Commun. Math. Stat. (2018) 6:163–184
https://doi.org/10.1007/s40304-018-0130-4

CrossMark

Hypothesis Testing for Independence Under Blocked Compound Symmetric Covariance Structure

Shin-ichi Tsukada[1]

Received: 14 September 2017 / Revised: 2 January 2018 / Accepted: 15 March 2018 /
Published online: 30 April 2018

Abstract One type of covariance structure is known as blocked compound symmetry. Recently, Roy et al. (J Multivar Anal 144:81–90, 2016) showed that, assuming this covariance structure, unbiased estimators are optimal under normality and described hypothesis testing for independence as an open problem. In this paper, we derive the distributions of unbiased estimators and consider hypothesis testing for independence. Representative test statistics such as the likelihood ratio criterion, Wald statistic, Rao's score statistic, and gradient statistic are derived, and we evaluate the accuracy of the test using these statistics through numerical simulations. The power of the Wald test is the largest when the dimension is high, and the power of the likelihood ratio test is the largest when the dimension is low.

Keywords Hypothesis testing · Asymptotic distribution · Independence · Blocked compound symmetric covariance structure

Mathematics Subject Classification 62H15 · 62E15 · 62E20

1 Introduction

In multivariate statistical analysis, the covariance matrix can have various specific structures. One of these is the blocked compound symmetric (BCS) covariance structure. The BCS covariance structure for doubly multivariate observations is a multivariate generalization of the compound symmetric covariance structure for mul-

✉ Shin-ichi Tsukada
tsukada@ed.meisei-u.ac.jp

1 School of Education, Meisei University, 2-1-1 Hodokubo Hino, Tokyo 191-8506, Japan

tivariate observations. The BCS covariance structure is defined as:

$$\boldsymbol{\Sigma} = \boldsymbol{I}_u \otimes (\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1) + \boldsymbol{J}_u \otimes \boldsymbol{\Sigma}_1 = \begin{pmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \cdots & \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \cdots & \boldsymbol{\Sigma}_1 \\ \vdots & \vdots & & \vdots \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_1 & \cdots & \boldsymbol{\Sigma}_0 \end{pmatrix},$$

where $\boldsymbol{I}_u$ is the $u \times u$ identity matrix, $\mathbf{1}_u$ is a $u \times 1$ vector of ones, $\boldsymbol{J}_u = \mathbf{1}_u \mathbf{1}_u'$, and $\otimes$ denotes the Kronecker product. We assume that $u \geq 2$, $\boldsymbol{\Sigma}_0$ is a positive-definite symmetric $p \times p$ matrix, and $\boldsymbol{\Sigma}_1$ is a symmetric $p \times p$ matrix. We also assume that $\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_0 + (u-1)\boldsymbol{\Sigma}_1$ are positive-definite matrices so that $\boldsymbol{\Sigma}$ is a positive-definite matrix. Arnold [2] studied this covariance structure in the general linear model when the error vectors are assumed to be exchangeable and normally distributed. Szatrowski [13] discussed the BCS covariance structure and used a model to analyze an educational testing problem. Leiva [8] derived maximum likelihood estimates (MLEs) of the BCS covariance structure, developed classification rules for doubly multivariate observations and generalized Fisher's linear discrimination method under the BCS covariance structure.

Recently, the BCS covariance structure has been actively researched. For three-level multivariate data, Roy and Leiva [10] and Coelho and Roy [3] have developed hypothesis testing frameworks for a covariance structure. Roy et al. [11] and Zezula et al. [15] studied hypothesis testing for the equality of mean vectors in two populations under the BCS covariance structure. Roy et al. [12] proved that the unbiased estimators of the BCS covariance structure are optimal under normality.

We consider hypothesis testing for independence under the BCS covariance structure, i.e.,

$$H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{O} \text{ versus } H_1 : \boldsymbol{\Sigma}_1 \neq \boldsymbol{O},$$

where $\boldsymbol{O}$ is a $p \times p$ zero matrix. This problem is the extension of an independence test for a covariance matrix to an independence test for a blocked covariance matrix. We investigate the properties of the unbiased estimator of the covariance matrix and use them to derive the Wald statistic. We also derive the likelihood ratio criterion (LRC), the modified LRC using the Bartlett correction, Rao's score statistic, and Terrel's [14] gradient statistic. The asymptotic behavior of these test statistics is similar, but the accuracy of their convergence to the significance level and the powers of test using these statistics are investigated for finite samples through numerical simulations. From the simulation results, we find that the accuracy of convergence to the significance level differs depending on the statistic. Therefore, we also simulate the bootstrap test using these test statistics. Simulation results show that the tests using these statistics converge to the significance level for large samples, the power of the test using the Wald statistic is the largest when the dimension is high, and the power of the likelihood ratio test is the largest when the dimension is low.

The remainder of this article is organized as follows. The properties of the unbiased estimator are obtained in Sect. 2. In Sect. 3, the LRC, modified LRC, Wald statistic, Rao's score statistic and gradient criterion are derived, and the process of the bootstrap test using the relevant statistics is described. Numerical simulations and an application to real data are reported in Sect. 4. Finally, Sect. 5 contains our conclusions.

2 Estimators

We assume that $\boldsymbol{x}_{r,s}$ is a p-variate vector of measurements on the r-th individual at the s-th site ($r = 1, \ldots, n$, $s = 1, \ldots, u$). The n individuals are all independent. Let $\boldsymbol{x}_r = (\boldsymbol{x}'_{r,1}, \ldots, \boldsymbol{x}'_{r,u})'$ be the up-variate vector of all measurements corresponding to the r-th individual. Finally, we assume that $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ be a random sample of size n drawn from the population $N_{up}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = (\boldsymbol{\mu}'_1, \ldots, \boldsymbol{\mu}'_u)'$ is a $up \times 1$ vector and $\boldsymbol{\Sigma}$ is a $up \times up$ positive-definite matrix that has the BCS covariance structure (cf. Leiva [8]).

In this section, we discuss estimators under the BCS covariance structure. Roy et al. [12] derive unbiased estimators as follows:

Theorem 2.1 (Roy et al. [12]) *Assume that $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ is a random sample of size n drawn from the population $N_{up}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Let $\bar{\boldsymbol{x}} = (\bar{\boldsymbol{x}}'_1, \bar{\boldsymbol{x}}'_2, \ldots, \bar{\boldsymbol{x}}'_u)'$,*

$$\boldsymbol{C}_0 = \sum_{s=1}^{u} \sum_{r=1}^{n} \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right) \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right)',$$

$$\boldsymbol{C}_1 = \sum_{\substack{s=1 \\ s\neq s^*}}^{u} \sum_{s^*=1}^{u} \sum_{r=1}^{n} \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right) \left(\boldsymbol{x}_{r,s^*} - \bar{\boldsymbol{x}}_{s^*}\right)',$$

where $\bar{\boldsymbol{x}}_s = \sum_{r=1}^{n} \boldsymbol{x}_{r,s}/n$ ($s = 1, \ldots, u$). Then, $\bar{\boldsymbol{x}}$ is distributed as $N_{up}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$ and is the unbiased estimator for the mean vector $\boldsymbol{\mu}$. The estimators

$$\tilde{\boldsymbol{\Sigma}}_0 = \frac{1}{u(n-1)} \boldsymbol{C}_0 \quad \textit{and} \quad \tilde{\boldsymbol{\Sigma}}_1 = \frac{1}{u(u-1)(n-1)} \boldsymbol{C}_1$$

are unbiased estimators for $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$, respectively.

Therefore, the unbiased estimator for $\boldsymbol{\Sigma}$ is

$$\tilde{\boldsymbol{\Sigma}} = \boldsymbol{I}_u \otimes (\tilde{\boldsymbol{\Sigma}}_0 - \tilde{\boldsymbol{\Sigma}}_1) + \boldsymbol{J}_u \otimes \tilde{\boldsymbol{\Sigma}}_1.$$

For further inference, we derive the distribution for these estimators under some assumptions. The distribution of an unbiased estimator for $\boldsymbol{\mu}$ is $N_{up}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$, but the estimators $\tilde{\boldsymbol{\Sigma}}_0$ and $\tilde{\boldsymbol{\Sigma}}_1$ do not follow a Wishart distribution, even when the population distribution is normal. We obtain the exact distribution of $\tilde{\boldsymbol{\Sigma}}_0$ and $\tilde{\boldsymbol{\Sigma}}_1$. Roy et al. [11] indicated that

$$\boldsymbol{W}_1 \equiv (n-1)(u-1) \left(\tilde{\boldsymbol{\Sigma}}_0 - \tilde{\boldsymbol{\Sigma}}_1\right) \sim W_p((n-1)(u-1), \boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1), \qquad (2.1)$$

$$\boldsymbol{W}_2 \equiv (n-1)\left\{\tilde{\boldsymbol{\Sigma}}_0 + (u-1)\tilde{\boldsymbol{\Sigma}}_1\right\} \sim W_p(n-1, \boldsymbol{\Sigma}_0 + (u-1)\boldsymbol{\Sigma}_1), \tag{2.2}$$

and these estimators are independent of each other. The estimator $\boldsymbol{W}_1$ is positive-definite when $(n-1)(u-1) \geq p$ and the estimator $\boldsymbol{W}_2$ is positive-definite when $n-1 \geq p$. When $n > p$, these inequalities are true for $u \geq 2$. Since

$$(n-1)u\tilde{\boldsymbol{\Sigma}}_0 = \boldsymbol{W}_1 + \boldsymbol{W}_2,$$
$$(n-1)u(u-1)\tilde{\boldsymbol{\Sigma}}_1 = (u-1)\boldsymbol{W}_2 - \boldsymbol{W}_1,$$

the exact distributions of $\tilde{\boldsymbol{\Sigma}}_0$ and $\tilde{\boldsymbol{\Sigma}}_1$ are obtained as the sum and the difference of Wishart matrices.

Lemma 2.2 *Let* $\boldsymbol{\Delta}_1 = \boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1$, *and* $\boldsymbol{\Delta}_2 = \boldsymbol{\Sigma}_0 + (u-1)\boldsymbol{\Sigma}_1$. *When* $u \geq 2$ *and* $n > p$, *the exact distribution of* $\tilde{\boldsymbol{\Sigma}}_0$ *is as follows*:

$$K_0 \operatorname{etr}\left[-\frac{(n-1)u}{2}\boldsymbol{\Delta}_1^{-1}\tilde{\boldsymbol{\Sigma}}_0\right]\left|\tilde{\boldsymbol{\Sigma}}_0\right|^{u(n-1)/2-p-1}$$
$$\times\ {}_1F_1\left[\frac{1}{2}(n-1); \frac{1}{2}(n-1)u; \frac{(n-1)u}{2}\left(\boldsymbol{\Delta}_1^{-1} - \boldsymbol{\Delta}_2^{-1}\right)\tilde{\boldsymbol{\Sigma}}_0\right],$$

where $\operatorname{etr}(\boldsymbol{H}) = \exp[\operatorname{tr}(\boldsymbol{H})]$,

$$K_0 = \left[\left\{\frac{2}{(n-1)u}\right\}^{u(n-1)p/2}\Gamma_p\left[\frac{1}{2}u(n-1)\right]|\boldsymbol{\Delta}_1|^{(n-1)(u-1)/2}\,|\boldsymbol{\Delta}_2|^{(n-1)/2}\right]^{-1},$$

and ${}_1F_1[a; b; \boldsymbol{H}]$ *is the hypergeometric function of a matrix argument defined by* (5.1).

Proof The details of the proof are described in "Appendix A". □

Lemma 2.3 *When* $u \geq 2$ *and* $n > p$, *the exact distribution of* $\tilde{\boldsymbol{\Sigma}}_1$ *is as follows*:

$$K_1 \operatorname{etr}\left[-\frac{(n-1)u}{2}\boldsymbol{\Delta}_2^{-1}\tilde{\boldsymbol{\Sigma}}_1\right]\left|\tilde{\boldsymbol{\Sigma}}_1\right|^{u(n-1)/2-p-1}$$
$$\times\ \Psi\left[\frac{1}{2}(n-1)(u-1), \frac{1}{2}(n-1)u; \frac{1}{2}\left\{(n-1)u\boldsymbol{\Delta}_2^{-1} + (n-1)u(u-1)\boldsymbol{\Delta}_1^{-1}\right\}\tilde{\boldsymbol{\Sigma}}_1\right],$$

where $\Psi[a, c; \boldsymbol{R}]$ *is the confluent hypergeometric function defined by* (5.2), *and*

$$K_1 = \left[\left\{\frac{2}{(n-1)u}\right\}^{u(n-1)p/2}\Gamma_p\left[\frac{1}{2}(n-1)\right]\left(\frac{1}{u-1}\right)^{(n-1)(u-1)p/2}\right.$$
$$\left.\times\ |\boldsymbol{\Delta}_1|^{(n-1)(u-1)/2}\,|\boldsymbol{\Delta}_2|^{(n-1)/2}\right]^{-1}.$$

Proof The details of the proof are described in "Appendix A". □

The exact distributions of $\tilde{\boldsymbol{\Sigma}}_0$ and $\tilde{\boldsymbol{\Sigma}}_1$ contain a hypergeometric function of the matrix argument, which is generally difficult to calculate. We may need the asymptotic distribution of the estimators.

Since the estimators vec $\left(\tilde{\boldsymbol{\Sigma}}_0\right)$ and vec $\left(\tilde{\boldsymbol{\Sigma}}_1\right)$ are represented as follows:

$$\begin{aligned}\mathrm{vec}\left(\tilde{\boldsymbol{\Sigma}}_0\right) &= \mathrm{vec}\left(\boldsymbol{\Sigma}_0\right) + \frac{u-1}{u}\mathrm{vec}\left(\tilde{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right) + \frac{1}{u}\mathrm{vec}\left(\tilde{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right),\\ \mathrm{vec}\left(\tilde{\boldsymbol{\Sigma}}_1\right) &= \mathrm{vec}\left(\boldsymbol{\Sigma}_1\right) - \frac{1}{u}\mathrm{vec}\left(\tilde{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right) + \frac{1}{u}\mathrm{vec}\left(\tilde{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right),\end{aligned}$$

the following theorem can be obtained using the properties of Wishart matrices.

Theorem 2.4 *Let*

$$\begin{aligned}\boldsymbol{\Phi}_0 &= \frac{u-1}{u^2}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)\left(\boldsymbol{\Delta}_1 \otimes \boldsymbol{\Delta}_1\right) + \frac{1}{u^2}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)\left(\boldsymbol{\Delta}_2 \otimes \boldsymbol{\Delta}_2\right),\\ \boldsymbol{\Phi}_1 &= \frac{1}{u^2(u-1)}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)\left(\boldsymbol{\Delta}_1 \otimes \boldsymbol{\Delta}_1\right) + \frac{1}{u^2}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)\left(\boldsymbol{\Delta}_2 \otimes \boldsymbol{\Delta}_2\right),\end{aligned}$$

where $\boldsymbol{K}_{p,p}$ is the commutation matrix.

The vectors $(n-1)^{1/2}$vec $\left(\tilde{\boldsymbol{\Sigma}}_0 - \boldsymbol{\Sigma}_0\right)$ and $(n-1)^{1/2}$vec $\left(\tilde{\boldsymbol{\Sigma}}_1 - \boldsymbol{\Sigma}_1\right)$ are asymptotically distributed as a $p(p+1)/2$-variate normal distribution with mean vector $\mathbf{0}$ and covariance matrices $\boldsymbol{\Phi}_0$ and $\boldsymbol{\Phi}_1$, respectively.

Proof The details of the proof are described in "Appendix B". □

3 Test Statistics and Bootstrap Test

In general, no test in multivariate analysis is uniformly the most powerful. Thus, in this section, we derive the fundamental test statistics, i.e., the LRC, Wald statistic, Rao's score statistic, and gradient statistic for testing the hypothesis

$$H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{O} \text{ versus } H_1 : \boldsymbol{\Sigma}_1 \neq \boldsymbol{O}.$$

Finally, we explain the process of the bootstrap test using these statistics.

3.1 Likelihood Ratio Criterion

Based on the work of Leiva [8], we derive the LRC and the moment of the LRC. Furthermore, we obtain the modified LRC using the moment of the LRC. The likelihood function is

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-nup/2}\left|\boldsymbol{\Sigma}\right|^{-n/2}\exp\left[-\frac{1}{2}\sum_{r=1}^{n}\left(\boldsymbol{x}_r - \boldsymbol{\mu}\right)'\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{x}_r - \boldsymbol{\mu}\right)\right]. \quad (3.1)$$

Since we have assumed that the covariance matrix $\boldsymbol{\Sigma}$ is BCS, the inverse matrix of the covariance matrix $\boldsymbol{\Sigma}$ is

$$\boldsymbol{\Sigma}^{-1} = \boldsymbol{I}_u \otimes \boldsymbol{A} + \boldsymbol{J}_u \otimes \boldsymbol{B},$$

where

$$\begin{aligned}
\boldsymbol{A} &= (\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1)^{-1} = \boldsymbol{\Delta}_1^{-1},\\
\boldsymbol{B} &= \frac{1}{u}\left[\{\boldsymbol{\Sigma}_0 + (u-1)\boldsymbol{\Sigma}_1\}^{-1} - (\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1)^{-1}\right] = \frac{1}{u}\left(\boldsymbol{\Delta}_2^{-1} - \boldsymbol{\Delta}_1^{-1}\right).
\end{aligned}$$

We denote Q_n as the sum of the quadratic forms in (3.1), and can rearrange Q_n as follows:

$$\begin{aligned}
Q_n &= \sum_{r=1}^{n} (\boldsymbol{x}_r - \boldsymbol{\mu})' \, \boldsymbol{\Sigma}^{-1} \, (\boldsymbol{x}_r - \boldsymbol{\mu})\\
&= \sum_{r=1}^{n}\sum_{s=1}^{u} \left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right)' (\boldsymbol{A} + \boldsymbol{B}) \left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right)\\
&\quad + \sum_{r=1}^{n}\sum_{s=1}^{u}\sum_{\substack{s^*=1\\ s\neq s^*}}^{u} \left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right)' \boldsymbol{B} \left(\boldsymbol{x}_{r,s^*} - \boldsymbol{\mu}_{s^*}\right)\\
&= \operatorname{tr}\left[(\boldsymbol{A} + \boldsymbol{B}) \sum_{r=1}^{n}\sum_{s=1}^{u} \left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right)\left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right)'\right]\\
&\quad + \operatorname{tr}\left[\boldsymbol{B} \sum_{r=1}^{n}\sum_{s=1}^{u}\sum_{\substack{s^*=1\\ s\neq s^*}}^{u} \left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right)\left(\boldsymbol{x}_{r,s^*} - \boldsymbol{\mu}_{s^*}\right)'\right].
\end{aligned}$$

Since $\bar{\boldsymbol{x}}_s = \sum_{r=1}^{n} \boldsymbol{x}_{r,s}/n$, we have $\sum_{r=1}^{n}\left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right) = \boldsymbol{0}$. Since

$$\begin{aligned}
&\sum_{r=1}^{n}\sum_{s=1}^{u} \left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right)\left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right)'\\
&\quad = \sum_{r=1}^{n}\sum_{s=1}^{u} \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right)\left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right)' + n\sum_{s=1}^{u} \left(\bar{\boldsymbol{x}}_s - \boldsymbol{\mu}_s\right)\left(\bar{\boldsymbol{x}}_s - \boldsymbol{\mu}_s\right)'\\
&\quad \equiv \boldsymbol{C}_0 + n\sum_{s=1}^{u} \left(\bar{\boldsymbol{x}}_s - \boldsymbol{\mu}_s\right)\left(\bar{\boldsymbol{x}}_s - \boldsymbol{\mu}_s\right)', \qquad (3.2)\\
&\sum_{r=1}^{n}\sum_{s=1}^{u}\sum_{\substack{s^*=1\\ s\neq s^*}}^{u} \left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right)\left(\boldsymbol{x}_{r,s^*} - \boldsymbol{\mu}_{s^*}\right)'
\end{aligned}$$

$$= \sum_{r=1}^{n} \sum_{\substack{s=1 \\ s \neq s^*}}^{u} \sum_{s^*=1}^{u} \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right) \left(\boldsymbol{x}_{r,s^*} - \bar{\boldsymbol{x}}_{s^*}\right)' + n \sum_{\substack{s=1 \\ s \neq s^*}}^{u} \sum_{s^*=1}^{u} + \left(\bar{\boldsymbol{x}}_s - \boldsymbol{\mu}_s\right) \left(\bar{\boldsymbol{x}}_{s^*} - \boldsymbol{\mu}_{s^*}\right)'$$

$$\equiv \boldsymbol{C}_1 + n \sum_{\substack{s=1 \\ s \neq s^*}}^{u} \sum_{s^*=1}^{u} \left(\bar{\boldsymbol{x}}_s - \boldsymbol{\mu}_s\right) \left(\bar{\boldsymbol{x}}_{s^*} - \boldsymbol{\mu}_{s^*}\right)', \tag{3.3}$$

letting $\boldsymbol{\Sigma}_* = \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}$, we can rearrange Q_n as follows:

$$\begin{aligned} Q_n &= \sum_{r=1}^{n} \left(\boldsymbol{x}_r - \boldsymbol{\mu}\right)' \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{x}_r - \boldsymbol{\mu}\right) \\ &= \operatorname{tr}\left(\boldsymbol{\Sigma}_*^{-1} \boldsymbol{C}\right) + \left(\mathbf{1}_n \otimes (\bar{\boldsymbol{x}} - \boldsymbol{\mu})\right)' \boldsymbol{\Sigma}_*^{-1} \left(\mathbf{1}_n \otimes (\bar{\boldsymbol{x}} - \boldsymbol{\mu})\right), \end{aligned} \tag{3.4}$$

where

$$\boldsymbol{C} = \boldsymbol{I}_{nu} \otimes \frac{1}{nu} \left(\boldsymbol{C}_0 - \frac{1}{u-1} \boldsymbol{C}_1\right) + \left(\boldsymbol{I}_n \otimes \boldsymbol{J}_u\right) \otimes \frac{1}{nu(u-1)} \boldsymbol{C}_1.$$

Therefore, Q_n is minimized when $\hat{\boldsymbol{\mu}} = \bar{\boldsymbol{x}}$, and then the log-likelihood function reduces to

$$\log L(\bar{\boldsymbol{x}}, \boldsymbol{\Sigma}_*) = -\frac{nup}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Sigma}_*| - \frac{1}{2} \operatorname{tr}\left(\boldsymbol{\Sigma}_*^{-1} \boldsymbol{C}\right). \tag{3.5}$$

From Lemma 3.2.2 of Anderson [1], the log-likelihood function is maximized when

$$\hat{\boldsymbol{\Sigma}}_* = \boldsymbol{C}.$$

Thus, the maximum of the likelihood function is

$$L(\bar{\boldsymbol{x}}, \hat{\boldsymbol{\Sigma}}_*) = \frac{e^{-nup/2}}{(2\pi)^{nup/2} \left|\hat{\boldsymbol{\Sigma}}\right|^{n/2}}. \tag{3.6}$$

From (3.4), the maximum likelihood estimators of $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$ are

$$\hat{\boldsymbol{\Sigma}}_0 = \frac{1}{nu} \boldsymbol{C}_0 = \frac{1}{nu} \sum_{r=1}^{n} \sum_{s=1}^{u} \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right) \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right)', \tag{3.7}$$

$$\hat{\boldsymbol{\Sigma}}_1 = \frac{1}{nu(u-1)} \boldsymbol{C}_1 = \frac{1}{nu(u-1)} \sum_{r=1}^{n} \sum_{\substack{s=1 \\ s \neq s^*}}^{u} \sum_{s^*=1}^{u} \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right) \left(\boldsymbol{x}_{r,s^*} - \bar{\boldsymbol{x}}_{s^*}\right)'. \tag{3.8}$$

Next, we consider the maximum of the likelihood function under the null hypothesis $H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{O}$. Under H_0, we have

$$\boldsymbol{\Sigma} = \boldsymbol{I}_u \otimes \boldsymbol{\Sigma}_0, \quad \boldsymbol{\Sigma}^{-1} = \boldsymbol{I}_u \otimes \boldsymbol{\Sigma}_0^{-1}, \quad |\boldsymbol{\Sigma}| = |\boldsymbol{\Sigma}_0|^n .$$

Thus, the likelihood function is

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma}_0) = (2\pi)^{-nup/2} |\boldsymbol{\Sigma}|^{-n/2} \exp\left[-\frac{1}{2} \sum_{r=1}^{n} \sum_{s=1}^{u} \left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right)' \boldsymbol{\Sigma}_0^{-1} \left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right) \right]. \tag{3.9}$$

We denote the sum of the quadratic forms in (3.9) as Q, and arrange this as follows:

$$\begin{aligned} Q &= \sum_{r=1}^{n} \sum_{s=1}^{u} \left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right)' \boldsymbol{\Sigma}_0^{-1} \left(\boldsymbol{x}_{r,s} - \boldsymbol{\mu}_s\right) \\ &= \operatorname{tr}\left[\boldsymbol{\Sigma}_0^{-1} \sum_{r=1}^{n} \sum_{s=1}^{u} \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right) \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right)' + n \boldsymbol{\Sigma}_0^{-1} \sum_{s=1}^{u} \left(\bar{\boldsymbol{x}}_s - \boldsymbol{\mu}_s\right) \left(\bar{\boldsymbol{x}}_s - \boldsymbol{\mu}_s\right)' \right] \\ &= \operatorname{tr}\left[\boldsymbol{\Sigma}_0^{-1} \sum_{r=1}^{n} \sum_{s=1}^{u} \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right) \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right)' \right] + n \left(\bar{\boldsymbol{x}} - \boldsymbol{\mu}\right)' \boldsymbol{\Sigma}^{-1} \left(\bar{\boldsymbol{x}} - \boldsymbol{\mu}\right) . \end{aligned}$$

When $\hat{\boldsymbol{\mu}} = \bar{\boldsymbol{x}}$, Q is minimized. Then, the log-likelihood function reduces to

$$\begin{aligned} \log L(\bar{\boldsymbol{x}}, \boldsymbol{\Sigma}_0) = &-\frac{nup}{2} \log(2\pi) - \frac{nu}{2} \log |\boldsymbol{\Sigma}_0| \\ &- \frac{1}{2} \operatorname{tr}\left[\boldsymbol{\Sigma}_0^{-1} \sum_{r=1}^{n} \sum_{s=1}^{u} \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right) \left(\boldsymbol{x}_{r,s} - \bar{\boldsymbol{x}}_s\right)' \right]. \end{aligned}$$

From Lemma 3.2.2 of Anderson [1], the log-likelihood function is maximized when $\hat{\boldsymbol{\Sigma}}_0 = \boldsymbol{C}_0/(nu)$, and the maximum of the likelihood function is

$$L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}_0) = \frac{e^{-nup/2}}{(2\pi)^{nup/2} \left| \hat{\boldsymbol{\Sigma}}_0 \right|^{nu/2}}. \tag{3.10}$$

From the maximums (3.6) and (3.10), the LRC Λ is

$$\Lambda = \frac{\max_{H_0} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\max L(\boldsymbol{\mu}, \boldsymbol{\Sigma})} = \frac{\left| \hat{\boldsymbol{\Sigma}}_0 - \hat{\boldsymbol{\Sigma}}_1 \right|^{n(u-1)/2} \left| \hat{\boldsymbol{\Sigma}}_0 + (u-1) \hat{\boldsymbol{\Sigma}}_1 \right|^{n/2}}{\left| \hat{\boldsymbol{\Sigma}}_0 \right|^{nu/2}}. \tag{3.11}$$

Therefore, we have

$$-2\log\Lambda = nu\log\left|\hat{\boldsymbol{\Sigma}}_0\right| - n(u-1)\log\left|\hat{\boldsymbol{\Sigma}}_0 - \hat{\boldsymbol{\Sigma}}_1\right| - n\log\left|\hat{\boldsymbol{\Sigma}}_0 + (u-1)\hat{\boldsymbol{\Sigma}}_1\right|. \tag{3.12}$$

Next, we obtain the h-th moment of Λ to derive the modified LRC. We express the LRC using $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ as follows:

$$\Lambda = \frac{(nu)^{nup/2}}{\{n(u-1)\}^{n(u-1)p/2}\, n^{np/2}} \cdot \frac{|\boldsymbol{W}_1|^{n(u-1)/2}\,|\boldsymbol{W}_2|^{n/2}}{|\boldsymbol{W}_1 + \boldsymbol{W}_2|^{nu/2}}. \tag{3.13}$$

Letting

$$\lambda = \frac{|\boldsymbol{W}_1|^{n(u-1)/2}\,|\boldsymbol{W}_2|^{n/2}}{|\boldsymbol{W}_1 + \boldsymbol{W}_2|^{nu/2}},$$

the h-th moment of λ is

$$E[\lambda^h] = \frac{\Gamma_p\left[\frac{1}{2}(n-1)u\right]}{\Gamma_p\left[\frac{1}{2}(n-1)u(1+h)\right]} \cdot \frac{\Gamma_p\left[\frac{1}{2}(n-1)(u-1)(1+h)\right]}{\Gamma_p\left[\frac{1}{2}(n-1)(u-1)\right]} \cdot \frac{\Gamma_p\left[\frac{1}{2}(n-1)(1+h)\right]}{\Gamma_p\left[\frac{1}{2}(n-1)\right]}$$

in the same way as in Section 10.4 of Anderson [1]. Since we can write the criterion as

$$\Lambda = \left\{\frac{nu}{n(u-1)}\right\}^{\frac{1}{2}pn(u-1)} \left(\frac{nu}{n}\right)^{\frac{1}{2}pn} \lambda = \left\{\left(\frac{1}{k_1}\right)^{k_1} \left(\frac{1}{k_2}\right)^{k_2}\right\}^{\frac{1}{2}pn} \lambda,$$

where $k_1 = (u-1)/u$ and $k_2 = 1/u$, the h-th moment of Λ is as follows:

$$E[\Lambda^h] = \left\{\left(\frac{1}{k_1}\right)^{k_1} \left(\frac{1}{k_2}\right)^{k_2}\right\}^{\frac{1}{2}pnh} E[\lambda^h].$$

Using the general theory of asymptotic expansions from Section 8.5 of Anderson [1], we have the modified LRC $-2\rho\log\Lambda$, which converges quickly to the chi-squared distribution compared to $-2\log\Lambda$, where

$$\rho = 1 - \frac{u^2 - u + 1}{(n-1)u(u-1)} \cdot \frac{2p^2 + 3p - 1}{6(p+1)}. \tag{3.14}$$

The effect of this modification is confirmed in the simulation described in Sect. 4.

3.2 Wald Statistic

From Theorem 2.4, we can construct the Wald statistic. Since we have $\boldsymbol{\Delta}_1 = \boldsymbol{\Delta}_2 = \boldsymbol{\Sigma}_0$ under the null hypothesis H_0, the asymptotic covariance matrix is

$$\boldsymbol{\Phi}_1 = \frac{1}{u(u-1)} \left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right) (\boldsymbol{\Sigma}_0 \otimes \boldsymbol{\Sigma}_0).$$

Hence, we obtain the following theorem.

Theorem 3.1 *When the null hypothesis H_0 is true, the vector $(n-1)^{1/2}vec\left(\tilde{\boldsymbol{\Sigma}}_1\right)$ is asymptotically distributed as a $p(p+1)/2$-variate normal distribution with mean vector $\mathbf{0}$ and covariance matrix $\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)(\boldsymbol{\Sigma}_0 \otimes \boldsymbol{\Sigma}_0)/\{u(u-1)\}$.*

Noting that

$$\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)^- = \frac{1}{4}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right), \quad \left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right) \mathrm{vec}\left(\tilde{\boldsymbol{\Sigma}}_1\right) = 2\mathrm{vec}\left(\tilde{\boldsymbol{\Sigma}}_1\right),$$

using Theorem 3.1, the Wald statistic

$$W = \frac{(n-1)u(u-1)}{2} \mathrm{vec}'\left(\tilde{\boldsymbol{\Sigma}}_1\right)\left(\tilde{\boldsymbol{\Sigma}}_0^{-1} \otimes \tilde{\boldsymbol{\Sigma}}_0^{-1}\right) \mathrm{vec}\left(\tilde{\boldsymbol{\Sigma}}_1\right) \tag{3.15}$$

is asymptotically distributed as a chi-squared distribution with $p(p+1)/2$ degrees of freedom, where $\boldsymbol{A}^-$ denotes the Moore–Penrose inverse matrix of $\boldsymbol{A}$.

3.3 Rao's Score Statistic

Assuming the BCS covariance structure, the log-likelihood function (3.5) is represented as follows:

$$\begin{aligned}\log L(\bar{\boldsymbol{x}}, \boldsymbol{\Sigma}_*) = &-\frac{nup}{2}\log(2\pi) - \frac{n}{2}\log|\boldsymbol{\Sigma}| \\ &- \frac{1}{2}\mathrm{tr}\left[(\boldsymbol{I}_n \otimes \boldsymbol{I}_u \otimes \boldsymbol{A})\boldsymbol{C}\right] - \frac{1}{2}\mathrm{tr}\left\{(\boldsymbol{I}_n \otimes \boldsymbol{J}_u \otimes \boldsymbol{B})\boldsymbol{C}\right\}.\end{aligned} \tag{3.16}$$

Details are given in "Appendix C", but the derivative of the log-likelihood function with respect to $\boldsymbol{\Sigma}_1$ is

$$\begin{aligned}\boldsymbol{U}(\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2) &= \frac{\partial}{\partial \boldsymbol{\Sigma}_1}\log L(\bar{\boldsymbol{x}}, \boldsymbol{\Sigma}_*) \\ &= -\frac{n(u-1)}{2}\left\{\boldsymbol{\Delta}_1^{-1}\left(\hat{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right)\boldsymbol{\Delta}_1^{-1}\right\} \\ &\quad + \frac{n(u-1)}{2}\left\{\boldsymbol{\Delta}_2^{-1}\left(\hat{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right)\boldsymbol{\Delta}_2^{-1}\right\}.\end{aligned} \tag{3.17}$$

From this result, the information matrix is as follows:

$$\begin{aligned}\boldsymbol{I}(\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2) &= E\left[\mathrm{vec}\,(\boldsymbol{U}(\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2))\,\mathrm{vec}'\,(\boldsymbol{U}(\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2))\right] \\ &= \frac{(n-1)(u-1)}{4}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)\left(\boldsymbol{\Delta}_1^{-1} \otimes \boldsymbol{\Delta}_1^{-1}\right) \\ &\quad + \frac{(n-1)(u-1)^2}{4}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)\left(\boldsymbol{\Delta}_2^{-1} \otimes \boldsymbol{\Delta}_2^{-1}\right). \end{aligned} \tag{3.18}$$

Let $\check{\boldsymbol{\Delta}}_1$ and $\check{\boldsymbol{\Delta}}_2$ be MLEs of $\boldsymbol{\Delta}_1$ and $\boldsymbol{\Delta}_2$, respectively, under the null hypothesis H_0. When the null hypothesis H_0 is true, we have

$$\check{\boldsymbol{\Delta}}_1 = \check{\boldsymbol{\Delta}}_2 = \hat{\boldsymbol{\Sigma}}_0.$$

Since the score $\mathrm{vec}\left(\boldsymbol{U}(\check{\boldsymbol{\Delta}}_1, \check{\boldsymbol{\Delta}}_2)\right)$ is

$$\mathrm{vec}\left(\boldsymbol{U}(\check{\boldsymbol{\Delta}}_1, \check{\boldsymbol{\Delta}}_2)\right) = \frac{nu(u-1)}{2}\left(\hat{\boldsymbol{\Sigma}}_0^{-1} \otimes \hat{\boldsymbol{\Sigma}}_0^{-1}\right)\mathrm{vec}\left(\hat{\boldsymbol{\Sigma}}_1\right), \tag{3.19}$$

Rao's score statistic is

$$S = \frac{nu(u-1)}{2}\mathrm{vec}'\left(\hat{\boldsymbol{\Sigma}}_1\right)\left(\hat{\boldsymbol{\Sigma}}_0^{-1} \otimes \hat{\boldsymbol{\Sigma}}_0^{-1}\right)\mathrm{vec}\left(\hat{\boldsymbol{\Sigma}}_1\right). \tag{3.20}$$

Using the score (3.19) under the null hypothesis H_0, we find that the gradient statistic is the same as Rao's score statistic.

3.4 Bootstrap Test

Following Efron and Tibshirani [4], we perform the bootstrap test using the criteria $-2\log\Lambda$, $-2\rho\log\Lambda$, W, and S as follows:

(i) Calculate the mean vector $\bar{\boldsymbol{x}}$, the unbiased covariance matrix $\tilde{\boldsymbol{\Sigma}}_0$, and the criteria $-2\log\Lambda_x$, $-2\rho\log\Lambda_x$, W_x, and S_x from the original sample $\boldsymbol{x}$.
(ii) Form B bootstrap datasets $\boldsymbol{y}$ of size n from the normal population $N(\bar{\boldsymbol{x}}, \boldsymbol{I}_u \otimes \tilde{\boldsymbol{\Sigma}}_0)$.
(iii) Evaluate the criteria $-2\log\Lambda_y$, $-2\rho\log\Lambda_y$, W_y, and S_y from each dataset $\boldsymbol{y}$.
(iv) Approximate an achieved significance level (ASL) as:

$$\widehat{ASL}_1 = \frac{\#\left\{-2\log\Lambda_y > -2\log\Lambda_x\right\}}{B},$$
$$\widehat{ASL}_2 = \frac{\#\left\{-2\rho\log\Lambda_y > -2\rho\log\Lambda_x\right\}}{B},$$
$$\widehat{ASL}_3 = \frac{\#\left\{W_y > W_x\right\}}{B}, \quad \widehat{ASL}_4 = \frac{\#\left\{S_y > S_x\right\}}{B}.$$

If the value of $\widehat{ASL}$ is less than the significance level α, we reject the null hypothesis.

We use the bootstrap test in our simulations because it can be allied to hypothesis testing using these statistics, and the ASLs of the bootstrap test are guaranteed to be accurate as the sample size becomes large.

4 Numerical Example

In this section, we investigate the accuracy of the test using the above criteria and apply them to real data. The simulation uses 100,000 samples.

4.1 Numerical Simulation

First, we investigate the accuracy of the significance level for the test using the criteria $-2\log\Lambda$, $-2\rho\log\Lambda$, W, and S under the null hypothesis. Letting

$$\boldsymbol{\Sigma}_0 = \sigma^2 \begin{pmatrix} 1 & \varrho & \cdots & \varrho^{p-1} \\ \varrho & 1 & \cdots & \varrho^{p-2} \\ \vdots & \vdots & & \vdots \\ \varrho^{p-1} & \varrho^{p-2} & \cdots & 1 \end{pmatrix}, \tag{4.1}$$

where $\sigma = 2$ and $\varrho = 0.5$, we set the population distribution such that the mean vector $\boldsymbol{\mu}$ is the zero vector and the covariance matrix is

$$\boldsymbol{\Sigma} = \boldsymbol{I}_u \otimes \boldsymbol{\Sigma}_0.$$

We change the dimension p and the number u of sites, and set the sample size n for each case. Table 1 presents the ASLs using the 95th percentile of the chi-squared distribution.

The results show that the ASLs of the likelihood ratio test and the modified likelihood ratio test are greater than 0.05, meaning that these tests fail to control the significance level. In contrast, the ASL of the Wald test is less than 0.05 and the Rao's score test retains the significance level. We have found that the correction using ρ improves the convergence to the significance level.

We consider the bootstrap test using these test statistics because their ASLs are different.

Table 2 presents the ASL for the bootstrap test using these statistics for the significance level $\alpha = 0.05$. The number of bootstrap replications is 1000. The results show that the ASLs of the bootstrap test using $-2\log\Lambda$, $-2\rho\log\Lambda$, and W are greater than 0.05 and the ASL of the bootstrap test using S is less than 0.05. We have found that the bootstrap test is dominant in terms of ensuring the stability of the significance level. When the sample size is large, the bootstrap test using $-2\rho\log\Lambda$ or W retains the significance level.

Next, we investigate the power of the test in two cases. We set the sample size n, dimension p, and number u of sites as for the situation under the null hypothesis. Since the convergence of each statistic to the significance level is different, we cannot make

Table 1 Achieved significance level of normal test for $\alpha = 0.05$

Parameters				Normal test			
n	p	u	ρ	$-2\log\Lambda$	$-2\rho\log\Lambda$	W	S
20	3	2	.9145	.0900	.0630	.0259	.0356
40	3	2	.9583	.0676	.0560	.0372	.0430
60	3	2	.9725	.0606	.0536	.0413	.0452
40	5	3	.9468	.0844	.0603	.0437	.0521
80	5	3	.9737	.0643	.0541	.0459	.0504
120	5	3	.9826	.0607	.0542	.0485	.0514
160	5	3	.9870	.0559	.0511	.0465	.0486
100	9	3	.9631	.0804	.0546	.0444	.0496
200	9	3	.9816	.0634	.0522	.0472	.0500
300	9	3	.9878	.0582	.0510	.0475	.0492
400	9	3	.9908	.0558	.0506	.0485	.0498

Table 2 Achieved significance level of bootstrap test for $\alpha = 0.05$

Parameters				Bootstrap test			
n	p	u	ρ	$-2\log\Lambda$	$-2\rho\log\Lambda$	W	S
20	3	2	.9145	.0505	.0505	.0506	.0381
40	3	2	.9583	.0507	.0507	.0511	.0446
60	3	2	.9725	.0507	.0507	.0508	.0465
40	5	3	.9468	.0521	.0521	.0527	.0441
80	5	3	.9737	.0502	.0502	.0512	.0467
120	5	3	.9826	.0521	.0521	.0521	.0493
160	5	3	.9870	.0495	.0495	.0494	.0475
100	9	3	.9631	.0498	.0498	.0501	.0450
200	9	3	.9816	.0499	.0499	.0504	.0478
300	9	3	.9878	.0498	.0498	.0504	.0487
400	9	3	.9908	.0504	.0504	.0505	.0493

a simple comparison of the powers of the test, but instead compare the powers of the bootstrap test by taking the convergence to the significance level into consideration. Let $\boldsymbol{\Sigma}_0$ be as in (4.1), and consider Case 1: $\boldsymbol{\Sigma}_1 = \tau_1 \boldsymbol{I}_p$ and Case 2: $\boldsymbol{\Sigma}_1 = \tau_2 \mathbf{1}_p \mathbf{1}_p'$. We set τ_1 and τ_2 as shown in the following table.

	$p = 3$	$p = 5$	$p = 9$
τ_1	0.8	0.3	0.15
τ_2	1.3	0.5	0.2

Table 3 Power of bootstrap test (Significance level: $\alpha = 0.05$)

Parameters				Case 1			
n	p	u	ρ	$-2\log\Lambda$	$-2\rho\log\Lambda$	W	S
20	3	2	.9145	.3965	.3965	.3714	.3166
40	3	2	.9583	.8007	.8007	.7919	.7746
60	3	2	.9725	.9544	.9544	.9526	.9491
40	5	3	.9468	.3894	.3894	.5566	.5289
80	5	3	.9737	.7796	.7796	.8623	.8544
120	5	3	.9826	.9451	.9451	.9689	.9673
160	5	3	.9870	.9901	.9901	.9948	.9945
100	9	3	.9631	.3030	.3030	.4788	.4594
200	9	3	.9816	.6711	.6711	.7945	.7880
300	9	3	.9878	.8976	.8976	.9454	.9438
400	9	3	.9908	.9768	.9768	.9888	.9885
Parameters				Case 2			
n	p	u	ρ	$-2\log\Lambda$	$-2\rho\log\Lambda$	W	S
20	3	2	.9145	.3958	.3958	.3610	.3049
40	3	2	.9583	.7924	.7924	.7749	.7561
60	3	2	.9725	.9493	.9493	.9450	.9406
40	5	3	.9468	.3827	.3827	.4963	.4679
80	5	3	.9737	.7574	.7574	.8274	.8187
120	5	3	.9826	.9312	.9312	.9554	.9537
160	5	3	.9870	.9851	.9851	.9910	.9907
100	9	3	.9631	.2326	.2326	.3083	.2906
200	9	3	.9816	.5177	.5177	.6071	.5981
300	9	3	.9878	.7635	.7635	.8246	.8207
400	9	3	.9908	.9071	.9071	.9372	.9359

Since the alternative hypothesis, the population covariance matrix is

$$\boldsymbol{\Sigma} = \boldsymbol{I}_u \otimes (\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1) + \boldsymbol{J}_u \otimes \boldsymbol{\Sigma}_1.$$

The upper part of Table 3 presents the powers of the test in Case 1. Since the criteria $-2\log\Lambda$ and $-2\rho\log\Lambda$ are essentially the same, the powers of the bootstrap test using these criteria are equal. When the dimension is high, the power of the bootstrap test using W is largest followed by the power of the bootstrap test using S. The powers of the modified likelihood ratio test are the largest when the dimension is low.

The lower part of Table 3 presents the powers of the test in Case 2. The same tendencies as in Case 1 can be observed. The power of the bootstrap test using W is largest, followed by the power of the bootstrap test using S; the power of the bootstrap test using the modified LRC is the third largest when the dimension is high. The powers of the modified likelihood ratio test are largest when the dimension is low, but the powers of the bootstrap test using the modified LRC, W and S are almost the same in the case of a low dimension and large sample.

4.2 Example Using Real Data

We apply hypothesis testing using real data taken from Johnson and Wichern [7]. To examine whether dietary supplements would slow bone loss in 25 older women, the mineral content of bones (radius, humerus, and ulna) was measured by photon absorptiometry. Measurements were recorded for three bones on the dominant and non-dominant sides, i.e., $p = 3$ and $u = 2$. Roy and Leiva [10] demonstrated that the data fail to reject the null hypothesis that the covariance structure is of BCS form (p-value = 0.5786). The unbiased estimator for $\boldsymbol{\mu}$ is

$$(0.8438,\ 1.7927,\ 0.7044,\ 0.8183,\ 1.7348,\ 0.6938)',$$

and the unbiased estimators for $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$ are

$$\tilde{\boldsymbol{\Sigma}}_0 = \begin{pmatrix} 0.0122 & 0.0217 & 0.0090 \\ 0.0217 & 0.0749 & 0.0168 \\ 0.0090 & 0.0168 & 0.0111 \end{pmatrix}, \quad \tilde{\boldsymbol{\Sigma}}_1 = \begin{pmatrix} 0.0104 & 0.0193 & 0.0082 \\ 0.0193 & 0.0668 & 0.0153 \\ 0.0082 & 0.0153 & 0.0081 \end{pmatrix}.$$

The maximum likelihood estimators are

$$\hat{\boldsymbol{\Sigma}}_0 = \begin{pmatrix} 0.0117 & 0.0209 & 0.0087 \\ 0.0209 & 0.0719 & 0.0161 \\ 0.0087 & 0.0161 & 0.0106 \end{pmatrix}, \quad \hat{\boldsymbol{\Sigma}}_1 = \begin{pmatrix} 0.0100 & 0.0185 & 0.0079 \\ 0.0185 & 0.0641 & 0.0147 \\ 0.0079 & 0.0147 & 0.0077 \end{pmatrix}.$$

Noting that $\rho = 0.9323$, the criteria are

$$\begin{aligned} -2\log\Lambda &= 71.7279, \quad -2\rho\log\Lambda = 66.8713, \\ W_1 &= 38.3102, \quad W_2 = 24.8056, \quad S = 39.9065. \end{aligned}$$

Since the upper 5% point of the chi-squared distribution with 6 degrees of freedom is 12.5916, we reject the null hypothesis $\boldsymbol{\Sigma}_1 = \boldsymbol{O}$ with a significance level 0.05. We also applied the bootstrap test using the same criteria. The ASL values for each statistic are approximately 0.0000, and the result is the same as for the previous test.

5 Conclusions

We have treated hypothesis testing for independence under the BCS covariance structure. The LRC, modified LRC, Wald statistic, and Rao's score statistic have been derived. We have shown that the test using these statistics is effective in specific situations. In particular, we found that the bootstrap test is superior in terms of convergence to the significance level, that the power of the bootstrap test using the Wald statistic is largest when the dimension is high, that the power of the bootstrap test using the modified LRC is largest when the dimension is low, and that the power of the bootstrap test using the Wald statistic is the same as the power of the bootstrap test using the modified LRC when the dimension is low and the sample size is large. We recommend the bootstrap test using the Wald statistic.

Recently, high-dimensional multivariate analysis has been extensively studied (see Fujikoshi and Ulyanov [5] and Pourahmadi [9]). It may also be possible to study hypothesis testing for independence under the BCS covariance structure under high-dimensional situations ($up > n$). However, we cannot employ statistics using the determinant, such as the LRC, because the matrices $\boldsymbol{W}_i$ are singular under high-dimensional conditions. Thus, it is necessary to consider new test statistics using the trace of $\boldsymbol{W}_i$ for hypothesis testing, which is left as a future problem.

Acknowledgements The author thanks Stuart Jenkinson, Ph.D., from Edanz Group (www.edanz-editing.com/ac) for editing a draft of this manuscript, and is grateful to three anonymous referees for comments to revise the original manuscript.

Appendix

A Proof of Lemmas 2.2 and 2.3

First, we show three lemmas to derive the exact distribution of $\tilde{\boldsymbol{\Sigma}}_0$ and $\tilde{\boldsymbol{\Sigma}}_1$.

Lemma 5.1 (Theorem 3.3.1 in Gupta and Nagar [6]) *When $a > 0$ and $\boldsymbol{S} \sim W_p(n, \boldsymbol{\Sigma})$, we have $a\boldsymbol{S} \sim W_p(n, a\boldsymbol{\Sigma})$.*

Lemma 5.2 (p. 127 in Gupta and Nagar [6]) *When $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are independent of each other, $\boldsymbol{S}_1 \sim W_p(n_1, \boldsymbol{\Sigma}_1)$, and $\boldsymbol{S}_2 \sim W_p(n_2, \boldsymbol{\Sigma}_2)$, the distribution of $\boldsymbol{P} = \boldsymbol{S}_1 + \boldsymbol{S}_2$ is as follows:*

$$\left\{2^{(n_1+n_2)p/2}\Gamma_p\left[\frac{1}{2}(n_1+n_2)\right]|\boldsymbol{\Sigma}_1|^{n_1/2}|\boldsymbol{\Sigma}_2|^{n_2/2}\right\}^{-1}$$
$$\times \mathrm{etr}\left[-\frac{1}{2}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{P}\right]|\boldsymbol{P}|^{(n_1+n_2)/2-p-1}\,{}_1F_1\left[\frac{1}{2}n_2; \frac{1}{2}(n_1+n_2); \frac{1}{2}\left(\boldsymbol{\Sigma}_1^{-1}-\boldsymbol{\Sigma}_2^{-1}\right)\boldsymbol{P}\right],$$

where

$${}_1F_1\left[a; b; \boldsymbol{H}\right] = \frac{\Gamma_p(b)}{\Gamma_p(a)\Gamma_p(b-a)}\int_{\boldsymbol{O}<\boldsymbol{Y}<\boldsymbol{I}_p}\mathrm{etr}(\boldsymbol{Y}\boldsymbol{H})|\boldsymbol{Y}|^{a-(p+1)/2}|\boldsymbol{I}_p-\boldsymbol{Y}|^{b-a-(p+1)/2}d\boldsymbol{Y}. \quad (5.1)$$

Proof Letting $\boldsymbol{P} = \boldsymbol{S}_1 + \boldsymbol{S}_2$ and $\boldsymbol{Q} = \boldsymbol{S}_2$, we transform the simultaneous density function of $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ into the simultaneous density function of $\boldsymbol{P}$ and $\boldsymbol{Q}$. We obtain the distribution of $\boldsymbol{P}$ by integrating the simultaneous density function of $\boldsymbol{P}$ and $\boldsymbol{Q}$ with respect to $\boldsymbol{Q}$. □

Lemma 5.3 *When $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are independent of each other, $\boldsymbol{S}_1 \sim W_p(n_1, \boldsymbol{\Sigma}_1)$, and $\boldsymbol{S}_2 \sim W_p(n_2, \boldsymbol{\Sigma}_2)$, the distribution of $\boldsymbol{M} = \boldsymbol{S}_1 - \boldsymbol{S}_2$ is as follows*:

$$\left\{2^{(n_1+n_2)p/2}\Gamma_p\left[\frac{1}{2}n_1\right]|\boldsymbol{\Sigma}_1|^{\frac{1}{2}n_1}\,|\boldsymbol{\Sigma}_2|^{\frac{1}{2}n_2}\right\}^{-1}$$
$$\times \operatorname{etr}\left[-\frac{1}{2}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{M}\right]|\boldsymbol{M}|^{(n_1+n_2)/2-p-1}\,\Psi\left[\frac{1}{2}n_2,\frac{1}{2}(n_1+n_2);\frac{1}{2}\left(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_2^{-1}\right)\boldsymbol{M}\right],$$

where

$$\Psi[a, c; \boldsymbol{R}] = \frac{1}{\Gamma_p\,[a]}\int_{\boldsymbol{S}>0}\operatorname{etr}(-\boldsymbol{R}\boldsymbol{S})\,|\boldsymbol{S}|^{a-(p+1)/2}\left|\boldsymbol{I}_p+\boldsymbol{S}\right|^{c-a-(p+1)/2}d\boldsymbol{S}. \tag{5.2}$$

Proof Letting $\boldsymbol{M} = \boldsymbol{S}_1 - \boldsymbol{S}_2$ and $\boldsymbol{Q} = \boldsymbol{S}_2$, we transform the simultaneous density function of $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ into the simultaneous density function of $\boldsymbol{M}$ and $\boldsymbol{Q}$. We obtain the distribution of $\boldsymbol{M}$ by integrating the simultaneous density function of $\boldsymbol{M}$ and $\boldsymbol{Q}$ with respect to $\boldsymbol{Q}$. □

We derive the distribution of $\tilde{\boldsymbol{\Sigma}}_0$. We have

$$\tilde{\boldsymbol{\Sigma}}_0 = \frac{1}{(n-1)u}\boldsymbol{W}_1 + \frac{1}{(n-1)u}\boldsymbol{W}_2,$$

and

$$\frac{1}{(n-1)u}\boldsymbol{W}_1 \sim W_p\left((n-1)(u-1), \frac{1}{(n-1)u}\boldsymbol{\Delta}_1\right),$$
$$\frac{1}{(n-1)u}\boldsymbol{W}_2 \sim W_p\left(n-1, \frac{1}{(n-1)u}\boldsymbol{\Delta}_2\right),$$

from Lemma 5.1. From Lemma 5.2, the distribution of $\tilde{\boldsymbol{\Sigma}}_0 = (\boldsymbol{W}_1 + \boldsymbol{W}_2)/\{(n-1)u\}$ is

$$\left\{2^{u(n-1)p/2}\Gamma_p\left[\frac{1}{2}u(n-1)\right]\left|\frac{1}{(n-1)u}\boldsymbol{\Delta}_1\right|^{(n-1)(u-1)/2}\left|\frac{1}{(n-1)u}\boldsymbol{\Delta}_2\right|^{(n-1)/2}\right\}^{-1}$$
$$\times \operatorname{etr}\left[-\frac{(n-1)u}{2}\boldsymbol{\Delta}_1^{-1}\tilde{\boldsymbol{\Sigma}}_0\right]\left|\tilde{\boldsymbol{\Sigma}}_0\right|^{u(n-1)/2-p-1}$$
$$\times\, {}_1F_1\left[\frac{1}{2}(n-1);\frac{1}{2}(n-1)u;\frac{1}{2}\left\{(n-1)u\boldsymbol{\Delta}_1^{-1}-(n-1)u\boldsymbol{\Delta}_2^{-1}\right\}\tilde{\boldsymbol{\Sigma}}_0\right].$$

Similarly, we have

$$\tilde{\boldsymbol{\Sigma}}_1 = \frac{1}{(n-1)u}\boldsymbol{W}_2 - \frac{1}{(n-1)u(u-1)}\boldsymbol{W}_1,$$

and

$$\frac{1}{(n-1)u}\boldsymbol{W}_2 \sim W_p\left(n-1, \frac{1}{(n-1)u}\boldsymbol{\Delta}_2\right),$$
$$\frac{1}{(n-1)u(u-1)}\boldsymbol{W}_1 \sim W_p\left((n-1)(u-1), \frac{1}{(n-1)u(u-1)}\boldsymbol{\Delta}_1\right),$$

from Lemma 5.1. From Lemma 5.3, the distribution of $\tilde{\boldsymbol{\Sigma}}_1 = \boldsymbol{W}_2/\{(n-1)u\} - \boldsymbol{W}_1/\{(n-1)u(u-1)\}$ is

$$\left\{2^{u(n-1)p/2}\Gamma_p\left[\frac{1}{2}(n-1)\right]\left|\frac{1}{(n-1)u}\boldsymbol{\Delta}_2\right|^{(n-1)/2}\left|\frac{1}{(n-1)u(u-1)}\boldsymbol{\Delta}_1\right|^{(n-1)(u-1)/2}\right\}^{-1}$$
$$\times \operatorname{etr}\left[-\frac{(n-1)u}{2}\boldsymbol{\Delta}_2^{-1}\tilde{\boldsymbol{\Sigma}}_1\right]\left|\tilde{\boldsymbol{\Sigma}}_1\right|^{u(n-1)/2-p-1}$$
$$\times \Psi\left[\frac{1}{2}(n-1)(u-1); \frac{1}{2}(n-1)u; \frac{1}{2}\left\{(n-1)u\boldsymbol{\Delta}_2^{-1} + (n-1)u(u-1)\boldsymbol{\Delta}_1^{-1}\right\}\tilde{\boldsymbol{\Sigma}}_1\right].$$

B Covariance Matrix of the Unbiased Estimator

From the result of Roy et al. [11], we have

$$\tilde{\boldsymbol{\Delta}}_1 \sim W_p\left((n-1)(u-1), \frac{1}{(n-1)(u-1)}\boldsymbol{\Delta}_1\right),$$
$$\tilde{\boldsymbol{\Delta}}_2 \sim W_p\left(n-1, \frac{1}{n-1}\boldsymbol{\Delta}_2\right).$$

Consequently, we have

$$E\left[\operatorname{vec}\left(\tilde{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right)\operatorname{vec}'\left(\tilde{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right)\right] = \frac{1}{(n-1)(u-1)}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)(\boldsymbol{\Delta}_1 \otimes \boldsymbol{\Delta}_1),$$
$$E\left[\operatorname{vec}\left(\tilde{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right)\operatorname{vec}'\left(\tilde{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right)\right] = \frac{1}{n-1}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)(\boldsymbol{\Delta}_2 \otimes \boldsymbol{\Delta}_2).$$

First, we calculate the covariance matrix of $\operatorname{vec}\left(\tilde{\boldsymbol{\Sigma}}_0 - \boldsymbol{\Sigma}_0\right)$. Since

$$\operatorname{vec}\left(\tilde{\boldsymbol{\Sigma}}_0 - \boldsymbol{\Sigma}_0\right) = \frac{u-1}{u}\operatorname{vec}\left(\tilde{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right) + \frac{1}{u}\operatorname{vec}\left(\tilde{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right),$$

the covariance matrix of $\operatorname{vec}\left(\tilde{\boldsymbol{\Sigma}}_0 - \boldsymbol{\Sigma}_0\right)$ is as follows:

$$E\left[\operatorname{vec}\left(\tilde{\boldsymbol{\Sigma}}_0 - \boldsymbol{\Sigma}_0\right)\operatorname{vec}'\left(\tilde{\boldsymbol{\Sigma}}_0 - \boldsymbol{\Sigma}_0\right)\right]$$

$$= \frac{u-1}{(n-1)u^2}\left(\boldsymbol{I}_{p^2}+\boldsymbol{K}_{p,p}\right)(\boldsymbol{\Delta}_1\otimes\boldsymbol{\Delta}_1)+\frac{1}{(n-1)u^2}\left(\boldsymbol{I}_{p^2}+\boldsymbol{K}_{p,p}\right)(\boldsymbol{\Delta}_2\otimes\boldsymbol{\Delta}_2)\,. \tag{5.3}$$

Similarly, since

$$\mathrm{vec}\left(\tilde{\boldsymbol{\Sigma}}_1-\boldsymbol{\Sigma}_1\right)=\frac{1}{u}\mathrm{vec}\left(\tilde{\boldsymbol{\Delta}}_2-\boldsymbol{\Delta}_2\right)-\frac{1}{u}\mathrm{vec}\left(\tilde{\boldsymbol{\Delta}}_1-\boldsymbol{\Delta}_1\right),$$

the covariance matrix of $\mathrm{vec}\left(\tilde{\boldsymbol{\Sigma}}_1-\boldsymbol{\Sigma}_1\right)$ is as follows:

$$\begin{aligned} &E\left[\mathrm{vec}\left(\tilde{\boldsymbol{\Sigma}}_1-\boldsymbol{\Sigma}_1\right)\mathrm{vec}'\left(\tilde{\boldsymbol{\Sigma}}_1-\boldsymbol{\Sigma}_1\right)\right]\\ &=\frac{1}{(n-1)u^2(u-1)}\left(\boldsymbol{I}_{p^2}+\boldsymbol{K}_{p,p}\right)(\boldsymbol{\Delta}_1\otimes\boldsymbol{\Delta}_1)\\ &\quad+\frac{1}{(n-1)u^2}\left(\boldsymbol{I}_{p^2}+\boldsymbol{K}_{p,p}\right)(\boldsymbol{\Delta}_2\otimes\boldsymbol{\Delta}_2)\,. \end{aligned} \tag{5.4}$$

C The Score and the Information Matrix

Assuming the BCS covariance structure, the log-likelihood function (3.5) is represented as follows:

$$\begin{aligned} \log L(\bar{\boldsymbol{x}},\boldsymbol{\Sigma}_*) &= -\frac{nup}{2}\log(2\pi)-\frac{n}{2}\log|\boldsymbol{\Sigma}|\\ &\quad-\frac{1}{2}\mathrm{tr}\left[(\boldsymbol{I}_n\otimes\boldsymbol{I}_u\otimes\boldsymbol{A})\,\boldsymbol{C}\right]-\frac{1}{2}\mathrm{tr}\left\{(\boldsymbol{I}_n\otimes\boldsymbol{J}_u\otimes\boldsymbol{B})\,\boldsymbol{C}\right\}, \end{aligned} \tag{5.5}$$

where

$$\begin{aligned} \boldsymbol{A} &= (\boldsymbol{\Sigma}_0-\boldsymbol{\Sigma}_1)^{-1}=\boldsymbol{\Delta}_1^{-1},\\ \boldsymbol{B} &= \frac{1}{u}\left[\{\boldsymbol{\Sigma}_0+(u-1)\boldsymbol{\Sigma}_1\}^{-1}-(\boldsymbol{\Sigma}_0-\boldsymbol{\Sigma}_1)^{-1}\right]=\frac{1}{u}\left(\boldsymbol{\Delta}_2^{-1}-\boldsymbol{\Delta}_1^{-1}\right). \end{aligned}$$

We show the following lemma used to derive the score function.

Lemma 5.4 *Let $\boldsymbol{X}$ be a $p\times p$ matrix and $\boldsymbol{H}$ be a $p\times p$ constant matrix. Then, we have*

(1) $\dfrac{d}{d\boldsymbol{X}}\log|\boldsymbol{X}|=(\boldsymbol{X}^{-1})'$,

(2) $\dfrac{d}{d\boldsymbol{X}}tr(\boldsymbol{X}^{-1}\boldsymbol{H})=-(\boldsymbol{X}^{-1}\boldsymbol{H}\boldsymbol{X}^{-1})'$.

Since the second term of the log-likelihood function (5.5) can be rewritten as:

$$-\frac{n}{2}\log|\boldsymbol{\Sigma}|=-\frac{n}{2}\log|\boldsymbol{\Delta}_1|^{u-1}|\boldsymbol{\Delta}_2|=-\frac{n(u-1)}{2}\log|\boldsymbol{\Delta}_1|-\frac{n}{2}\log|\boldsymbol{\Delta}_2|\,,$$

using Lemma 5.4 (1), we have

$$\frac{\partial}{\partial \boldsymbol{\Sigma}_1}\left(-\frac{n}{2}\log|\boldsymbol{\Sigma}|\right) = \frac{n(u-1)}{2}\boldsymbol{\Delta}_1^{-1} - \frac{n(u-1)}{2}\boldsymbol{\Delta}_2^{-1}. \tag{5.6}$$

We can rewrite the third term of the log-likelihood function (5.5) as follows:

$$-\frac{1}{2}\mathrm{tr}\left[(\boldsymbol{I}_n \otimes \boldsymbol{I}_u \otimes \boldsymbol{A})\,\boldsymbol{C}\right] = -\frac{nu}{2}\mathrm{tr}\left(\boldsymbol{A}\hat{\boldsymbol{\Sigma}}_0\right),$$

and so Lemma 5.4 (2) implies that

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{\Sigma}_1}\left[-\frac{1}{2}\mathrm{tr}\left[(\boldsymbol{I}_n \otimes \boldsymbol{I}_u \otimes \boldsymbol{A})\,\boldsymbol{C}\right]\right] &= \frac{\partial}{\partial \boldsymbol{\Sigma}_1}\left[-\frac{nu}{2}\mathrm{tr}\left(\boldsymbol{\Delta}_1^{-1}\hat{\boldsymbol{\Sigma}}_0\right)\right] \\ &= -\frac{nu}{2}\left(\boldsymbol{\Delta}_1^{-1}\hat{\boldsymbol{\Sigma}}_0\boldsymbol{\Delta}_1^{-1}\right).\end{aligned} \tag{5.7}$$

We can rewrite the fourth term of log-likelihood function (5.5) as follows:

$$\begin{aligned}\mathrm{tr}\left\{(\boldsymbol{I}_n \otimes \boldsymbol{J}_u \otimes \boldsymbol{B})\,\boldsymbol{C}\right\} &= n\mathrm{tr}\left[(\boldsymbol{J}_u \otimes \boldsymbol{B})\,\hat{\boldsymbol{\Sigma}}\right] = nu\mathrm{tr}\left[\boldsymbol{B}\hat{\boldsymbol{\Sigma}}_0 + (u-1)\boldsymbol{B}\hat{\boldsymbol{\Sigma}}_1\right] \\ &= n\mathrm{tr}\left(\boldsymbol{\Delta}_2^{-1}\hat{\boldsymbol{\Sigma}}_0\right) - n\mathrm{tr}\left(\boldsymbol{\Delta}_1^{-1}\hat{\boldsymbol{\Sigma}}_0\right) + n(u-1)\mathrm{tr}\left(\boldsymbol{\Delta}_2^{-1}\hat{\boldsymbol{\Sigma}}_1\right) - n(u-1)\mathrm{tr}\left(\boldsymbol{\Delta}_1^{-1}\hat{\boldsymbol{\Sigma}}_1\right).\end{aligned}$$

Therefore, we have

$$\begin{aligned}&\frac{\partial}{\partial \boldsymbol{\Sigma}_1}\left[-\frac{1}{2}\mathrm{tr}\left\{(\boldsymbol{I}_n \otimes \boldsymbol{J}_u \otimes \boldsymbol{B})\,\boldsymbol{C}\right\}\right] \\ &\quad= \frac{n(u-1)}{2}\boldsymbol{\Delta}_2^{-1}\hat{\boldsymbol{\Sigma}}_0\boldsymbol{\Delta}_2^{-1} + \frac{n}{2}\boldsymbol{\Delta}_1^{-1}\hat{\boldsymbol{\Sigma}}_0\boldsymbol{\Delta}_1^{-1} \\ &\qquad+ \frac{n(u-1)^2}{2}\boldsymbol{\Delta}_2^{-1}\hat{\boldsymbol{\Sigma}}_1\boldsymbol{\Delta}_2^{-1} + \frac{n(u-1)}{2}\boldsymbol{\Delta}_1^{-1}\hat{\boldsymbol{\Sigma}}_1\boldsymbol{\Delta}_1^{-1} \\ &\quad= \frac{n(u-1)}{2}\boldsymbol{\Delta}_2^{-1}\left\{\hat{\boldsymbol{\Sigma}}_0 + (u-1)\hat{\boldsymbol{\Sigma}}_1\right\}\boldsymbol{\Delta}_2^{-1} + \frac{n}{2}\boldsymbol{\Delta}_1^{-1}\left\{\hat{\boldsymbol{\Sigma}}_0 + (u-1)\hat{\boldsymbol{\Sigma}}_1\right\}\boldsymbol{\Delta}_1^{-1}.\end{aligned} \tag{5.8}$$

From (5.6), (5.7), and (5.8), the derivative of the log-likelihood function is

$$\begin{aligned}\boldsymbol{U}(\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2) &= \frac{\partial}{\partial \boldsymbol{\Sigma}_1}\log L(\bar{\boldsymbol{x}}, \boldsymbol{\Sigma}_*) \\ &= \frac{n(u-1)}{2}\boldsymbol{\Delta}_1^{-1} - \frac{n(u-1)}{2}\boldsymbol{\Delta}_2^{-1} - \frac{nu}{2}\left(\boldsymbol{\Delta}_1^{-1}\hat{\boldsymbol{\Sigma}}_0\boldsymbol{\Delta}_1^{-1}\right) \\ &\quad+ \frac{n(u-1)}{2}\boldsymbol{\Delta}_2^{-1}\left\{\hat{\boldsymbol{\Sigma}}_0 + (u-1)\hat{\boldsymbol{\Sigma}}_1\right\}\boldsymbol{\Delta}_2^{-1} \\ &\quad+ \frac{n}{2}\boldsymbol{\Delta}_1^{-1}\left\{\hat{\boldsymbol{\Sigma}}_0 + (u-1)\hat{\boldsymbol{\Sigma}}_1\right\}\boldsymbol{\Delta}_1^{-1}.\end{aligned} \tag{5.9}$$

Since

$$\hat{\boldsymbol{\Sigma}}_0 = \frac{u-1}{u}\hat{\boldsymbol{\Delta}}_1 + \frac{1}{u}\hat{\boldsymbol{\Delta}}_2,$$

we have

$$\begin{aligned}\text{vec}\,(\boldsymbol{U}(\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2)) = &-\frac{n(u-1)}{2}\left(\boldsymbol{\Delta}_1^{-1} \otimes \boldsymbol{\Delta}_1^{-1}\right)\text{vec}\left(\hat{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right) \\ &+\frac{n(u-1)}{2}\left(\boldsymbol{\Delta}_2^{-1} \otimes \boldsymbol{\Delta}_2^{-1}\right)\text{vec}\left(\hat{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right). \qquad (5.10)\end{aligned}$$

Before we calculate the information matrix, we obtain the expectations and the covariance matrices of $\hat{\boldsymbol{\Delta}}_1$ and $\hat{\boldsymbol{\Delta}}_2$. Since $\tilde{\boldsymbol{\Delta}}_i = n\hat{\boldsymbol{\Delta}}_i/(n-1)$, we have

$$\begin{aligned}\hat{\boldsymbol{\Delta}}_1 &\sim W_p\left((n-1)(u-1), \frac{1}{n(u-1)}\boldsymbol{\Delta}_1\right), \\ \hat{\boldsymbol{\Delta}}_2 &\sim W_p\left(n-1, \frac{1}{n}\boldsymbol{\Delta}_2\right).\end{aligned}$$

The expectations of $\text{vec}\left(\hat{\boldsymbol{\Delta}}_1\right)$ and $\text{vec}\left(\hat{\boldsymbol{\Delta}}_2\right)$ are

$$\begin{aligned}E\left[\text{vec}\left(\hat{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right)\right] &= \frac{n-1}{n}\text{vec}\,(\boldsymbol{\Delta}_1) - \text{vec}\,(\boldsymbol{\Delta}_1) = -\frac{1}{n}\text{vec}\,(\boldsymbol{\Delta}_1)\,, \\ E\left[\text{vec}\left(\hat{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right)\right] &= \frac{n-1}{n}\text{vec}\,(\boldsymbol{\Delta}_2) - \text{vec}\,(\boldsymbol{\Delta}_2) = -\frac{1}{n}\text{vec}\,(\boldsymbol{\Delta}_2)\,,\end{aligned}$$

and the covariance matrices of $\text{vec}\left(\hat{\boldsymbol{\Delta}}_1\right)$ and $\text{vec}\left(\hat{\boldsymbol{\Delta}}_2\right)$ are

$$\begin{aligned}&E\left[\text{vec}\left(\hat{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right)\text{vec}'\left(\hat{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right)\right] \\ &\quad= E\left[\left\{\text{vec}\left(\hat{\boldsymbol{\Delta}}_1 - \frac{n-1}{n}\boldsymbol{\Delta}_1\right) - \frac{1}{n}\text{vec}\,(\boldsymbol{\Delta}_1)\right\}\right. \\ &\qquad\left.\left\{\text{vec}\left(\hat{\boldsymbol{\Delta}}_1 - \frac{n-1}{n}\boldsymbol{\Delta}_1\right) - \frac{1}{n}\text{vec}\,(\boldsymbol{\Delta}_1)\right\}'\right] \\ &\quad= \frac{n-1}{n^2(u-1)}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)(\boldsymbol{\Delta}_1 \otimes \boldsymbol{\Delta}_1) + \frac{1}{n^2}\text{vec}\,(\boldsymbol{\Delta}_1)\,\text{vec}'\,(\boldsymbol{\Delta}_1)\,, \\ &E\left[\text{vec}\left(\hat{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right)\text{vec}'\left(\hat{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right)\right] \\ &\quad= E\left[\left\{\text{vec}\left(\hat{\boldsymbol{\Delta}}_2 - \frac{n-1}{n}\boldsymbol{\Delta}_2\right) - \frac{1}{n}\text{vec}\,(\boldsymbol{\Delta}_2)\right\}\right. \\ &\qquad\left.\left\{\text{vec}\left(\hat{\boldsymbol{\Delta}}_2 - \frac{n-1}{n}\boldsymbol{\Delta}_2\right) - \frac{1}{n}\text{vec}\,(\boldsymbol{\Delta}_2)\right\}'\right] \\ &\quad= \frac{n-1}{n^2}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)(\boldsymbol{\Delta}_2 \otimes \boldsymbol{\Delta}_2) + \frac{1}{n^2}\text{vec}\,(\boldsymbol{\Delta}_2)\,\text{vec}'\,(\boldsymbol{\Delta}_2)\,.\end{aligned}$$

Therefore, the information matrix is

$$\begin{aligned}
\boldsymbol{I}\left(\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2\right) &= E\left[\operatorname{vec}\left(\boldsymbol{U}(\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2)\right)\operatorname{vec}'\left(\boldsymbol{U}(\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2)\right)\right] \\
&= \frac{n^2(u-1)^2}{4}\left(\boldsymbol{\Delta}_1^{-1} \otimes \boldsymbol{\Delta}_1^{-1}\right) E\left[\operatorname{vec}\left(\hat{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right)\operatorname{vec}'\left(\hat{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right)\right]\left(\boldsymbol{\Delta}_1^{-1} \otimes \boldsymbol{\Delta}_1^{-1}\right) \\
&\quad - \frac{n^2(u-1)^2}{4}\left(\boldsymbol{\Delta}_1^{-1} \otimes \boldsymbol{\Delta}_1^{-1}\right) E\left[\operatorname{vec}\left(\hat{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right)\operatorname{vec}'\left(\hat{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right)\right]\left(\boldsymbol{\Delta}_2^{-1} \otimes \boldsymbol{\Delta}_2^{-1}\right) \\
&\quad - \frac{n^2(u-1)^2}{4}\left(\boldsymbol{\Delta}_2^{-1} \otimes \boldsymbol{\Delta}_2^{-1}\right) E\left[\operatorname{vec}\left(\hat{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right)\operatorname{vec}'\left(\hat{\boldsymbol{\Delta}}_1 - \boldsymbol{\Delta}_1\right)\right]\left(\boldsymbol{\Delta}_1^{-1} \otimes \boldsymbol{\Delta}_1^{-1}\right) \\
&\quad + \frac{n^2(u-1)^2}{4}\left(\boldsymbol{\Delta}_2^{-1} \otimes \boldsymbol{\Delta}_2^{-1}\right) E\left[\operatorname{vec}\left(\hat{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right)\operatorname{vec}'\left(\hat{\boldsymbol{\Delta}}_2 - \boldsymbol{\Delta}_2\right)\right]\left(\boldsymbol{\Delta}_2^{-1} \otimes \boldsymbol{\Delta}_2^{-1}\right) \\
&= \frac{(n-1)(u-1)}{4}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)\left(\boldsymbol{\Delta}_1^{-1} \otimes \boldsymbol{\Delta}_1^{-1}\right) \\
&\quad + \frac{(n-1)(u-1)^2}{4}\left(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p}\right)\left(\boldsymbol{\Delta}_2^{-1} \otimes \boldsymbol{\Delta}_2^{-1}\right). \qquad (5.11)
\end{aligned}$$

References

1. Anderson, T.W.: An Introduction to Multivariate Statistical Analysis, 2nd edn. Wiley, New York (1984)
2. Arnold, S.F.: Linear models with exchangeably distributed errors. J. Am. Stat. Assoc. **74**, 194–199 (1979)
3. Coelho, C.A., Roy, A.: Testing the hypothesis of a block compound symmetric covariance matrix for elliptically contoured distributions. TEST **26**, 308–330 (2017)
4. Efron, B., Tibshirani, R.J.: An Introduction to the Bootstrap. CRC Press, New York (1994)
5. Fujikoshi, Y., Ulyanov, V.V., Shimizu, R.: Multivariate Statistics: High-Dimensional and Large-Sample Approximations. Wiley, Hoboken (2010)
6. Gupta, A.K., Nagar, D.K.: Matrix Variate Distributions. Chapman and Hall, Boca Raton (1999)
7. Johnson, R.A., Wichern, D.W.: Applied Multivariate Statistical Analysis, 6th edn. Pearson Prentice Hall, Englewood Cliffs (2007)
8. Leiva, R.: Linear discrimination with equicorrelated training vectors. J. Multivar. Anal. **98**, 384–409 (2007)
9. Pourahmadi, M.: High-Dimensional Covariance Estimation: With High-Dimensional Data. Wiley, New York (2013)
10. Roy, A., Leiva, R.: Estimating and testing a structured covariance matrix for three-level multivariate data. Commun. Stat. Theory Methods **40**, 1945–1963 (2011)
11. Roy, A., Leiva, R., Žežula, I., Klein, D.: Testing the equality of mean vectors for paired doubly multivariate observations in blocked compound symmetric covariance matrix setup. J. Multivar. Anal. **137**, 50–60 (2015)
12. Roy, A., Zmyślony, R., Fonseca, M., Leiva, R.: Optimal estimation for doubly multivariate data in blocked compound symmetric covariance structure. J. Multivar. Anal. **144**, 81–90 (2016)
13. Szatrowski, T.H.: Testing and estimation in the block compound symmetry problem. J. Educ. Stat. **7**, 3–18 (1982)
14. Terrell, G.R.: The gradient statistic. Comput. Sci. Stat. **34**, 206–215 (2002)
15. Zezula, I., Klein, D., Roy, A.: Testing of multivariate repeated measures data with block exchangeable covariance structure. TEST (2017). https://doi.org/10.1007/s11749-017-0549-z