $g_2(\hat{\mathbf{\Psi}})$ is minimized at $\hat{\mathbf{\Psi}}_{ii} = \mathbf{B}_{2iB}^{1/2}\mathbf{T}_{ii}^{-1}$ for $i = 1, \ldots, k$ and $\hat{\mathbf{\Psi}}_{j1} = -\hat{\mathbf{\Psi}}_{jj}\mathbf{W}_{j1}\mathbf{W}_{j11}^{-1}$ for $j = 2, \ldots, k$. Thus the proof is completed. □

Remark 4. Similar to Remark 2, the best $\mathcal{G}$-equivariant estimator $\hat{\mathbf{\Omega}}_{2B}$ is also minimax with respect to the entropy loss L_2^*.

Remark 5. Similar to Remark 3, the MLE $\hat{\mathbf{\Omega}}_M$ and the unbiased estimator $\hat{\mathbf{\Omega}}_U$ are still inadmissible under the entropy loss L_2^*.

6. Best equivariant estimator of $\mathbf{\Omega}$ under the symmetric loss

We need the following lemma, which is a direct result of Lemma 2.2 in [7].

Lemma 1. *Let $\mathcal{A} = \{\mathbf{A} \in \mathbf{R}^{p\times p} \mid \mathbf{A}$ is lower-triangular with positive diagonal elements$\}$. If $\mathbf{\Delta}$ and $\mathbf{\Lambda}$ are both positive definite, then*

$$\min_{A\in\mathcal{A}}\left\{\mathrm{tr}(\mathbf{A}\mathbf{\Delta}\mathbf{A}') + \mathrm{tr}((\mathbf{A}')^{-1}\mathbf{\Lambda}\mathbf{A}^{-1})\right\} = 2\,\mathrm{tr}(\mathbf{\Lambda}^{1/2}\mathbf{\Delta}\mathbf{\Lambda}^{1/2})^{1/2}$$

is achieved by taking $\mathbf{A}$ as the inverse of Cholesky decomposition of $\mathbf{\Lambda}^{-1/2}(\mathbf{\Lambda}^{1/2}\mathbf{\Delta}\mathbf{\Lambda}^{1/2})^{1/2}\mathbf{\Lambda}^{-1/2}$. Specifically, if $\mathbf{\Delta}$ and $\mathbf{\Lambda}$ are both diagonal, then the minimum will be achieved at $\mathbf{A} = \mathbf{\Lambda}^{1/4}\mathbf{\Delta}^{-1/4}$.

Theorem 4. *Under the symmetric loss L_3^*, the best $\mathcal{G}$-equivariant estimator of $\mathbf{\Omega}$ is given by*

$$\hat{\mathbf{\Omega}}_{3B} = \mathbf{R}'\mathbf{B}_{3B}\mathbf{R}, \tag{56}$$

where $\mathbf{B}_{3B} = diag(\mathbf{B}_{31B}, \mathbf{B}_{32B},\ldots,\mathbf{B}_{3kB})$ with $\mathbf{B}_{31B} = \mathbf{B}_{11B}^{1/2}(\mathbf{B}_{11B}^{1/2}\mathbf{B}_{21B}^{-1}\mathbf{B}_{11B}^{1/2})^{-1/2}\mathbf{B}_{11B}^{1/2}$ and $\mathbf{B}_{3iB} = \mathbf{B}_{1iB}^{1/2}\mathbf{B}_{2iB}^{1/2}$, $i = 2, \ldots, k$, where $\mathbf{B}_{1iB}$ and $\mathbf{B}_{2iB}$ are given by Theorems 2 *and* 3, *respectively*, $i = 1, 2, \ldots, k$.

Proof. The Bayesian estimator of $\mathbf{\Omega}$ under the symmetric loss L_3^* will be produced by minimizing the posterior risk

$$b_3(\hat{\mathbf{\Omega}})= \int \left[\text{tr} \left\{ \hat{\mathbf{\Omega}}(\mathbf{\Psi}'\mathbf{\Psi})^{-1} \right\} + \text{tr} \left\{ \hat{\mathbf{\Omega}}^{-1}(\mathbf{\Psi}'\mathbf{\Psi}) \right\} - 2p \right] f(\mathbf{\Psi} \mid \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)\, d\mathbf{\Psi},$$

which is equivalent to minimize

$$g_3(\hat{\mathbf{\Psi}})= \int \left[\text{tr}\left\{ (\hat{\mathbf{\Psi}}\mathbf{\Psi}^{-1})(\hat{\mathbf{\Psi}}\mathbf{\Psi}^{-1})' \right\} + \text{tr}\left\{ (\mathbf{\Psi}\hat{\mathbf{\Psi}}^{-1})(\mathbf{\Psi}\hat{\mathbf{\Psi}}^{-1})' \right\} \right] f(\mathbf{\Psi}|\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_k)\, d\mathbf{\Psi}.$$

Combining (55) with (51), it yields

$$\begin{aligned} g_3(\hat{\mathbf{\Psi}}) &= \sum_{i=1}^{k} \Big[\text{tr}\big(\hat{\mathbf{\Psi}}_{ii} \mathbf{T}_{ii} \mathbf{B}_{2iB}^{-1} \mathbf{T}_{ii}' \hat{\mathbf{\Psi}}_{ii}' \big) + \text{tr}\big\{ (\hat{\mathbf{\Psi}}_{ii}')^{-1} (\mathbf{T}_{ii}')^{-1} \mathbf{B}_{1iB} \mathbf{T}_{ii}^{-1} \hat{\mathbf{\Psi}}_{ii}^{-1} \big\} \Big] \\ &\quad + \sum_{i=2}^{k} \text{tr}\big\{ (\hat{\mathbf{\Psi}}_{i1} + \hat{\mathbf{\Psi}}_{ii} \mathbf{W}_{i1} \mathbf{W}_{i11}^{-1})' (\hat{\mathbf{\Psi}}_{i1} + \hat{\mathbf{\Psi}}_{ii} \mathbf{W}_{i1} \mathbf{W}_{11}^{-1}) \mathbf{T}_{11} \mathbf{B}_{21B}^{-1} \mathbf{T}_{11}' \big\} \\ &\quad + \sum_{i=2}^{k} \text{tr}\big\{ (\mathbf{W}_{i1} \mathbf{W}_{11}^{-1} + \hat{\mathbf{\Psi}}_{ii}^{-1} \hat{\mathbf{\Psi}}_{i1}) (\hat{\mathbf{\Psi}}_{11}' \hat{\mathbf{\Psi}}_{11})^{-1} \\ &\quad \times (\mathbf{W}_{i1} \mathbf{W}_{11}^{-1} + \hat{\mathbf{\Psi}}_{ii}^{-1} \hat{\mathbf{\Psi}}_{i1})' (\mathbf{T}_{ii}')^{-1} \mathbf{B}_{1iB} \mathbf{T}_{ii}^{-1} \big\} \\ &\geqslant \sum_{i=1}^{k} \Big[\text{tr}\big(\hat{\mathbf{\Psi}}_{ii} \mathbf{T}_{ii} \mathbf{B}_{2iB}^{-1} \mathbf{T}_{ii}' \hat{\mathbf{\Psi}}_{ii}' \big) + \text{tr}\big\{ (\hat{\mathbf{\Psi}}_{ii}')^{-1} (\mathbf{T}_{ii})^{-1} \mathbf{B}_{1iB} \mathbf{T}_{ii}^{-1} \hat{\mathbf{\Psi}}_{ii}^{-1} \big\} \Big], \end{aligned}$$

and the equality holds if we take $\hat{\mathbf{\Psi}}_{i1} = -\hat{\mathbf{\Psi}}_{ii} \mathbf{W}_{i1} \mathbf{W}_{i11}^{-1}$, $i = 2, 3, \ldots, k$. Thus, by Lemma 1, we can easily see that $g_3(\hat{\mathbf{\Psi}})$ attaches minimum at $\hat{\mathbf{\Psi}}_{11} = \mathbf{Q}_{31} \mathbf{T}_{11}^{-1}$ with $\mathbf{Q}_{31}$ being the inverse of Cholesky decomposition of $\mathbf{B}_{11B}^{-1/2} (\mathbf{B}_{11B}^{1/2} \mathbf{B}_{21B}^{-1} \mathbf{B}_{11B}^{1/2})^{1/2} \mathbf{B}_{11B}^{-1/2}$, $\hat{\mathbf{\Psi}}_{ii} = \mathbf{B}_{3iB}^{1/2} \mathbf{T}_{ii}^{-1}$ and $\hat{\mathbf{\Psi}}_{i1} = -\hat{\mathbf{\Psi}}_{ii} \mathbf{W}_{i1} \mathbf{W}_{i11}^{-1}$, $i = 2, 3, \ldots, k$. Thus the proof is completed. □

Remark 6. Similar to Remarks 2 and 4, the best $\mathcal{G}$-equivariant estimators $\hat{\mathbf{\Omega}}_2$ is also minimax with respect to the symmetric loss L_3^*.

Remark 7. Similar to Remarks 3 and 5, both $\hat{\mathbf{\Omega}}_M$ and $\hat{\mathbf{\Omega}}_U$ are inadmissible under the symmetric loss L_3^*.

7. Estimating the covariance matrix

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

Corollary 1. *Under the loss* L_i, $i = 1, 2, 3$, *the* $\mathcal{G}$*-equivariant estimator of* $\mathbf{\Sigma}$ *is given by*

$$\hat{\mathbf{\Sigma}}_{iB} = \mathbf{T} \mathbf{B}_{iB}^{-1} \mathbf{T}', \tag{57}$$

where $\mathbf{T}$ *is given by* (42) *and* $\mathbf{B}_{iB}$, $i = 1, 2, 3$ *is shown by Theorems* 2, 3, *and* 4, *respectively*.

Remark 8. For fixed $i = 1, 2, 3$, both the MLE $\hat{\mathbf{\Sigma}}_M$ and the unbiased estimator $\hat{\mathbf{\Sigma}}_U$ are inadmissible under the loss L_i.

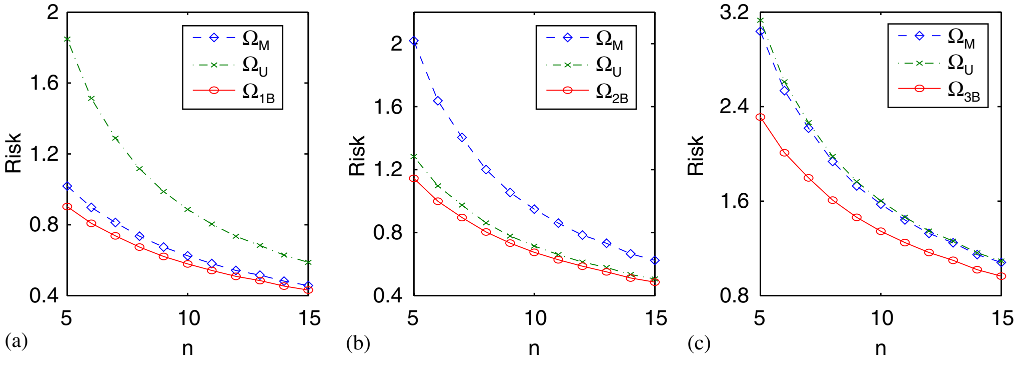

Fig. 1. Risk comparisons when $(p_1, p_2, p_3) = (2, 1, 1)$, $(n_1, n_2, n_3) = (3, 4, 5)$ and $5 \leqslant n \leqslant 15$: (a) under L_1^*, (b) under L_2^*, and (c) under L_3^*.

Remark 9. For $i = 1, 2, 3$, the best $\mathcal{G}$-equivariant estimator $\hat{\mathbf{\Sigma}}_i$ is minimax with respect to L_i.

8. Simulation results

In this section, we will compare the risks of MLE $\hat{\mathbf{\Omega}}_M$, the unbiased estimator $\hat{\mathbf{\Omega}}_U$ and the best equivariant estimator $\hat{\mathbf{\Omega}}_{iB}$ under each L_i^*, $i = 1, 2, 3$. Each risk will be denoted as R_{iM}, R_{iU}, R_{iB} under L_i^*, respectively.

Unlike the model with complete data studied by Sun and Sun [23], it is rather complicated to derive a closed form expression for the risks of the above estimates under any L_i^*. So we compare their risks by simulation. Because all of these estimators are equivariant, without loss of generality, we may take $\mathbf{\Sigma} = \mathbf{I}_p$ in our simulation. Risks of the three estimators $\hat{\mathbf{\Omega}}_M$, $\hat{\mathbf{\Omega}}_U$ and $\hat{\mathbf{\Omega}}_i$ under losses L_1^*, L_2^* and L_3^* for various combinations of n_i and p_i, and $k = 3$ and 5 are plotted in Figs. 1–9. The numerical values are computed based on 10, 000 simulated samples. From the simulation study, $\hat{\mathbf{\Omega}}_M$ is superior to $\hat{\mathbf{\Omega}}_U$ under the loss L_1^*, but $\hat{\mathbf{\Omega}}_U$ is superior to $\hat{\mathbf{\Omega}}_M$ under the loss L_2^*. There is no difference between $\hat{\mathbf{\Omega}}_M$ and $\hat{\mathbf{\Omega}}_U$ under loss L_3^*. Furthermore, the improvement over the risks of $\hat{\mathbf{\Omega}}_M$ and $\hat{\mathbf{\Omega}}_U$ by $\hat{\mathbf{\Omega}}_i$ under all three losses is quite substantial.

9. Concluding remarks

This paper deals with the problem of estimating the covariance matrix and the precision matrix under the three common loss functions in a star-shape model with missing data. Using a type of Cholesky decomposition of the precision matrix $\mathbf{\Omega} = \mathbf{\Psi}'\mathbf{\Psi}$, we easily obtained the MLEs of the covariance matrix and the precision matrix. Also, we get the closed forms of the best equivariant estimators of $\mathbf{\Omega}$ under the Stein loss, entropy loss and symmetric loss, respectively. This method is quite powerful in estimating the covariance matrix or the precision matrix.

Although our sample plan is restricted to taking observations from $\mathbf{X}, \mathbf{X}_1, (\mathbf{X}_1', \mathbf{X}_i')'$, $i = 2, \ldots, k$, which is popular in economic studies, we can deal with other cases such as taking observations from $\mathbf{X}, \mathbf{X}_1, (\mathbf{X}_1', \mathbf{X}_{i_1}', \mathbf{X}_{i_2}', \ldots, \mathbf{X}_{i_j}')'$, $2 \leqslant i_1 < \cdots < i_j \leqslant k$ by applying the similar method. In these cases, the monotone missing data pattern is not required when the covariance matrix has a special structure, which is different from the case of the covariance matrix with no

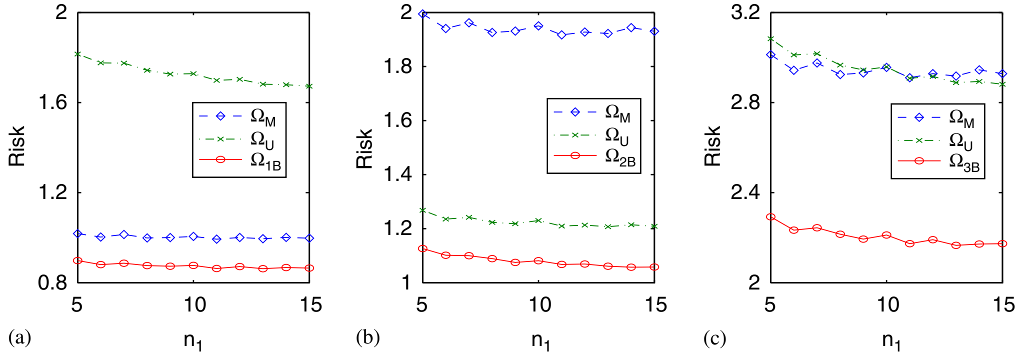

Fig. 2. Risk comparisons when $(p_1, p_2, p_3) = (2, 1, 1)$, $(n, n_2, n_3) = (5, 4, 5)$ and $5 \leqslant n_1 \leqslant 15$: (a) under L_1^*, (b) under L_2^*, and (c) under L_3^*.

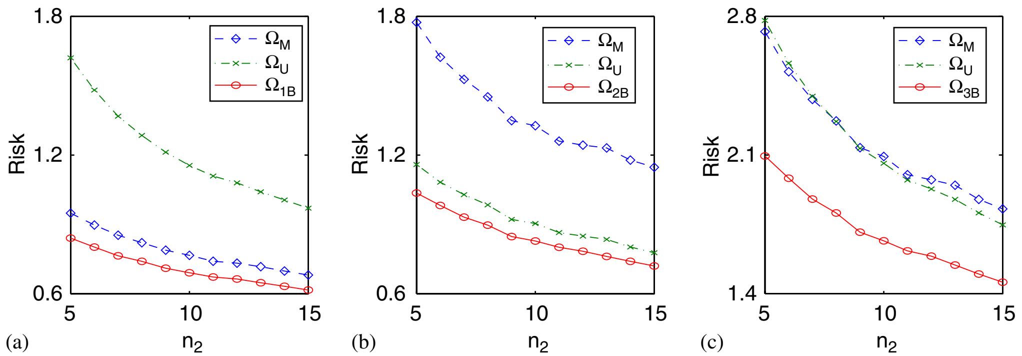

Fig. 3. Risk comparisons when $(p_1, p_2, p_3) = (2, 1, 1)$, $(n, n_1, n_3) = (5, 4, 5)$ and $5 \leqslant n_2 \leqslant 15$: (a) under L_1^*, (b) under L_2^*, and (c) under L_3^*.

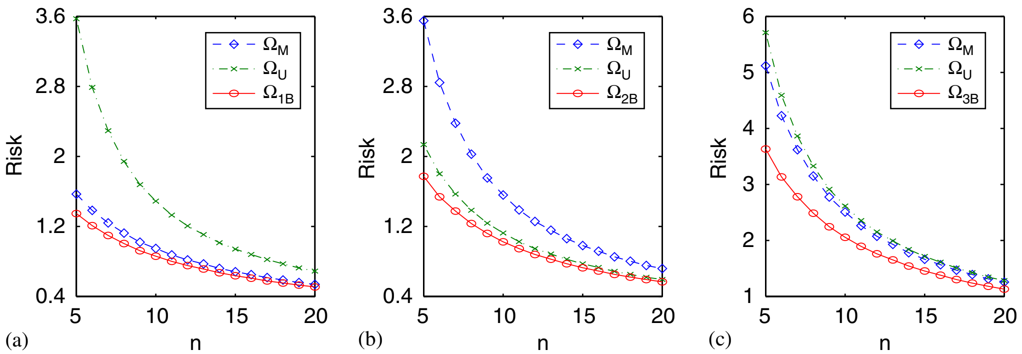

Fig. 4. Risk comparisons when $(p_1, p_2, p_3) = (2, 2, 1)$, $(n_1, n_2, n_3) = (3, 5, 4)$ and $5 \leqslant n \leqslant 20$: (a) under L_1^*, (b) under L_2^*, and (c) under L_3^*.

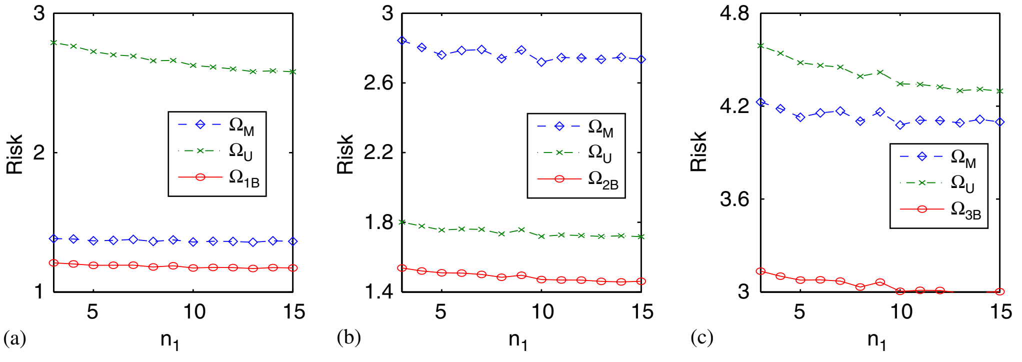

Fig. 5. Risk comparisons when $(p_1, p_2, p_3) = (2, 2, 1)$, $(n, n_2, n_3) = (6, 5, 4)$ and $3 \leqslant n_1 \leqslant 15$: (a) under L_1^*, (b) under L_2^*, and (c) under L_3^*.

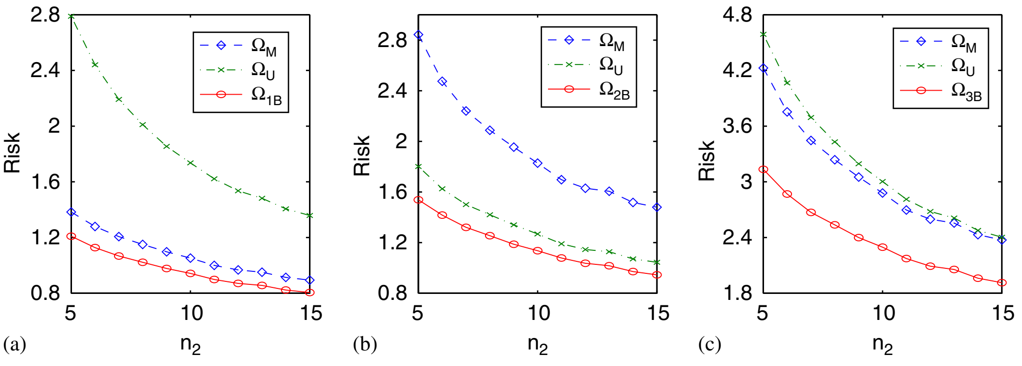

Fig. 6. Risk comparisons when $(p_1, p_2, p_3) = (2, 2, 1)$, $(n, n_1, n_3) = (6, 3, 4)$ and $5 \leqslant n_2 \leqslant 15$: (a) under L_1^*, (b) under L_2^*, and (c) under L_3^*.

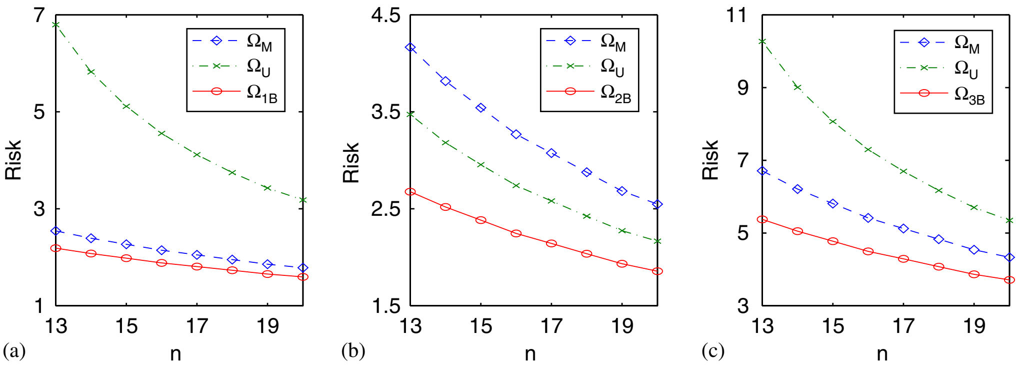

Fig. 7. Risk comparisons when $(p_1, p_2, p_3, p_4, p_5) = (2, 2, 3, 4, 1)$, $(n_1, n_2, n_3, n_4, n_5) = (3, 5, 6, 7, 4)$ and $13 \leqslant n \leqslant 20$: (a) under L_1^*, (b) under L_2^*, and (c) under L_3^*.

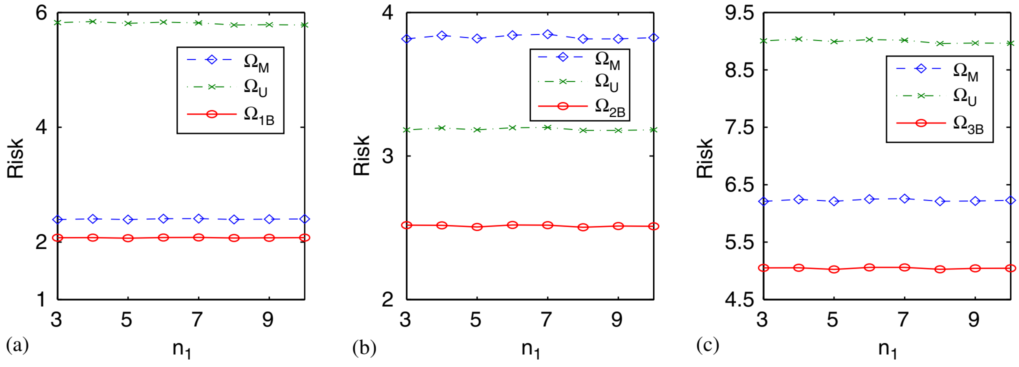

Fig. 8. Risk comparisons when $(p_1, p_2, p_3, p_4, p_5) = (2, 2, 3, 4, 1)$, $(n, n_2, n_3, n_4, n_5) = (14, 5, 6, 7, 4)$ and $3 \leqslant n_1 \leqslant 10$: (a) under L_1^*, (b) under L_2^*, and (c) under L_3^*.

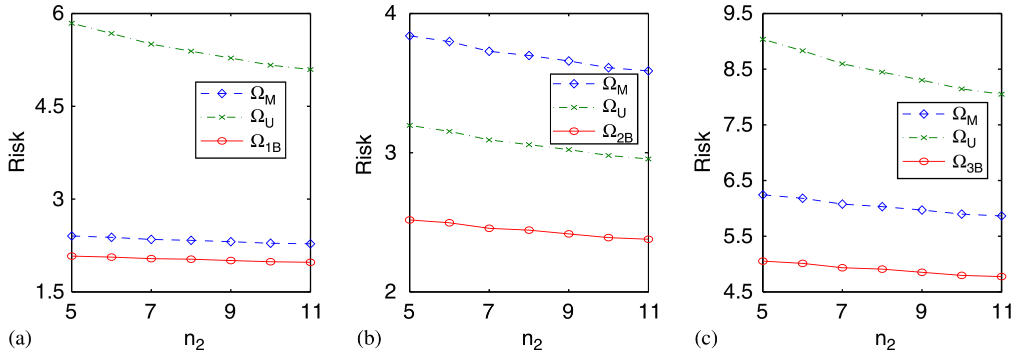

Fig. 9. Risk comparisons when $(p_1, p_2, p_3, p_4, p_5) = (2, 2, 3, 4, 1)$, $(n, n_1, n_3, n_4, n_5) = (14, 4, 6, 7, 4)$ and $13 \leqslant n_2 \leqslant 20$: (a) under L_1^*, (b) under L_2^*, and (c) under L_3^*.

restriction by Anderson [1] and Liu [18] and so on. In addition, for convenience, we assume that the sample sizes satisfying $n > p$, $n_1 > p_1$ and $n_i > p_1 + p_i$, $i = 2, \ldots, k$ in this paper. The essential conditions are $m_1 > p_1 + 1$ and $m_i > p_1 + p_i + 1$.

The investigation on a star-shape model with missing data is, nevertheless, far from being complete, and there are many important and interesting questions to be further studied. An interesting but difficult problem is whether the best equivariant estimate $\hat{\Sigma}_{iB}$ is admissible under the corresponding loss L_i. Other Bayesian estimates by using appropriate priors will be considered in the future.

Acknowledgements

The project was partially supported by the National Science Foundation Grants SES-0095919 and SES-0351523 and a grant from the Missouri Department of Conservation. The authors wish to express deep appreciation to the editor and a referee for helpful comments on an earlier version of the paper. The author is grateful to Dr. Karon F. Speckman for her carefully proof-reading this paper.

References

[1] T.W. Anderson, Maximum likelihood estimates for a multivariate normal distribution when some observations are missing, J. Amer. Statist. Assoc. 52 (1957) 200–203.
[2] S.A. Andersson, M.D. Perlman, Lattice models for conditional independence in a multivariate normal distribution, Ann. Statist. 21 (1993) 1318–1358.
[3] J.V. Bondar, P. Milnes, Amenability: A survey for statistical applications of Hunt–Stein and related conditions on groups, Z. Wahr. Verw. Geb. [Became: @J(ProbTher)] 57 (1981) 103–128.
[4] A.M. Dempster, Covariance selection, Biometrics 28 (1972) 157–175.
[5] F. Domonici, G. Parmigiani, M. Clyde, Conjugate analysis of multivariate normal data incomplete observations, Canad. J. Statist. 28 (2000) 533–550.
[6] M.L. Eaton, Some problems in covariance estimation (preliminary report), Technical Report 49, Department of Statistics, Stanford University, 1970.
[7] M.L. Eaton, Group invariance applications in statistics, Institute of Mathematical Statistics, 1989.
[8] A.K. Gupta, S. Ofori-Nyarko, Improved minimax estimators of normal covariance and precision matrices, Statistics [Formerly: @J(MaOpfStS)] 26 (1995) 19–25.
[9] L.R. Haff, Empirical Bayes estimation of the multivariate normal covariance matrix, Ann. Statist. 8 (1980) 586–597.
[10] W. James, C. Stein, Estimation with quadratic loss, in: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, 1961.
[11] J. Kiefer, Invariance, minimax sequential estimation, and continuous time process, Ann. Math. Statist. 28 (1957) 573 –601.
[12] Y. Konno, Estimation of a normal covariance matrix with incomplete data under Stein's loss, J. Multivariate Anal. 52 (1995) 308–324.
[13] Y. Konno, Inadmissibility of the maximum likelihood estimator of normal covariance matrices with the lattice conditional independence, J. Multivariate Anal. 79 (2001) 33–51.
[14] K. Krishnamoorthy, A.K. Gupta, Improved minimax estimation of a normal precision matrix, Canad. J. Statist. 17 (1989) 91–102.
[15] T. Kubokawa, Y. Konno, Estimating the covariance matrix and the generalized variance under a symmetric loss, Ann. Inst. Statist. Math. 42 (1990) 331–343.
[16] S.L. Lauritzen, Graphical Models, Oxford University Press, Oxford, 1996.
[17] R.J.A. Little, D.B. Rubin, Statistical Analysis with Missing Data, Wiley, New York, 2002.
[18] C. Liu, Efficient ML estimation of the multivariate normal distribution from incomplete data, J. Multivariate Anal. 69 (1999) 206–217.
[19] I. Olkin, J.B. Selliah, Estimating covariances in a multivariate distribution, in: Statistical Decision Theory and Related Topics, Proceedings of a Symposium at Purdue University, May, 1976, in two volumes, vol. 2, 1977.
[20] C.P. Robert, The Bayesian Choice: a Decision-theoretic Motivation, Springer, Berlin, 1994.
[21] D. Sharma, K. Krishnamoorthy, Orthogonal equivariant minimax estimators of bivariate normal covariance matrix and precision matrix, Calcutta Statist. Assoc. Bull. 32 (1983) 23–45.
[22] B.K. Sinha, M. Ghosh, Inadmissibility of the best equivariant estimators of the variance–covariance matrix, the precision matrix, and the generalized variance under entropy loss, Statist. Decisions 5 (1987) 201–227.
[23] D. Sun, X. Sun, Estimation of the multivariate normal precision and covariance matrices in a star-shape model, Ann. Inst. Statist. Math. 57 (2005) 455–484.
[24] J. Whittaker, Graphical Models in Applied Multivariate Statistics, Wiley, New York, 1990.
[25] R. Yang, J.O. Berger, Estimation of a covariance matrix using the reference prior, Ann. Statist. 22 (1994) 1195–1211.