Tests for high-dimensional covariance structures under the SSE model

Aki Ishii[a], Kazuyoshi Yata[b] and Makoto Aoshima[b]

[a] Department of Information Sciences, Tokyo University of Science
[b] Institute of Mathematics, University of Tsukuba

1 Introduction

Suppose we take samples, $\boldsymbol{x}_j = (x_{1j}, \ldots, x_{pj})^T,\ j = 1, \ldots, n$, of size $n\ (\geq 4)$, which are independent and identically distributed (i.i.d.) as a $p\ (\geq 2)$-variate distribution. We assume that $\boldsymbol{x}_j$ has an unknown mean vector $\boldsymbol{\mu}$ and unknown (positive-semidefinite) covariance matrix $\boldsymbol{\Sigma}$. We have that $\boldsymbol{\Sigma} = \boldsymbol{H\Lambda H}^T$, where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, ..., \lambda_p)$ is a diagonal matrix of eigenvalues of $\boldsymbol{\Sigma}$, $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$, and $\boldsymbol{H} = (\boldsymbol{h}_1, ..., \boldsymbol{h}_p)$ is an orthogonal matrix of the corresponding eigenvectors. Let $\boldsymbol{x}_j = \boldsymbol{H\Lambda}^{1/2}\boldsymbol{z}_j + \boldsymbol{\mu}$, where $\boldsymbol{z}_j = (z_{1j}, \ldots, z_{pj})^T$ is considered as a sphered data vector having the zero mean vector and identity covariance matrix. Let $\sigma = \mathrm{tr}(\boldsymbol{\Sigma})/p$. Let σ_{ij} be the (i, j) element of $\boldsymbol{\Sigma}$ for $i, j = 1, \ldots, p$. We assume that $\sigma_{jj} \in (0, \infty)$ as $p \to \infty$ for all j. For a function, $f(\cdot)$, "$f(p) \in (0, \infty)$ as $p \to \infty$" implies that $\liminf_{p\to\infty} f(p) > 0$ and $\limsup_{p\to\infty} f(p) < \infty$. Then, it holds that $\sigma \in (0, \infty)$ as $p \to \infty$. Let $\rho = \sum_{i\neq j}^{p} \sigma_{ij}/\{\sigma p(p-1)\}$. Note that

$$\frac{\boldsymbol{1}_p^T\boldsymbol{\Sigma}\boldsymbol{1}_p}{p} = \sigma\{1 + \rho(p-1)\} \tag{1}$$

and $\rho \in [-(p-1)^{-1}, 1]$, where $\boldsymbol{1}_p = (1, \ldots, 1)^T$. We denote the identity matrix of dimension p by $\boldsymbol{I}_p$.

In this paper, we consider testing

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_* \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_*, \tag{2}$$

where $\boldsymbol{\Sigma}_*$ is a candidate (positive-semidefinite) covariance matrix. For $\boldsymbol{\Sigma}_*$ we consider the following covariance structures: (i) identity matrix, (ii) scaled identity matrix, (iii) diagonal matrix, and (iv) intraclass covariance matrix. Let

$$\boldsymbol{\Sigma}_{\mathrm{S}} = \sigma\boldsymbol{I}_p,\ \boldsymbol{\Sigma}_{\mathrm{D}} = \mathrm{diag}(\sigma_{11}, \ldots, \sigma_{pp}) \ \text{ and } \ \boldsymbol{\Sigma}_{\mathrm{IC}} = \sigma\{(1-\rho)\boldsymbol{I}_p + \rho\boldsymbol{1}_p\boldsymbol{1}_p^T\}.$$

Ledoit and Wolf (2002) gave test procedures for

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{I}_p \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{I}_p \tag{3}$$

and

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{S}} \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_{\mathrm{S}} \tag{4}$$

when $p/n \to c > 0$ and $\boldsymbol{x}_j$ is Gaussian. Schott (2005) gave a test procedure for

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{D}} \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_{\mathrm{D}} \tag{5}$$

when $p/n \to c > 0$ and $\boldsymbol{x}_j$ is Gaussian. Srivastava, Kollo and Rosen (2011) considered test procedures for (3) to (5) when $n/p \to 0$ under an assumption that is stronger than (A-ii) given in Section 2. On the other hand, Srivastava and Reid (2012) gave a test procedure for

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{IC}} \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_{\mathrm{IC}} \tag{6}$$

when $n/p \to 0$ and $\boldsymbol{x}_j$ is Gaussian. Meanwhile, Zhong et al. (2017) considered a high-dimensional regression model and testing (6) for the covariance matrix associated with error vectors when the error vectors are Gaussian. However, it is known that those test statistics do not give a preferable performance unless $\boldsymbol{x}_j$ is Gaussian. As for a nonparametric approach, Chen, Zhang and Zhong (2010) considered test statistics based on the U-statistic for (3) and (4). In the current paper, we take a different nonparametric approach and produce a new test statistic for (2). We utilize the extended cross-data-matrix (ECDM) method developed by Yata and Aoshima (2013) which is an extension of the cross-data-matrix methodology created by Yata and Aoshima (2010). The ECDM method is a nonparametric method to produce an unbiased estimator for a function of $\boldsymbol{\Sigma}$ at a low computational cost even for ultra high-dimensional data. In addition, the ECDM method possesses a high versatility in high-dimensional data analysis. See Yata and Aoshima (2016) for the details.

In this paper, we consider constructing a new test procedure for (6) by using the ECDM method. In Section 2, we produce an ECDM test statistic when $\boldsymbol{\Sigma}_*$ is known. We show that the ECDM test statistic is an unbiased estimator of its test parameter even in a high-dimensional setting. In Section 3, we produce an ECDM test statistic when $\boldsymbol{\Sigma}_*$ involves unknown parameters. We propose a new test procedure based on the test statistic and evaluate its asymptotic size and power theoretically. In Section 4, we apply the new test procedure to testing (6).

2 Test procedure for (2) when Σ_* is known

In this section, we propose a test procedure for (2) when $\boldsymbol{\Sigma}_*$ is known and evaluate its asymptotic size and power theoretically. Let

$$\boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma} - \boldsymbol{\Sigma}_* \quad \text{and} \quad \Delta = \|\boldsymbol{\Sigma}_0\|_F^2 = \mathrm{tr}(\boldsymbol{\Sigma}_0^2),$$

where $\|\cdot\|_F$ is the Frobenius norm. Note that $\Delta = 0$ under H_0 and $\Delta > 0$ under H_1. We regard Δ as a test parameter and construct a test procedure for (2) by using an estimator of Δ.

2.1 Unbiased estimator of Δ

We first give an unbiased estimator of Δ by using the ECDM method. Let $n_{(1)} = \lceil n/2 \rceil$ and $n_{(2)} = n - n_{(1)}$, where $\lceil x \rceil$ denotes the smallest integer $\geq x$. Let

$$\boldsymbol{V}_{n(1)(k)} = \begin{cases} \{\lfloor k/2 \rfloor - n_{(1)} + 1, \ldots, \lfloor k/2 \rfloor\} & \text{if } \lfloor k/2 \rfloor \geq n_{(1)}, \\ \{1, \ldots, \lfloor k/2 \rfloor\} \cup \{\lfloor k/2 \rfloor + n_{(2)} + 1, \ldots, n\} & \text{otherwise;} \end{cases}$$

$$\boldsymbol{V}_{n(2)(k)} = \begin{cases} \{\lfloor k/2 \rfloor + 1, \ldots, \lfloor k/2 \rfloor + n_{(2)}\} & \text{if } \lfloor k/2 \rfloor \leq n_{(1)}, \\ \{1, \ldots, \lfloor k/2 \rfloor - n_{(1)}\} \cup \{\lfloor k/2 \rfloor + 1, \ldots, n\} & \text{otherwise} \end{cases}$$

for $k = 3, \ldots, 2n-1$, where $\lfloor x \rfloor$ denotes the largest integer $\leq x$. Let $\#\boldsymbol{S}$ denote the number of elements in a set $\boldsymbol{S}$. Note that $\#\boldsymbol{V}_{n(l)(k)} = n_{(l)}$, $l = 1, 2$, $\boldsymbol{V}_{n(1)(k)} \cap \boldsymbol{V}_{n(2)(k)} = \emptyset$ and $\boldsymbol{V}_{n(1)(k)} \cup \boldsymbol{V}_{n(2)(k)} = \{1, \ldots, n\}$ for $k = 3, \ldots, 2n-1$. Also, note that $i \in \boldsymbol{V}_{n(1)(i+j)}$ and $j \in \boldsymbol{V}_{n(2)(i+j)}$ for $i < j \ (\leq n)$. Let

$$\overline{\boldsymbol{x}}_{(1)(k)} = n_{(1)}^{-1} \sum_{j \in \boldsymbol{V}_{n(1)(k)}} \boldsymbol{x}_j \quad \text{and} \quad \overline{\boldsymbol{x}}_{(2)(k)} = n_{(2)}^{-1} \sum_{j \in \boldsymbol{V}_{n(2)(k)}} \boldsymbol{x}_j$$

for $k = 3, \dots, 2n-1$. Let $u_{n(l)} = n_{(l)}/(n_{(l)}-1)$ for $l = 1, 2$,

$$\boldsymbol{y}_{ij(1)} = \boldsymbol{x}_i - \overline{\boldsymbol{x}}_{(1)(i+j)} \quad \text{and} \quad \boldsymbol{y}_{ij(2)} = \boldsymbol{x}_j - \overline{\boldsymbol{x}}_{(2)(i+j)}$$

for all $i < j$. We note that $u_{n(l)} E(\boldsymbol{y}_{ij(l)} \boldsymbol{y}_{ij(l)}^T) = \boldsymbol{\Sigma}$ for $l = 1, 2$, and $\boldsymbol{y}_{ij(1)}$ and $\boldsymbol{y}_{ij(2)}$ are independent for all $i < j$. For example, Yata and Aoshima (2013) gave an estimator of tr$(\boldsymbol{\Sigma}^2)$ as

$$W_n = \frac{2u_{n(1)}u_{n(2)}}{n(n-1)} \sum_{i<j}^{n} (\boldsymbol{y}_{ij(1)}^T \boldsymbol{y}_{ij(2)})^2 \tag{7}$$

by the ECDM method. Then, it holds that $E(W_n) = \text{tr}(\boldsymbol{\Sigma}^2)$.

Now, we can give an unbiased estimator of Δ as

$$\widehat{\Delta}_n = 2 \sum_{i<j}^{n} \frac{\text{tr}\big\{(u_{n(1)} \boldsymbol{y}_{ij(1)} \boldsymbol{y}_{ij(1)}^T - \boldsymbol{\Sigma}_*)(u_{n(2)} \boldsymbol{y}_{ij(2)} \boldsymbol{y}_{ij(2)}^T - \boldsymbol{\Sigma}_*)\big\}}{n(n-1)} \tag{8}$$

by the ECDM method. Note that $E(\widehat{\Delta}_n) = \Delta$. Here, we write that

$$\widehat{\Delta}_n = W_n + \text{tr}(\boldsymbol{\Sigma}_*^2) - 2 \sum_{i<j}^{n} \frac{\big(u_{n(1)} \boldsymbol{y}_{ij(1)}^T \boldsymbol{\Sigma}_* \boldsymbol{y}_{ij(1)} + u_{n(2)} \boldsymbol{y}_{ij(2)}^T \boldsymbol{\Sigma}_* \boldsymbol{y}_{ij(2)}\big)}{n(n-1)}. \tag{9}$$

The computational cost of $\widehat{\Delta}_n$ by (9) is much lower than that by (8) when $n = o(p)$.

2.2 Asymptotic properties of $\widehat{\Delta}_n$

We assume the following model:

$$\boldsymbol{x}_j = \boldsymbol{\Gamma} \boldsymbol{w}_j + \boldsymbol{\mu},$$

where $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_d)$ is a $p \times d$ matrix for some $d > 0$ such that $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T = \boldsymbol{\Sigma}$, and $\boldsymbol{w}_j = (w_{1j}, \dots, w_{dr})^T$, $j = 1, \dots, n$, are i.i.d. random vectors having $E(\boldsymbol{w}_j) = \boldsymbol{0}$ and Var$(\boldsymbol{w}_j) = \boldsymbol{I}_d$. Let Var$(w_{sj}^2) = M_s$ for $s = 1, \dots, d$. We assume that $\limsup_{p\to\infty} M_s < \infty$ for all s. Similar to Bai and Saranadasa (1996), Chen and Qin (2010) and Aoshima and Yata (2015), we assume that

(A-i) $E(w_{sj}^2 w_{tj}^2) = E(w_{sj}^2)E(w_{tj}^2) = 1$ and $E(w_{sj} w_{tj} w_{uj} w_{vj}) = 0$ for all $s \neq t, u, v$.

We assume the following assumption instead of (A-i) as necessary:

(A-ii) $E(w_{s_1 j}^{\alpha_1} w_{s_2 j}^{\alpha_2} \cdots w_{s_v j}^{\alpha_v}) = E(w_{s_1 j}^{\alpha_1}) E(w_{s_2 j}^{\alpha_2}) \cdots E(w_{s_v j}^{\alpha_v})$ for all $s_1 \neq s_2 \neq \cdots \neq s_v \in [1, d]$ and $\alpha_i \in [1, 4]$, $i = 1, \dots, v$, where $v \leq 8$ and $\sum_{i=1}^{v} \alpha_i \leq 8$.

Note that (A-ii) implies (A-i). When $\boldsymbol{x}_j$ is Gaussian and $\boldsymbol{\Gamma} = \boldsymbol{H}\boldsymbol{\Lambda}^{1/2}$, it holds that $\boldsymbol{w}_j = \boldsymbol{z}_j$ and $\boldsymbol{z}_j$ is distributed as $N_p(\boldsymbol{0}, \boldsymbol{I}_p)$, so that (A-ii) is naturally satisfied.

For $\boldsymbol{\Sigma}$ we assume the following condition as necessary:

(C-i) $\dfrac{\text{tr}(\boldsymbol{\Sigma}^4)}{\text{tr}(\boldsymbol{\Sigma}^2)^2} \to 0$ as $p \to \infty$.

Note that (C-i) is equivalent to "$\lambda_1/\text{tr}(\boldsymbol{\Sigma}^2)^{1/2} \to 0$ as $p \to \infty$". Aoshima and Yata (2018) called (C-i) the "non-strongly spiked eigenvalue (NSSE) model". When $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{S}}$ or $\boldsymbol{\Sigma}_{\text{D}}$, (C-i) holds. On the other hand, when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{IC}}$ with $\liminf_{p\to\infty} \rho > 0$, (C-i) does not hold since it follows that

$$\liminf_{p\to\infty} \left\{ \frac{\lambda_1}{\text{tr}(\boldsymbol{\Sigma}^2)^{1/2}} \right\} > 0 \tag{10}$$

from the facts that $\lambda_1 = \sigma\{1 + (p-1)\rho\}$ and $\mathrm{tr}(\boldsymbol{\Sigma}^2) = O(p^2)$. Aoshima and Yata (2018) called (10) the "strongly spiked eigenvalue (SSE) model".

Let

$$m = \min\{p, n\}.$$

We consider the divergence condition as

$$p \to \infty \ \text{ and } \ n \to \infty,$$

which is equivalent to $m \to \infty$. Let

$$K = 4\mathrm{tr}(\boldsymbol{\Sigma}^2)^2/n^2 \ \text{ and } \ K_* = 4\mathrm{tr}(\boldsymbol{\Sigma}_*^2)^2/n^2.$$

We assume one of the following assumptions as necessary:

(C-ii) $\dfrac{K^{1/2}}{\Delta} \to 0$ as $m \to \infty$; (C-iii) $\displaystyle\limsup_{m\to\infty} \frac{\Delta}{K^{1/2}} < \infty$.

Note that (C-iii) holds under H_0 in (2).

For $\widehat{\Delta}_n$ in (8), we have the following results.

Lemma 2.1. *Assume (A-i). Then, it holds that as $m \to \infty$*

$$\mathit{Var}(\widehat{\Delta}_n) = K\{1 + o(1)\} + O\Big(\frac{\mathit{tr}(\boldsymbol{\Sigma}^4)^{1/2}\Delta}{n} + \frac{\mathit{tr}(\boldsymbol{\Sigma}^4)}{n^2}\Big).$$

Furthermore, under (C-i) and (C-iii), it holds that as $m \to \infty$

$$\mathit{Var}(\widehat{\Delta}_n) = K\{1 + o(1)\}.$$

From Lemma 2.1 we obtain the following result under (C-ii).

Theorem 2.1. *Assume (A-i) and (C-ii). Then, it holds that as $m \to \infty$*

$$\widehat{\Delta}_n/\Delta = 1 + o_P(1).$$

On the other hand, we obtain the following result under (C-iii).

Theorem 2.2. *Assume (A-ii), (C-i) and (C-iii). Then, it holds that as $m \to \infty$*

$$\frac{\widehat{\Delta}_n - \Delta}{K^{1/2}} \Rightarrow N(0, 1),$$

where "$\Rightarrow$" denotes the convergence in distribution and $N(0,1)$ denotes a random variable distributed as the standard normal distribution.

2.3 Test procedure based on $\widehat{\Delta}_n$

Note that $\mathrm{tr}(\boldsymbol{\Sigma}^2) = \mathrm{tr}(\boldsymbol{\Sigma}_*^2)$ under H_0 in (2). Let

$$T_n = \frac{n\widehat{\Delta}_n}{2\mathrm{tr}(\boldsymbol{\Sigma}_*^2)}. \tag{11}$$

From Theorem 2.2 we propose a test procedure for (2) by

$$\text{rejecting } H_0 \Longleftrightarrow T_n > z_\alpha, \tag{12}$$

where z_α is a constant such that $P\{N(0,1) > z_\alpha\} = \alpha$ with $\alpha \in (0, 1/2)$. Then, we have the following result.

Theorem 2.3. *Assume (A-ii) and (C-i). For the test procedure (12), we have that*

$$\textit{Size} = \alpha + o(1) \quad \textit{and} \quad \textit{Power} = \Phi\bigg(\frac{\Delta}{K^{1/2}} - z_\alpha \frac{K_*^{1/2}}{K^{1/2}}\bigg) + o(1) \ \ \textit{as } m \to \infty, \tag{13}$$

where $\Phi(\cdot)$ denotes the c.d.f. of $N(0,1)$.

Corollary 2.1. *Assume (A-i). Assume also (C-ii) under H_1 in (2). For the test procedure (12), we have that*

$$\textit{Power} \to 1 \ \ \textit{as } m \to \infty. \tag{14}$$

3 Test procedure for (2) when $\boldsymbol{\Sigma}_*$ involves unknown parameters

When $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\text{IC}}$, the eigenstructures are identified. Otherwise they involve unknown parameters. In this section, we construct an unbiased estimator of Δ through the eigenstructures and propose a test procedure by using the unbiased estimator.

3.1 Test procedure

Let $\boldsymbol{A}_j$ be a $p \times p$ known idempotent matrix with rank r_j (≥ 1) for $j = 1, \dots, q$, such that $\sum_{j=1}^q r_j = p$ and $\sum_{j=1}^q \boldsymbol{A}_j = \boldsymbol{I}_p$, where $r_1 \leq \cdots \leq r_q$ when $q \geq 2$. Note that $\text{tr}(\boldsymbol{A}_j) = r_j$, $\boldsymbol{A}_j^2 = \boldsymbol{A}_j$ and $\boldsymbol{A}_j \boldsymbol{A}_{j'} = \boldsymbol{O}$ for all j $(\neq j')$. Let κ_j (≥ 0) be an unknown scalar such that $\text{tr}(\boldsymbol{\Sigma}\boldsymbol{A}_j) = r_j\kappa_j$ for all j. Hereafter, we assume that $\boldsymbol{\Sigma}_*$ has the following structure:

$$\boldsymbol{\Sigma}_* = \kappa_1 \boldsymbol{A}_1 + \cdots + \kappa_q \boldsymbol{A}_q. \tag{15}$$

Note that $\text{tr}(\boldsymbol{\Sigma}_*^2) = \sum_{j=1}^q r_j\kappa_j^2$ and $\Delta = \text{tr}(\boldsymbol{\Sigma}^2) - \text{tr}(\boldsymbol{\Sigma}_*^2)$, so that $\text{tr}(\boldsymbol{\Sigma}^2) \geq \text{tr}(\boldsymbol{\Sigma}_*^2)$. As for $\boldsymbol{\Sigma}_* = \boldsymbol{\Sigma}_{\text{IC}}$,

$\boldsymbol{A}_1 = \boldsymbol{1}_p\boldsymbol{1}_p^T/p$, $\boldsymbol{A}_2 = \boldsymbol{I}_p - \boldsymbol{1}_p\boldsymbol{1}_p^T/p$, $\kappa_1 = \sigma\{1+(p-1)\rho\}$, $\kappa_2 = \sigma(1-\rho)$, $r_1 = 1$, $r_2 = p-1$ and $q = 2$.

We note that $\text{tr}(\boldsymbol{\Sigma}\boldsymbol{A}_j)^2/r_j = r_j\kappa_j^2$ for all j. Then, we give an estimator of $\text{tr}(\boldsymbol{\Sigma}_*^2)$ as

$$U_n = 2\sum_{s=1}^q \sum_{i<j}^n \frac{u_{n(1)}u_{n(2)}\boldsymbol{y}_{ij(1)}^T\boldsymbol{A}_s\boldsymbol{y}_{ij(1)}\boldsymbol{y}_{ij(2)}^T\boldsymbol{A}_s\boldsymbol{y}_{ij(2)}}{r_s n(n-1)} \tag{16}$$

by the ECDM method. Note that $E(U_n) = \text{tr}(\boldsymbol{\Sigma}_*^2)$. Let

$$\widetilde{\Delta}_n = W_n - U_n,$$

where W_n is given by (7). Then, it holds that $E(\widetilde{\Delta}_n) = \Delta$. Here, we write that

$$\begin{aligned} U_n &= B_n(1) - \text{tr}(\boldsymbol{\Sigma}_*^2) \\ &\quad + 2\sum_{i<j}^n \frac{u_{n(1)}\boldsymbol{y}_{ij(1)}^T\boldsymbol{\Sigma}_*\boldsymbol{y}_{ij(1)} + u_{n(2)}\boldsymbol{y}_{ij(2)}^T\boldsymbol{\Sigma}_*\boldsymbol{y}_{ij(2)}}{n(n-1)}, \end{aligned} \tag{17}$$

where

$$B_n(t) = 2\sum_{i<j}^n \sum_{s=t}^q \frac{(u_{n(1)}\boldsymbol{y}_{ij(1)}^T\boldsymbol{A}_s\boldsymbol{y}_{ij(1)} - \kappa_s r_s)(u_{n(2)}\boldsymbol{y}_{ij(2)}^T\boldsymbol{A}_s\boldsymbol{y}_{ij(2)} - \kappa_s r_s)}{r_s n(n-1)}$$

for $t = 1, \dots, q$. By combining (9) and (17), we have that

$$\widetilde{\Delta}_n = \widehat{\Delta}_n - B_n(1). \tag{18}$$

Note that $E\{B_n(t)\} = 0$ for all t. Let us consider an asymptotic variance of $\widetilde{\Delta}_n$. Let $q_\star$ be the maximum integer such that

$$r_1 = \cdots = r_{q_\star} = 1 < r_{q_\star+1} \le \cdots \le r_q.$$

If $r_1 = \cdots = r_q = 1$, we set $q_\star = q$. We set $B_n(q_\star + 1) = 0$ when $q_\star = q$. If $r_j \ge 2$ for all j, we set $q_\star = 0$. Let

$$\boldsymbol{Y}_{ij(l),s} = u_{n(l)} \boldsymbol{y}_{ij(l)} \boldsymbol{y}_{ij(l)}^T \boldsymbol{A}_s - \kappa_s \boldsymbol{A}_s$$

for all i, j, l, s. Then, from (8), it follows that

$$\widetilde{\Delta}_n = \begin{cases} 2\sum_{i<j}^{n}\Big(\sum_{s\ne s'}^{q} \frac{\mathrm{tr}(\boldsymbol{Y}_{ij(1),s}\boldsymbol{Y}_{ij(2),s'})}{n(n-1)} + \sum_{s=q_\star+1}^{q} \frac{\mathrm{tr}(\boldsymbol{Y}_{ij(1),s}\boldsymbol{Y}_{ij(2),s})}{n(n-1)}\Big) & \\ -B_n(q_\star+1) & (q_\star < q), \\ 2\sum_{i<j}^{n}\sum_{s\ne s'}^{q} \frac{\mathrm{tr}(\boldsymbol{Y}_{ij(1),s}\boldsymbol{Y}_{ij(2),s'})}{n(n-1)} & (q_\star = q). \end{cases}$$

We have the following result.

Proposition 3.1. *Assume (A-i). Under H_0 in (2), it holds that as $m \to \infty$*

$$\mathit{Var}\{\widetilde{\Delta}_n + B_n(q_\star+1)\} = \frac{4\Psi}{n^2}\{1 + o(1)\} + O\Big(\frac{\mathit{tr}(\boldsymbol{\Sigma}_*^4)}{n^2}\Big),$$

where $\Psi = \mathit{tr}(\boldsymbol{\Sigma}_*^2)^2 - \sum_{s=1}^{q_\star} \kappa_s^4$ when $q_\star \ge 1$ and $\Psi = \mathit{tr}(\boldsymbol{\Sigma}_*^2)^2$ when $q_\star = 0$.

Since $B_n(q_\star+1)$ is a redundant term, we can regard "$4\Psi/n^2$" as an asymptotic variance of $\widetilde{\Delta}_n$ under H_0 in (2). We note that

$$E\Big(2\sum_{i<j}^{n} \frac{u_{n(1)}u_{n(2)}\boldsymbol{y}_{ij(1)}^T\boldsymbol{A}_s\boldsymbol{y}_{ij(1)}\boldsymbol{y}_{ij(2)}^T\boldsymbol{A}_s\boldsymbol{y}_{ij(2)}}{r_s n(n-1)}\Big) = r_s\kappa_s^2$$

in view of (16). We give an estimator of Ψ by

$$\widetilde{\Psi}_n = \begin{cases} U_n^2 - \sum_{s=1}^{q_\star}\Big(2\sum_{i<j}^{n} \frac{u_{n(1)}u_{n(2)}\boldsymbol{y}_{ij(1)}^T\boldsymbol{A}_s\boldsymbol{y}_{ij(1)}\boldsymbol{y}_{ij(2)}^T\boldsymbol{A}_s\boldsymbol{y}_{ij(2)}}{n(n-1)}\Big)^2 & (q_\star \ge 1), \\ U_n^2 & (q_\star = 0) \end{cases}$$

in view of $r_1 = \cdots = r_{q_\star} = 1$ when $q_\star \ge 1$. Note that $P(\widetilde{\Psi}_n \ge 0) = 1$. Let

$$\widetilde{T}_n = \frac{n\widetilde{\Delta}_n}{2\widetilde{\Psi}_n^{1/2}}. \tag{19}$$

Then, for (2) with (15), we propose a test procedure by

$$\text{rejecting } H_0 \Longleftrightarrow \widetilde{T}_n > z_\alpha. \tag{20}$$

3.2 Test procedure (20) under the SSE model

For the SSE model (10), we focus on the following model:

$$\frac{\lambda_1}{\text{tr}(\boldsymbol{\Sigma}^2)^{1/2}} \to 1 \ \text{ as } p \to \infty. \tag{21}$$

Note that (21) is one of the SSE models. When $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{IC}}$ and $\liminf_{p\to\infty} \rho > 0$, (21) is met. We call (21) the "uni-SSE (USSE) model". See Ishii, Yata and Aoshima (2016, 2019) for several statistical inferences under the USSE model.

We consider the test procedure (20) under the USSE model (21). One may suppose $r_1 = 1$. We assume the following condition:

(C-v) $\dfrac{\kappa_1}{\text{tr}(\boldsymbol{\Sigma}^2)^{1/2}} \to 1 \ \text{ as } p \to \infty.$

Note that (21) holds under (C-v) from the fact that $\lambda_1 \geq \kappa_1 = \text{tr}(\boldsymbol{\Sigma}\boldsymbol{A}_1)$. Let $\boldsymbol{A}_{(1)} = \boldsymbol{I}_p - \boldsymbol{A}_1$, $\boldsymbol{\Omega}_1 = \boldsymbol{A}_1\boldsymbol{\Sigma}\boldsymbol{A}_{(1)}$ and $\boldsymbol{\Omega}_2 = \boldsymbol{A}_{(1)}\boldsymbol{\Sigma}\boldsymbol{A}_{(1)}$. Note that (C-v) is equivalent to

$$\text{tr}(\boldsymbol{\Omega}_2^2)^{1/2}/\kappa_1 \to 0 \ \text{ as } p \to \infty$$

from the facts that $\text{tr}(\boldsymbol{\Sigma}^2) = \kappa_1^2 + \text{tr}(\boldsymbol{\Omega}_2^2) + 2\|\boldsymbol{\Omega}_1\|_F^2$ and $\|\boldsymbol{\Omega}_1\|_F^2 \leq \kappa_1 \text{tr}(\boldsymbol{\Omega}_2^2)^{1/2}$. As for $\boldsymbol{\Omega}_2$, we assume the following model:

(C-i') $\dfrac{\text{tr}(\boldsymbol{\Omega}_2^4)}{\text{tr}(\boldsymbol{\Omega}_2^2)^2} \to 0 \ \text{ as } p \to \infty.$

Note that (C-i') holds when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{IC}}$ and $\limsup_{p\to\infty} \rho < 1$ because $\boldsymbol{\Omega}_2 = \sigma(1-\rho)(\boldsymbol{I}_p - \boldsymbol{1}_p\boldsymbol{1}_p^T/p)$ and $\text{tr}(\boldsymbol{\Omega}_2^4)/\text{tr}(\boldsymbol{\Omega}_2^2)^2 = 1/\text{tr}(\boldsymbol{I}_p - \boldsymbol{1}_p\boldsymbol{1}_p^T/p) = 1/(p-1)$ when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{IC}}$. Here, we write that

$$\widetilde{\Delta}_n = 2\sum_{i<j}^{n}\Bigg(\sum_{s\neq s'}^{q} \frac{\text{tr}(\boldsymbol{Y}_{ij(1),s}\boldsymbol{Y}_{ij(2),s'})}{n(n-1)} + \sum_{s=2}^{q} \frac{\text{tr}(\boldsymbol{Y}_{ij(1),s}\boldsymbol{Y}_{ij(2),s})}{n(n-1)}\Bigg) - B_n(2).$$

Let

$$\Upsilon = 2\kappa_1^2\text{tr}(\boldsymbol{\Omega}_2^2) + \text{tr}(\boldsymbol{\Omega}_2^2)^2 + 2\|\boldsymbol{\Omega}_1\|_F^4 + 4\|\boldsymbol{\Omega}_1\|_F^2\text{tr}(\boldsymbol{\Omega}_2^2) \ \text{ and } \ L = 4\Upsilon/n^2.$$

Note that

$$\Upsilon = 2\kappa_1^2\sum_{s=2}^{q} r_s\kappa_s^2 + \Bigg(\sum_{s=2}^{q} r_s\kappa_s^2\Bigg)^2 = \text{tr}(\boldsymbol{\Sigma}_*^2)^2 - \kappa_1^4 \ \ (= \Psi_1, \text{ say})$$

when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_*$. Then, we have the following results.

Lemma 3.1. *Assume (A-i). It holds that as $m \to \infty$*

$$\mathit{Var}\{\widetilde{\Delta}_n + B_n(2)\} = L\{1+o(1)\} + O\Big(\frac{\Delta \mathit{tr}(\boldsymbol{\Omega}^4)^{1/4}\{\mathit{tr}(\boldsymbol{\Omega}^4)^{1/4} + \kappa_1\}}{n}\Big) + O\Big(\frac{\mathit{tr}(\boldsymbol{\Omega}_2^4)^{1/2}\{\kappa_1^2 + \mathit{tr}(\boldsymbol{\Omega}_2^4)^{1/2}\}}{n^2}\Big).$$

Furthermore, under (C-i') and

(C-iii') $\displaystyle\limsup_{m\to\infty} \frac{\Delta}{L^{1/2}} < \infty,$

it holds that as $m \to \infty$

$$\mathit{Var}\{\widetilde{\Delta}_n + B_n(2)\} = L\{1+o(1)\}.$$

Lemma 3.2. *Assume (A-ii), (C-i') and (C-v). Assume also (C-iii'). It holds that as* $m \to \infty$

$$\frac{\widetilde{\Delta}_n + B_n(2) - \Delta}{L^{1/2}} \Rightarrow N(0,1).$$

We note that $\sum_{j,j'=2}^{q}(\sum_{s=1}^{d} \boldsymbol{\gamma}_s^T \boldsymbol{A}_j \boldsymbol{\gamma}_s \boldsymbol{\gamma}_s^T \boldsymbol{A}_{j'} \boldsymbol{\gamma}_s)^2 \le \{\sum_{s=1}^{d}(\boldsymbol{\gamma}_s^T \boldsymbol{A}_{(1)} \boldsymbol{\gamma}_s)^2\}^2 \le \text{tr}\{(\boldsymbol{\Sigma}\boldsymbol{A}_{(1)})^2\}^2 = \text{tr}(\boldsymbol{\Omega}_2^2)^2$ and $\sum_{j,j'=2}^{q} \text{tr}(\boldsymbol{\Sigma}\boldsymbol{A}_j\boldsymbol{\Sigma}\boldsymbol{A}_{j'})^2 \le \text{tr}(\boldsymbol{\Omega}_2^2)^2$. Then, under (A-i) and (C-v), it holds that as $m \to \infty$

$$\text{Var}\{B_n(2)\} = o(L). \tag{22}$$

Thus, from Lemma 3.2, we have the following result.

Theorem 3.1. *Assume (A-ii), (C-i'), (C-iii') and (C-v). It holds that as* $m \to \infty$

$$\frac{\widetilde{\Delta}_n - \Delta}{L^{1/2}} \Rightarrow N(0,1).$$

Lemma 3.3. *Assume (A-i) and (C-v). It holds that* $\widetilde{\Psi}_n/\Psi_1 = 1 + o_P(1)$ *as* $m \to \infty$.

From Theorem 3.1 and Lemma 3.3, we have the following results.

Theorem 3.2. *Assume (A-ii), (C-i') and (C-v). For the test procedure (20), we have that*

$$\textit{Size} = \alpha + o(1) \quad \textit{and} \quad \textit{Power} = \Phi\left(\frac{\Delta}{L^{1/2}} - z_\alpha \frac{L_*^{1/2}}{L^{1/2}}\right) + o(1) \ \textit{ as } m \to \infty, \tag{23}$$

where $L_* = 4\Psi_1/n^2$.

Corollary 3.1. *Assume (A-i) and (C-v). Assume also*

(C-ii') $\dfrac{L^{1/2}}{\Delta} \to 0$ *as* $m \to \infty$

under H_1 *in (2). For the test procedure (20), we have (14).*

Acknowledgements

Research of the first author was partially supported by Grant-in-Aid for Young Scientists, Japan Society for the Promotion of Science (JSPS), under Contract Number 18K18015. Research of the second author was partially supported by Grant-in-Aid for Scientific Research (C), JSPS, under Contract Number 18K03409. Research of the third author was partially supported by Grants-in-Aid for Scientific Research (A) and Challenging Research (Exploratory), JSPS, under Contract Numbers 15H01678 and 19K22837.

References

Aoshima, M. and Yata, K. (2015). Asymptotic normality for inference on multisample, high-dimensional mean vectors under mild conditions. *Methodology and Computing in Applied Probability, 17*, 419–439.

Aoshima, M. and Yata, K. (2018). Two-sample tests for high-dimension, strongly spiked eigenvalue models. *Statistica Sinica, 28*, 43–62.

Bai, Z. and Saranadasa, H. (1996). Effect of high dimension: by an example of a two sample problem. *Statistica Sinica, 6*, 311–329.

Chen, S.X. and Qin, Y.-L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *The Annals of Statistics, 38*, 808–835.

Chen, S.X., Zhang, L.-X. and Zhong, P.-S. (2010). Tests for high-dimensional covariance matrices. *Journal of the American Statistical Association, 105*, 810–819.

Ishii, A., Yata, K. and Aoshima, M. (2016). Asymptotic properties of the first principal component and equality tests of covariance matrices in high-dimension, low-sample-size context. *Journal of Statistical Planning and Inference, 170*, 186–199.

Ishii, A., Yata, K. and Aoshima, M. (2019). Equality tests of high-dimensional covariance matrices under the strongly spiked eigenvalue model. *Journal of Statistical Planning and Inference, 202*, 99-111.

Ledoit, O. and Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Annals of Statistics, 30*, 1081–1102.

Schott, J. R. (2005). Testing for complete independence in high dimensions. *Biometrika, 92*, 951–956.

Srivastava, M. S., Kollo, T. and Rosen, V.D. (2011). Some tests for the covariance matrix with fewer observations than the dimension under non-normality. *Journal of Multivariate Analysis, 102*, 1090-1103.

Srivastava, M. S. and Reid, N. (2012). Testing the structure of the covariance matrix with fewer observations than the dimension. *Journal of Multivariate Analysis, 112*, 156-171.

Yata, K. and Aoshima, M. (2010). Effective PCA for high-dimension, low-sample-size data with singular value decomposition of cross data matrix. *Journal of Multivariate Analysis, 101*, 2060–2077.

Yata, K. and Aoshima, M. (2013). Correlation tests for high-dimensional data using extended cross-data-matrix methodology. *Journal of Multivariate Analysis, 117*, 313–331.

Yata, K. and Aoshima, M. (2016). High-dimensional inference on covariance structures via the extended cross-data-matrix methodology. *Journal of Multivariate Analysis, 151*, 151–166.

Zhong, P.-S., Lan, W., Song, X.K.P. and Tsai, C.H. (2017). Tests for covariance structures with high-dimensional repeated mesurements. *The Annals of Statistics, 45*, 1185–1213.