A Study of Two High-dimensional Likelihood Ratio Tests under Alternative Hypotheses

Huijun Chen[1] and Tiefeng Jiang[2]
Jilin University and University of Minnesota

Abstract

Let $N_p(\mu, \mathbf{\Sigma})$ be a p-dimensional normal distribution. Testing $\mathbf{\Sigma}$ equal to a given matrix or $(\mu, \mathbf{\Sigma})$ equal to a give pair through the likelihood ratio test (LRT) are classical problems in the multivariate analysis. When the population dimension p is fixed, it is known the LRT statistics go to χ^2-distributions. When p is large, simulation shows that the approximations are far from accurate. For the two LRT statistics, in the high dimensional cases, we obtain their central limit theorems under a big class of alternative hypotheses. In particular, the alternative hypotheses are not local ones. We do not need the assumption that p and n are proportional to each other. The condition $n - 1 > p \to \infty$ suffices in our results.

Keywords: Likelihood ratio test, alternative hypothesis, central limit theorem, high-dimensional data, multivariate normal distribution, multivariate Gamma function.
AMS 2000 Subject Classification: Primary 62H15; secondary 62H10.

[1]School of Mathematics, Jilin University, Changchun 130021, China, hchen@jlu.edu.cn.
[2]School of Statistics, University of Minnesota, 224 Church Street, S. E., MN55455, USA, jiang040@umn.edu.
The research of Tiefeng Jiang was supported in part by NSF Grant DMS-1209166 and DMS-1406279.

1 Introduction

Traditional likelihood ratio tests (LRT) have been used very popularly in statistics, see, for example, Muirhead (1982), Eaton (1983) and Anderson (2003) for book-length treatments. With the appearance of big data in recent years many statistical methods have to be modified to adjust the new structure of data. In the area of high-dimensional tests on multivariate normal distributions, there are some recent work to test mean vectors and covariance matrices, see, for example, Schott (2001, 2005, 2007), Ledoit and Wolf (2002), Srivastava (2005), Bai *et al.* (2009), Chen *et al.* (2010), Jiang *et al.* (2012), Jiang and Yang (2013) and Jiang and Qi (2015).

In this paper we will focus on two high-dimensional LRT problems. Let $\mathbf{x}_1, \cdots, \mathbf{x}_n$ be i.i.d. $\mathbb{R}^p$-valued random variables with normal distribution $N_p(\mu, \mathbf{\Sigma})$, where $\mu \in \mathbb{R}^p$ is the mean vector and $\mathbf{\Sigma}$ is the covariance matrix. We are interesting in the following tests:

$$H_0: \ \mathbf{\Sigma} = \mathbf{I}_p \ \text{ vs } \ H_a: \ \mathbf{\Sigma} \neq \mathbf{I}_p \tag{1.1}$$

with μ unspecified, and

$$H_0: \quad \mu = \mathbf{0} \text{ and } \mathbf{\Sigma} = \mathbf{I}_p \ \text{ vs } \ H_a: \quad \mu \neq \mathbf{0} \text{ or } \mathbf{\Sigma} \neq \mathbf{I}_p. \tag{1.2}$$

The test (1.1) is equivalent to a seemingly more general test that $\mathbf{\Sigma}$ is any given matrix. The test in (1.2) is essentially the same one as testing that $(\mu, \mathbf{\Sigma})$ is equal to any given pair. More details are given later. For large classes of alternative tests, as p is large, we prove that the LRT statistics satisfy the central limit theorems (CLT) with explicit means and covariance matrices. It is interesting to see that the alternatives are not needed to be local ones. They contain wide classes of covariance matrices in the first test and mean vectors and covariance matrices in the second test.

To make our presentation clear, we discuss the two tests in two different sections. Remarks are given afterwards.

1.1 Testing Specified Value for Covariance Matrix

Consider the test $H_0 : \mathbf{\Sigma} = \mathbf{\Sigma}_0$ where $\mathbf{\Sigma}_0$ is a known and non-singular matrix and μ is unspecified. Set $\mathbf{y}_i = \mathbf{\Sigma}_0^{-1/2}\mathbf{x}_i$ for $i = 1, 2 \cdots, n$. Then $\mathbf{y}_1, \cdots, \mathbf{y}_n$ are i.i.d. random vectors with distribution $N_p(\tilde{\mu}, \mathbf{I}_p)$, where $\tilde{\mu} = \mathbf{\Sigma}_0^{-1/2}\mu$. So this test is equivalent to (1.1).

Let

$$\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \ \text{ and } \ \mathbf{A} = \sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^*. \tag{1.3}$$

Evidently, $\mathbf{A}$ is a $p \times p$ matrix. Define

$$\Lambda_n^* = \Big(\frac{e}{n-1}\Big)^{(n-1)p/2} e^{-tr(\mathbf{A})/2} \cdot |\mathbf{A}|^{(n-1)/2}. \tag{1.4}$$

The quantity Λ_n^* is a modified version of LRT statistic. The advantage is that the rejection region $\{\Lambda_n^* \leq c_\alpha\}$ is unbiased; See, for example, page 357 from Muirhead (1982). The Wilks phenomenon is known for Λ_n^*: under H_0 from (1.1), for fixed dimension p,

$$-2\log\Lambda_n^* \text{ converges weakly to } \ \chi^2(f) \tag{1.5}$$

with $f = p(p+1)/2$ as $n \to \infty$. See, for example, Theorem 8.4.9 from Muirehead (1982) plus the Slutsky lemma. A limiting chi-square distribution is also obtained under the alternative hypothesis; see, for example, Theorem 8.4.10 from Muirhead (1982).

It is seen from Figure 1 that the approximation (1.5) is far from accurate as p is large. In fact, mathematically the χ^2-approximation is no longer true as $p \to \infty$. We will prove that it satisfies a central limit theorem under the null hypothesis and the alternative hypothesis as well. Let us first introduce some notation before the statement of the CLT.

For a $p \times p$ non-negative definite matrix $\mathbf{\Sigma}$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$. Let $\|\mathbf{\Sigma}\| = \lambda_1$ be the spectral norm, $\|\mathbf{\Sigma}\|_2 = (\sum_{i=1}^p \lambda_i^2)^{1/2}$ be the Frobenius norm and $|\mathbf{\Sigma}|$ be the determinant of $\mathbf{\Sigma}$. The symbol $\mathbf{\Sigma} > 0$ indicates that $\mathbf{\Sigma}$ is a positive definite matrix.

In the paper, when there is no confusion, we sometimes write p for p_n and $\mathbf{\Sigma}$ for $\mathbf{\Sigma}_n$, which is a $p \times p$ matrix. This enables our formulas to appear short.

THEOREM 1 *Recall Λ_n^* in (1.4). Assume $n > p_n + 1$ for $n \geq 3$ and $\lim_{n\to\infty} p_n = \infty$. If $\mathbf{\Sigma}_n$ is non-singular for each n and $\sup_{n\geq 3}\{\|\mathbf{\Sigma}_n\|\} < \infty$, then $(\log \Lambda_n^* - \mu_n)/(n\sigma_n) \to N(0,1)$ in distribution as $n \to \infty$, where*

$$\mu_n = -\frac{1}{4}(n-1)\big(2n-2p-3\big)\log\big(1-\frac{p}{n-1}\big) + \frac{1}{2}(n-1)\cdot\big(\log|\mathbf{\Sigma}_n| - tr(\mathbf{\Sigma}_n)\big);$$
$$\sigma_n^2 = -\frac{1}{2}\Big[\frac{p}{n-1} + \log\Big(1-\frac{p}{n-1}\Big)\Big] + \frac{1}{2n}\cdot tr\big((\mathbf{\Sigma}_n - \mathbf{I})^2\big)$$

and p denotes p_n for brevity.

Theorem 1 is proved through a combination of tools from Jiang and Yang (2013) and Jiang and Qi (2015). Note that the alternatives in Theorem 1 are not local alternatives ($\mathbf{\Sigma}_n$ has to be very close to $\mathbf{I}$), this seems a surprise. More importantly, the theorem allows us to evaluate the test with more options of alternatives.

Now, by taking $\mathbf{\Sigma}_n = \mathbf{I}$ in Theorem 1 we immediately obtain the following result.

COROLLARY 1 *Let $\mathbf{x}_1, \cdots, \mathbf{x}_n$ be i.i.d. random vectors with normal distribution $N_p(\mu, \mathbf{\Sigma})$. Assume H_0 in (1.1) holds. Suppose that $n > p_n + 1$ for all $n \geq 3$ and $\lim_{n\to\infty} p_n = \infty$. Then $(\log \Lambda_n^* - \mu_{n,0})/(n\sigma_{n,0})$ converges in distribution to $N(0,1)$ as $n \to \infty$, where*

$$\mu_{n,0} = -\frac{1}{4}(n-1)\Big[2p + \big(2n-2p-3\big)\log\big(1-\frac{p}{n-1}\big)\Big];$$
$$\sigma_{n,0}^2 = -\frac{1}{2}\Big[\frac{p}{n-1} + \log\Big(1-\frac{p}{n-1}\Big)\Big]$$

and p denotes p_n for brevity.

Assuming H_0 in (1.1), Zheng et al. (2015) obtain a central limit theorem for Λ_n^* under the assumption $\lim_{n\to\infty}\frac{p_n}{n} = y \in (0,1)$ and that $\mathbf{x}_1, \cdots, \mathbf{x}_n$ is a random sample not necessarily from a Gaussian distribution. In terms of population distributions, their result is more general than Corollary 1 because we study a normal population. In terms of the restriction on dimension p, Corollary 1 is more general, in particular, the corollary includes the case $y = 0$ and $y = 1$. Our emphasis is Theorem 1, which is true under both the null and the alternative hypotheses.

We run a simulation in Section 2 to discuss the sizes and powers of the tests based on the χ^2-approximation (1.5) and the CLT from Corollary 1. The discussions are presented in the same section.

For the LRT in (1.1), the LRT rejection region based on Λ_n^* is $R = \{\Lambda_n^* \le c\}$ for some value of c. Then, according to Corollary 1, the asymptotic size-α test is given by $R = \{\log \Lambda_n^* \le c_\alpha\}$ with $c_\alpha = \mu_{n,0} + n\sigma_{n,0}\Phi^{-1}(\alpha)$, where $\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^{x} e^{-t^2/2}\, dt$. By Theorem 1, the power function for the test is

$$\beta(\mathbf{\Sigma}_n) = P(\log \Lambda_n^* \le c_\alpha | \mathbf{\Sigma}_n) \sim \Phi\big(\frac{c_\alpha - \mu_n}{n\sigma_n}\big), \qquad \mathbf{\Sigma}_n > 0, \tag{1.6}$$

where μ_n and σ_n are as in Theorem 1. In particular,

$$\lim_{n\to\infty} \beta(\mathbf{\Sigma}_n) \to 1 \ \text{ if } \quad \lim_{n\to\infty} \frac{\mathrm{tr}(\mathbf{\Sigma}_n) - \log|\mathbf{\Sigma}_n| - p}{\sigma_n} = \infty. \tag{1.7}$$

Now, let us investigate a special alternative H_a. Let $0 < C_1 < C_2 < 1$ be two constants such that $C_1 \le \frac{p}{n} \le C_2$ for all $n \ge 3$. Suppose $a_n = 1 + \frac{c_n}{\sqrt{p}} > 0$, where $-\sqrt{p} < c_n \le A\sqrt{p}$ for all $n \ge 3$ and $A > 0$ is a constant not depending on n. Assuming $\mathbf{\Sigma}_n = a_n \mathbf{I}_p$. Then

$$\lim_{n\to\infty} \beta(\mathbf{\Sigma}_n) = \begin{cases} 1, & \text{if } \lim_{n\to\infty} |c_n| = \infty; \\ \alpha, & \text{if } \lim_{n\to\infty} c_n = 0. \end{cases} \tag{1.8}$$

The second statement says that H_a is almost H_0 if a_n is close to 1 enough, and hence the power is almost the type I error. In other words, we obtain the cut-off point such that the power goes to 1 or not around this point. The statements (1.7) and (1.8) will be verified in Section 3.4.

1.2 Testing Specified Values for Mean Vector and Covariance Matrix

Recall the notation used in the previous section: $\mathbf{x}_1, \cdots, \mathbf{x}_n$ are i.i.d. $\mathbb{R}^p$-valued random vectors from a normal distribution $N_p(\mu, \mathbf{\Sigma})$, where $\mu \in \mathbb{R}^p$ is the mean vector and $\mathbf{\Sigma}$ is the $p \times p$ covariance matrix. Consider the hypothesis test:

$$H_0: \quad \mu = \mu_0 \text{ and } \mathbf{\Sigma} = \mathbf{\Sigma}_0 \quad \text{vs} \quad H_a : H_0 \text{ is not true},$$

where μ_0 is a specified vector in $\mathbb{R}^p$ and $\mathbf{\Sigma}_0$ is a specified $p \times p$ non-singular matrix. By applying the transformation $\tilde{\mathbf{x}}_i = \mathbf{\Sigma}_0^{-1/2}(\mathbf{x}_i - \mu_0)$, the above test is equivalent to the simpler one in (1.2).

Let $\bar{\mathbf{x}}$ be the sample mean and $\mathbf{A}$ be the normalized covariance matrix as defined in (1.3). Assume $n > p$. Then the LRT statistic for H_0 in (1.2) is given by

$$\Lambda_n = \Big(\frac{e}{n}\Big)^{np/2} |\mathbf{A}|^{n/2} \cdot \exp\Big\{ -\frac{1}{2} tr(\mathbf{A}) - \frac{1}{2} n\overline{\mathbf{x}}'\overline{\mathbf{x}}\Big\}. \tag{1.9}$$

The test based on Λ_n is unbiased; see, e.g., Theorem 8.5.1 from Muirhead (1982). The condition that $n > p$ is required to ensure that Λ_n is non-degenerate. For a large class of alternative hypothesis, we obtain the following CLT.

Theorem 2 *Assume $n > p_n + 1$ for $n \ge 3$ and $\lim_{n\to\infty} p_n = \infty$. Let Λ_n be as in (1.9) and $\mathbf{x}_1 \sim N_p(\mu, \mathbf{\Sigma}_n)$ with $\mathbf{\Sigma}_n > 0$ and $\sup\{\|\mathbf{\Sigma}_n\|; n \ge 3\} < \infty$. Set*

$$\mu_n = -\frac{1}{4}\Big[n(2n - 2p - 3)\log(1 - \frac{p}{n-1}) + 2p\Big] + \frac{1}{2} n \cdot \big[\log|\mathbf{\Sigma}_n| - tr(\mathbf{\Sigma}_n)\big] - \frac{1}{2} n\|\mu\|^2;$$
$$\sigma_n^2 = -\frac{1}{2}\Big[\frac{p}{n-1} + \log(1 - \frac{p}{n-1})\Big] + \frac{1}{2n} \cdot tr\big[(\mathbf{\Sigma}_n - \mathbf{I})^2\big] + \frac{1}{n}(\mu'\mathbf{\Sigma}_n\mu)$$

and p denotes p_n for brevity. Then $(\log \Lambda_n - \mu_n)/(n\sigma_n) \to N(0, 1)$ in distribution provided one of the following holds: (a) $\|\mu\| = o\big((p^3/n)^{1/2}\big)$; (b) $\|\mathbf{\Sigma}_n^{-1}\| = o(p)$.

Here is a warning of the notation: μ is the mean vector from the normal population $N_p(\mu, \mathbf{\Sigma}_n)$, but $\mu_n \in \mathbb{R}$ is the asymptotic mean of the test statistic $\log \Lambda_n$. This should be clear from the context now and later.

Since $x + \log(1-x) < 0$ for all $x \in (0,1)$ and $\mu'\mathbf{\Sigma}\mu \geq 0$ due to the fact that $\mathbf{\Sigma}$ is non-negative definite, we know $\sigma_n > 0$. By taking $\mathbf{\Sigma} = \mathbf{I}$ and $\mu = \mathbf{0}$ in Theorem 2, we obtain Theorem 5 from Jiang and Qi (2015), which studies the asymptotic distribution of Λ_n under H_0.

The assumption in Theorem 2 can be easily satisfied. Given $\sup\{\|\mathbf{\Sigma}_n\|;\, n \geq 3\} < \infty$, the CLT holds if (a) $\|\mu\|$ is not very large but $\mathbf{\Sigma}_n$ is arbitrary or (b) the smallest eigenvalue of $\mathbf{\Sigma}_n$ is not very close to zero but μ is arbitrary. In particular, Theorem 2 holds if all of the eigenvalues of $\mathbf{\Sigma}_n$ lie in $[a, b]$ for any $n \geq 3$, where $0 < a < b < \infty$ are constants not depending on n.

As elaborated above, the alternative hypothesis in Theorem 2 is not a local one (μ has to be very close to $\mathbf{0}$ and $\mathbf{\Sigma}$ has to be very close to $\mathbf{I}$). Our central limit theorem of the LRT statistic holds for a big class of $(\mu, \mathbf{\Sigma})$.

Now we give the power of the LRT in (1.2). Take $\mu = \mathbf{0} \in \mathbb{R}^p$ and $\mathbf{\Sigma}_n = \mathbf{I}_p$ in the expressions of μ_n and σ_n from Theorem 2, we get

$$\mu_{n,0} = -\frac{1}{4}\Big[n(2n-2p-3)\log(1-\frac{p}{n-1}) + 2(n+1)p\Big];$$
$$\sigma_{n,0}^2 = -\frac{1}{2}\Big[\frac{p}{n-1} + \log(1-\frac{p}{n-1})\Big].$$

By Theorem 2, the asymptotic size-α test is given by $R = \{\log \Lambda_n \leq c_\alpha\}$ with $c_\alpha = \mu_{n,0} + n\sigma_{n,0}\Phi^{-1}(\alpha)$, where $\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^{x} e^{-t^2/2}\, dt$. Based upon Theorem 2, the asymptotic power function in terms of μ and $\mathbf{\Sigma}_n$ is given by

$$\beta(\mu, \mathbf{\Sigma}_n) = P(\log \Lambda_n^* \leq c_\alpha | \mu, \mathbf{\Sigma}_n) \sim \Phi\big(\frac{c_\alpha - \mu_n}{n\sigma_n}\big), \qquad \mu \in \mathbb{R}^p, \quad \mathbf{\Sigma}_n > 0, \tag{1.10}$$

where μ_n and σ_n are as in Theorem 2. This implies that

$$\lim_{n\to\infty} \beta(\mu, \mathbf{\Sigma}_n) \to 1 \quad \text{if} \quad \lim_{n\to\infty} \frac{\mathrm{tr}(\mathbf{\Sigma}_n) - \log|\mathbf{\Sigma}_n| - p + \|\mu\|^2}{\sigma_n} = \infty. \tag{1.11}$$

Let us look into a special alternative H_a. Let $0 < C_1 < C_2 < 1$ be two constants such that $C_1 \leq \frac{p}{n} \leq C_2$ for all $n \geq 3$. Suppose $a_n = 1 + \frac{c_n}{\sqrt{p}} > 0$, where $-\sqrt{p} < c_n \leq A\sqrt{p}$ for all $n \geq 3$ and $A > 0$ is a constant not depending on n. Assuming $\mathbf{\Sigma}_n = a_n \mathbf{I}_p$. Then

$$\lim_{n\to\infty} \beta(\mu, \mathbf{\Sigma}_n) = \begin{cases} 1, & \text{if } \lim_{n\to\infty} |c_n| = \infty \text{ and } \|\mu\| = O(\sqrt{p}); \\ \alpha, & \text{if } \lim_{n\to\infty} c_n = 0 \text{ and } \|\mu\| = o(1). \end{cases} \tag{1.12}$$

The second statement says that H_a is almost H_0 if a_n is close to 1 enough and $\|\mu\|$ is close to 0 enough. The power is then almost the type I error. On th other hand, if $\mathbf{\Sigma}_n$ is far enough from $\mathbf{I}_n$ and $\|\mu\|$ is not too large, the power goes to 1.

The statements (1.11) and (1.12) will be verified in Section 3.4.

Lastly we make some remarks.

1. The CLT's of six LRT statistics under null hypotheses are obtained by Jiang and Yang (2013) and Jiang and Qi (2015). In this paper, we have worked on two of them under alternative hypotheses. We expect similar CLT's for other test statistics also hold, however, the technical tools will be different. An understanding of the zonal polynomials (a special case of Jack polynomials) and hypergeometric functions of matrix arguments will be likely needed.

2. The study of the powers of hypothesis testing problems for fixed population dimensions is very rich. In comparison, there are not so many in the high-dimensional testing problems. This is because the mathematics becomes more involved when both dimension p and sample size n are large. This paper provides such type of results in the high-dimensional setting. Among other results in this direction, Onatski *et al.* (2014) investigate the powers of sphericity tests.

3. In this paper we investigate two hypothesis tests by using the method of LRT. Assuming that a random sample also comes from a normal population, Srivastava (2005) studies a couple of statistical tests. In particular, for H_0 from (1.1) the author obtains a central limit theorem for a different test statistic. The theorem allows the case $p > n$.

4. Based on our main results, some optimal hypothesis tests can be carried out. This is because our theories give CLT's under big classes of alternative hypotheses. One is refereed to, for example, Cai and Ma (2013) for a general strategy.

Finally, we describe the organization of the rest of paper. A simulation study is given in Section 2 to discuss the sizes and powers of the tests. The proof of Theorem 1 is presented in Section 3.2, and the proof of Theorem 2 is arranged in Section 3.3. Some facts are verified in Section 3.4.

2 Simulation Study: Sizes and Powers

In this section we study the sizes and powers of the tests based on the classical χ^2-test in (1.5) and the CLT in Theorem 1. The notation $\lfloor x \rfloor$ stands for the integer part of $x > 0$. Looking at Table 1 and Figure 1, we make a discussion as follows.

Table 1 gives the performance of the CLT in Corollary 1 and the classical χ^2-approximation on Λ_n^* as in (1.5). The chi-square test is better but the CLT is also reasonable as p is small. When p grows, the CLT surpasses the chi-square test and performs very well. This phenomenon can also been visualized from Figure 1: the histograms and their limiting χ^2-curves go in opposite directions and become farther as p become larger. In contrast, the CLT fits the histogram almost perfectly.

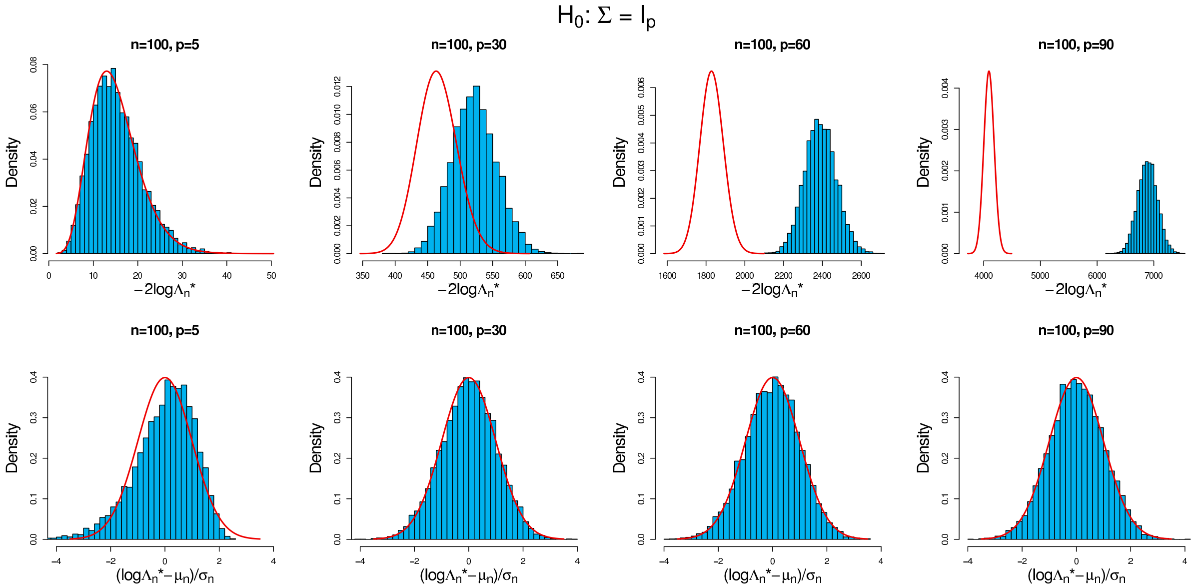

Figure 1: *Comparison between (1.5) and Corollary 1. We choose $n = 100$ with $p = 5, 30, 60, 90$. The pictures in the top row show that the χ^2 curves stay farther away from the histogram of $-2\log\Lambda_n^*$ as p grows. The bottom row shows that the $N(0,1)$-curve fits the histogram of $(\log\Lambda_n^* - \mu_n)/(n\sigma_n)$ better as p becomes larger.*

Table 1: Size and Power of LRT from Theorem 1 and (1.5)

	Size under H_0		Power under H_a	
	CLT	χ^2 approx.	CLT	χ^2 approx.
$n = 100, p = 5$	0.0808	0.0592	0.7142	0.6656
$n = 100, p = 30$	0.0565	0.5573	0.9152	0.9987
$n = 100, p = 60$	0.0546	1.0000	0.8482	1.0000
$n = 100, p = 90$	0.0497	1.0000	0.6073	1.0000

The sizes (alpha errors) are estimated based on $10,000$ simulations from $N_p(\mathbf{0}, \mathbf{I}_p)$. The powers are estimated under the alternative hypothesis that $\mathbf{\Sigma} = diag(1.44, \cdots, 1.44, 1, \cdots, 1)$, where the number of 1.44 on the diagonal is equal to $\lfloor p/2 \rfloor$.

3 Proofs

To prove the main results, we prepare some technical tools in Section 3.1. The proofs of Theorems 1 and 2 are presented in Sections 3.2 and 3.3, respectively. Finally, we verify some basic facts in Section 3.4.

3.1 Some Tools

Let $\Gamma(z)$ be the Gamma function defined on the complex plane $\mathbb{C}$. Define

$$\Gamma_p(z) := \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\Big(z - \frac{1}{2}(i-1)\Big) \tag{3.1}$$

for complex number z with $\mathrm{Re}(z) > \frac{1}{2}(p-1)$; see p. 62 from Muirhead (1982). The machineries in the derivations of our CLTs are the following on $\Gamma_p(z)$.

Lemma 3.1 *(Lemma 5.4 from Jiang and Yang, 2013) Let $\Gamma_p(z)$ be as in (3.1). Let $n > p = p_n$ and $r_n = (-\log(1-\frac{p}{n}))^{1/2}$. Assume $\frac{p}{n} \to y \in (0,1]$ and $s = s_n = O(1/r_n)$ and $t = t_n = O(1/r_n)$ as $n \to \infty$. Then*

$$\log \frac{\Gamma_p(\frac{n}{2}+t)}{\Gamma_p(\frac{n}{2}+s)} = p(t-s)(\log n - 1 - \log 2) + r_n^2\Big[(t^2 - s^2) - \Big(p - n + \frac{1}{2}\Big)(t-s)\Big] + o(1)$$

as $n \to \infty$.

When $\frac{p}{n} \to 0$, the analogue of Lemma 3.1 is the following.

Lemma 3.2 *(Proposition 5.1 from Jiang and Qi, 2015) Let $\{p = p_n \in \mathbb{N};\ n \geq 1\}$, $\{m = m_n \in \mathbb{N};\ n \geq 1\}$ and $\{t_n \in \mathbb{R};\ n \geq 1\}$ satisfy that (i) $p_n \to \infty$ and $p_n = o(n)$; (ii) there exists $\epsilon \in (0,1)$ such that $\epsilon \leq m_n/n \leq \epsilon^{-1}$ for large n; (iii) $t = t_n = O(n/p)$. Then, as $n \to \infty$,*

$$\log \frac{\Gamma_p(\frac{m-1}{2}+t)}{\Gamma_p(\frac{m-1}{2})} = \alpha_n t + \beta_n t^2 + \gamma_n(t) + o(1)$$

where

$$\alpha_n = -\big[2p + (m - p - \frac{3}{2})\log(1 - \frac{p}{m-1})\big]; \quad \beta_n = -\big[\frac{p}{m-1} + \log\big(1 - \frac{p}{m-1}\big)\big];$$
$$\gamma_n(t) = p\Big[\big(\frac{m-1}{2} + t\big)\log\big(\frac{m-1}{2} + t\big) - \frac{m-1}{2}\log\frac{m-1}{2}\Big].$$

3.2 Proof of Theorem 1

The methodology of our proof is the analysis of the moment generating function (m.g.f.) of $\log \Lambda_n^*$. We will show that it converges to $N(0,1)$. To do so, we first need to get the exact expression of its m.g.f., which is essentially the moment of Λ_n^*.

Lemma 3.3 *(Theorem 8.4.7 from Muirhead (1982)). Let Λ_n^* be given at (1.4). Then*

$$E[(\Lambda_n^*)^h] = \Big(\frac{2e}{q}\Big)^{pqh/2} \frac{|\mathbf{\Sigma}|^{qh/2}}{|\mathbf{I} + h\mathbf{\Sigma}|^{q(1+h)/2}} \frac{\Gamma_p\big(\frac{q(1+h)}{2}\big)}{\Gamma_p\big(\frac{q}{2}\big)}$$

for all $h > \frac{p-1}{q} - 1$, where $q = n - 1$.

The proof of Theorem 1 consists of two steps. The first one is to prove a special case of the theorem with $\mathbf{\Sigma}_n = \mathbf{I}$. We then prove the general situation afterwards.

LEMMA 3.4 *Theorem 1 holds if* $\mathbf{\Sigma}_n = \mathbf{I}$.

Proof of Lemma 3.4. We need to show

$$J_n := \frac{\log \Lambda_n^* - \mu_n}{n\sigma_n} \text{ converges to } N(0,1) \tag{3.2}$$

in distribution as $n \to \infty$. Equivalently, it suffices to prove that for any subsequence $\{n_k\}$, there is a further subsequence $\{n_{k_j}\}$ such that $J_{n_{k_j}}$ converges to $N(0,1)$ in distribution as $j \to \infty$. Now, noticing $p_n/n \in [0,1]$ for all n, for any subsequence n_k, take a further subsequence n_{k_j} such that $p_{n_{k_j}}/n_{k_j} \to y \in [0,1]$. So, without loss of generality, we only need to show (3.2) under the condition that $\lim_{n\to\infty} p_n/n = y \in [0,1]$.

Case 1: $y \in (0,1]$. First, since $\log(1-x) < -x$ for all $x < 1$, we know $\sigma_n^2 > 0$ for all $n > p+1$. Now, by assumption, it is easy to see

$$\lim_{n\to\infty} \sigma_n^2 = \begin{cases} -\frac{1}{2}\big[y + \log(1-y)\big], & \text{if } y \in (0,1); \\ \infty, & \text{if } y = 1, \end{cases}$$

which is always positive. Consequently,

$$\delta_0 := \inf\{\sigma_n;\, n \geq 3\} > 0.$$

To finish the proof, according to the method of the m.g.f. (see, e.g., page 408 from Billingsley, 1986), it suffices to show that

$$E\exp\Big\{\frac{\log \Lambda_n^* - \mu_n}{n\sigma_n}s\Big\} \to e^{s^2/2} \tag{3.3}$$

as $n \to \infty$ for all s such that $|s| < \delta_0/2$. Or equivalently,

$$\log E[(\Lambda_n^*)^t] - \mu_n t \to \frac{s^2}{2} \tag{3.4}$$

as $n \to \infty$, where $t = t_n := \frac{s}{n\sigma_n}$.

Now fix $|s| < \delta_0/2$. Then, by the definition of δ_0, we have $t > -\frac{1}{n} \geq \frac{p-1}{n} - 1$ for all $n \geq 3$. Set $q = n-1$. Take $\mathbf{\Sigma} = \mathbf{I}_p$ in Lemma 3.3 to see that

$$\log E[(\Lambda_n^*)^t] = \frac{pqt}{2}\big(1 + \log\frac{2}{q}\big) - \frac{pq(1+t)}{2}\log(1+t) + \log\frac{\Gamma_p\big(\frac{q(1+t)}{2}\big)}{\Gamma_p\big(\frac{q}{2}\big)}.$$

Let $r_q = (-\log(1-\frac{p}{q}))^{1/2}$. Evidently, $|\frac{1}{2}qt| \leq |\frac{s}{2\sigma_n}| = O(r_q^{-1})$ as $n \to \infty$. From Lemma 3.1 we see

$$\log\frac{\Gamma_p\big(\frac{q(1+t)}{2}\big)}{\Gamma_p\big(\frac{q}{2}\big)} = \frac{pqt}{2}\big(-1 + \log\frac{q}{2}\big) + r_q^2\Big[\frac{q^2t^2}{4} - \big(p - q + \frac{1}{2}\big)\frac{qt}{2}\Big] + o(1)$$

as $n \to \infty$. Now, expand $u(t) := (1+t)\log(1+t)$ at 0 by the Taylor expansion to obtain

$$-\frac{pq}{2}(1+t)\log(1+t) = -\frac{pq}{2}\Big(t + \frac{t^2}{2} + O(t^3)\Big) = -\frac{pq}{2}\Big(t + \frac{t^2}{2}\Big) + o(1)$$

since $t = O(n^{-1})$. This together with the above two formulas implies that

$$\begin{aligned}
\log E[(\Lambda_n^*)^t] &= -\frac{pq}{2}\big(t+\frac{t^2}{2}\big) + r_q^2\Big[\frac{q^2t^2}{4} - \big(p-q+\frac{1}{2}\big)\frac{qt}{2}\Big] + o(1) \\
&= \frac{q^2}{4}\Big(-\frac{p}{q} - \log\big(1-\frac{p}{q}\big)\Big)t^2 + \big(-\frac{1}{4}q\big)\Big[2p + (2q-2p-1)\log\big(1-\frac{p}{q}\big)\Big]t \\
&\quad + o(1) \\
&= \frac{q^2}{4}\cdot(2\sigma_n^2)t^2 + \mu_n t + o(1) = \frac{s^2}{2} + \mu_n t + o(1)
\end{aligned}$$

as $n \to \infty$. This gives (3.4).

Case 2: $y = 0$. Similar to *Case 1*, to prove the theorem, it is enough to show that

$$E\exp\Big\{\frac{\log\Lambda_n^* - \mu_n}{n\sigma_n}s\Big\} = \exp(-\frac{\mu_n s}{n\sigma_n})E[(\Lambda_n^*)^{\frac{s}{n\sigma_n}}] \to e^{s^2/2}$$

as $n \to \infty$ for all $s \in [-1, 1]$. Again, set $t = \frac{s}{n\sigma_n}$ and $q = n-1$. Then the above is equivalent to

$$\log E[(\Lambda_n^*)^t] - \mu_n t \to \frac{s^2}{2} \tag{3.5}$$

as $n \to \infty$ for all $s \in [-1, 1]$. Evidently, $\sigma_n \sim \frac{p}{2q}$ and hence $t \sim \frac{2s}{p} > \frac{p-1}{q} - 1$ as n is sufficiently large. Using Lemma 3.3 again, we get that

$$\log E[(\Lambda_n^*)^t] = \frac{pqt}{2}\big(1+\log\frac{2}{q}\big) - \frac{pq(1+t)}{2}\log(1+t) + \log\frac{\Gamma_p\big(\frac{q(1+t)}{2}\big)}{\Gamma_p\big(\frac{q}{2}\big)}.$$

Write $\frac{q(1+t)}{2} = \frac{n-1}{2} + \frac{qt}{2}$. By the fact $\frac{qt}{2} = O(n/p)$ we get from Lemma 3.2 that

$$\log\frac{\Gamma_p\big(\frac{q(1+t)}{2}\big)}{\Gamma_p\big(\frac{q}{2}\big)} = \alpha_n\frac{qt}{2} + \beta_n\Big(\frac{qt}{2}\Big)^2 + \gamma_n(t) + o(1)$$

where

$$\begin{aligned}
&\alpha_n = -\big[2p + (n-p-\frac{3}{2})\log(1-\frac{p}{n-1})\big]; \quad \beta_n = -\big[\frac{p}{n-1} + \log\big(1-\frac{p}{n-1}\big)\big]; \\
&\gamma_n(t) = p\Big[\frac{(n-1)(1+t)}{2}\log\frac{(n-1)(1+t)}{2} - \frac{n-1}{2}\log\frac{n-1}{2}\Big].
\end{aligned}$$

Observe that

$$\gamma_n(t) = \frac{p(n-1)t}{2}\log\frac{n-1}{2} + \frac{p(n-1)}{2}(1+t)\log(1+t).$$

Joining all of the above assertions it is not difficult to check that

$$\log E[(\Lambda_n^*)^t] = -\frac{1}{4}q(-2\alpha_n - 2p)t + \frac{1}{2}\cdot\frac{\beta_n q^2t^2}{2} + o(1) = \mu_n t + \frac{s^2}{2} + o(1)$$

as $n \to \infty$ since $\beta_n = 2\sigma_n^2$ and $t = \frac{s}{n\sigma_n}$. This leads to (3.5). ■

To finish the proof of Theorem 1, we need a couple of technical results.

Lemma 3.5 *Let p satisfy that $1 \le p < n-1$ for all $n \ge 3$. Let $\sigma_n > 0$ be such that $\sigma_n^2 = -\frac{1}{2}[\frac{p}{n-1} + \log(1-\frac{p}{n-1})]$. Then, $\sigma_n \ge \frac{1}{2}\frac{p}{n-1}$ for all $n \ge 3$ and*

$$\delta := \inf_{n\ge 3}\Big\{n\big(1-\frac{p-1}{n-1}\big)\sigma_n\Big\} \ge \frac{1}{4}.$$

Proof. First, since $\log(1-x) = -\sum_{k=1}^{\infty}\frac{x^k}{k}$ for $x<1$. We know $\log(1-x) \le -x - \frac{x^2}{2}$ for $x\in[0,1)$. This implies that $\sigma_n \ge \frac{1}{2}\frac{p}{n-1}$. Hence,

$$\begin{aligned} 2\delta &\ge \inf_{n\ge 3}\Big\{(n-1)\big(1-\frac{p-1}{n-1}\big)\cdot\frac{p-1}{n-1}\Big\} \\ &\ge \inf_{n\ge 3,\,1\le x\le n-2}\Big\{x\big(1-\frac{x}{n-1}\big)\Big\}. \end{aligned}$$

The function $f(x) := x\big(1-\frac{x}{n-1}\big)$ defined on $[1, n-2]$ has to take its minimum at either 1, $n-2$ or its stable point $\frac{n-1}{2}\in[1,n-2]$. Thus, $\min_{1\le x\le n-2} f(x) = \min\{f(1), f(n-2), f(\frac{n-1}{2})\} = 1-\frac{1}{n-1}$. This says that

$$\delta \ge \frac{1}{2}\cdot\inf_{n\ge 3}\big\{1-\frac{1}{n-1}\big\} = \frac{1}{4}.$$

The proof is complete. ■

The following lemma is a key part in the derivation of the CLTs of the LRT statistics under alternative hypotheses on covariance matrices.

Lemma 3.6 *Let $\mathbf{\Sigma}$ be a $p\times p$ non-negative definite matrix. Write*

$$\log\Big(\frac{(1+t)^p}{|\mathbf{I}+t\mathbf{\Sigma}|}\Big)^{(1+t)/2} = \frac{1}{2}[p-tr(\mathbf{\Sigma})]\,t + \frac{1}{4}tr[(\mathbf{\Sigma}-\mathbf{I})^2]\,t^2 + D\cdot t^3\,tr[(\mathbf{\Sigma}-\mathbf{I})^2].$$

Then, $|D| \le 4\|\mathbf{\Sigma}\|+5$ as $|t|\le(2\|\mathbf{\Sigma}\|+3)^{-1}$.

Proof. Let $\lambda_1,\cdots,\lambda_p$ be the eigenvalues of $\mathbf{\Sigma}$. Then

$$\log\Big(\frac{(1+t)^p}{|\mathbf{I}+t\mathbf{\Sigma}|}\Big)^{(1+t)/2} = -\frac{1+t}{2}\cdot\sum_{i=1}^{p}\log\Big(1+\frac{t}{1+t}(\lambda_i-1)\Big). \tag{3.6}$$

Write $\log(1+x) = x-\frac{1}{2}x^2+x^3\cdot\epsilon(x)$ for $|x|<1$. Then, by Taylor's expansion, $\epsilon(x) = \sum_{i=0}^{\infty}(-1)^i\frac{1}{i+3}x^i$. Thus, $|\epsilon(x)|\le\frac{1}{3}\sum_{i=0}^{\infty}|x|^i = \frac{1}{3(1-|x|)}$ for all $|x|<1$. So $\sup_{|x|\le 1/2}|\epsilon(x)|\le 1$. Observe $\max_{1\le i\le p}\big|\frac{t}{1+t}(\lambda_i-1)\big| \le \frac{|t|}{1-|t|}(\|\mathbf{\Sigma}\|+1)\le\frac{1}{2}$ if $|t|\le(2\|\mathbf{\Sigma}\|+3)^{-1}$. Consequently,

$$\begin{aligned} &\log\Big(1+\frac{t}{1+t}(\lambda_i-1)\Big) \\ &= \frac{t}{1+t}(\lambda_i-1) - \frac{1}{2}\frac{t^2}{(1+t)^2}(\lambda_i-1)^2 + t^3(\lambda_i-1)^2C_i \end{aligned}$$

for $1\le i\le p$, where

$$C_i := \frac{\lambda_i-1}{(1+t)^3}\epsilon\Big(\frac{t}{1+t}(\lambda_i-1)\Big)$$

satisfies that $\max_{1\le i\le p}|C_i| \le (\|\mathbf{\Sigma}\|+1)(1-|t|)^{-3}\le 4(\|\mathbf{\Sigma}\|+1)$ as $|t|\le\frac{1}{3}$, which is true if $|t|\le(2\|\mathbf{\Sigma}\|+3)^{-1}$. In the rest of the proof, we always assume $|t|\le(2\|\mathbf{\Sigma}\|+3)^{-1}$. Then

$$\begin{aligned} &\sum_{i=1}^{p}\log\Big(1+\frac{t}{1+t}(\lambda_i-1)\Big) \\ =\ &\frac{t}{1+t}[\mathrm{tr}(\mathbf{\Sigma})-p] - \frac{1}{2}\frac{t^2}{(1+t)^2}\mathrm{tr}[(\mathbf{\Sigma}-\mathbf{I})^2] + t^3\,\mathrm{tr}[(\mathbf{\Sigma}-\mathbf{I})^2]\cdot C \end{aligned}$$

where $|C| \le 4(\|\mathbf{\Sigma}\| + 1)$. This and (3.6) imply

$$\begin{aligned} & \log\Big(\frac{(1+t)^p}{|\mathbf{I}+t\mathbf{\Sigma}|}\Big)^{(1+t)/2} \\ = \;& \frac{1}{2}[p - \operatorname{tr}(\mathbf{\Sigma})]t + \frac{1}{4}\frac{t^2}{1+t}\operatorname{tr}[(\mathbf{\Sigma}-\mathbf{I})^2] + t^3 \operatorname{tr}[(\mathbf{\Sigma}-\mathbf{I})^2]\cdot C' \end{aligned}$$

with $|C'| \le |C|$. Let $\frac{1}{1+t} = 1 + \eta(t)$. Then $|\eta(t)| = \frac{|t|}{|1+t|} \le 2|t|$ as $|t| \le \frac{1}{3}$, which is guaranteed by the restriction $|t| \le (2\|\mathbf{\Sigma}\| + 3)^{-1}$. Hence,

$$\frac{1}{4}\frac{t^2}{1+t}\operatorname{tr}[(\mathbf{\Sigma}-\mathbf{I})^2] = \frac{t^2}{4}\operatorname{tr}[(\mathbf{\Sigma}-\mathbf{I})^2] + \frac{\eta(t)t^2}{4}\operatorname{tr}[(\mathbf{\Sigma}-\mathbf{I})^2]$$

and the absolute value of the last term is bounded by $\frac{1}{2}|t|^3 \operatorname{tr}[(\mathbf{\Sigma}-\mathbf{I})^2]$ as $|t| \le \frac{1}{3}$. Thus,

$$\begin{aligned} & \log\Big(\frac{(1+t)^p}{|\mathbf{I}+t\mathbf{\Sigma}|}\Big)^{(1+t)/2} \\ = \;& \frac{1}{2}[p - \operatorname{tr}(\mathbf{\Sigma})]t + \frac{1}{4}\operatorname{tr}[(\mathbf{\Sigma}-\mathbf{I})^2]t^2 + t^3 \operatorname{tr}[(\mathbf{\Sigma}-\mathbf{I})^2]\cdot C'' \end{aligned}$$

with $|C''| \le |C'| + \frac{1}{2}$. Our desired conclusion follows by noting $|C''| \le 4\|\mathbf{\Sigma}\| + 5$. ■

Proof of Theorem 1. First, set

$$\begin{aligned} \mu_{n,0} &= -\frac{1}{4}(n-1)\Big[2p + \big(2n-2p-3\big)\log\big(1 - \frac{p}{n-1}\big)\Big]; \\ \sigma_{n,0}^2 &= -\frac{1}{2}\Big[\frac{p}{n-1} + \log\Big(1 - \frac{p}{n-1}\Big)\Big] \end{aligned}$$

and

$$\begin{aligned} \mu_{n,1} &= \frac{1}{2}(n-1)\cdot\big(p + \log|\mathbf{\Sigma}| - \operatorname{tr}(\mathbf{\Sigma})\big); \\ \sigma_{n,1}^2 &= \frac{1}{2n}\cdot\operatorname{tr}\big[(\mathbf{\Sigma}-\mathbf{I})^2\big]. \end{aligned}$$

Then,

$$\mu_n = \mu_{n,0} + \mu_{n,1} \quad \text{and} \quad \sigma_n^2 = \sigma_{n,0}^2 + \sigma_{n,1}^2 \tag{3.7}$$

for all $n \ge 3$. We need to prove

$$J_n := \frac{\log\Lambda_n^* - \mu_n}{n\sigma_n} \text{ converges to } N(0,1) \tag{3.8}$$

in distribution as $n \to \infty$. Equivalently, it is enough to show that for any subsequence $\{n_k\}$, there is a further subsequence $\{n_{k_j}\}$ such that $J_{n_{k_j}}$ converges to $N(0,1)$ in distribution as $j \to \infty$. Now, noticing $\sigma_{n,0}^2/\sigma_n^2 \in [0,1]$ by (3.7) and $p_n/n \in [0,1]$ for all $n \ge 3$, for any subsequence n_k, take a further subsequence n_{k_j} such that $p_{n_{k_j}}/n_{k_j} \to y \in [0,1]$ and $\sigma_{n_{k_j},0}/\sigma_{n_{k_j}} \to y_0 \in [0,1]$. So, without loss of generality, we only need to show (3.8) under the condition that

$$\lim_{n\to\infty}\frac{\sigma_{n,0}}{\sigma_n} = y_0 \in [0,1] \quad \text{and} \quad \lim_{n\to\infty}\frac{p_n}{n} = y \in [0,1]. \tag{3.9}$$

Step 1. To prove (3.8), similar to the argument in (3.3), it is enough to show that

$$E\exp\Big\{\frac{\log\Lambda_n^* - \mu_n}{n\sigma_n}s\Big\} \to e^{s^2/2}$$

as $n \to \infty$ for all s such that

$$|s| < \frac{1}{4}, \tag{3.10}$$

or equivalently,

$$\log E[(\Lambda_n^*)^t] - \mu_n t \to \frac{s^2}{2} \tag{3.11}$$

as $n \to \infty$, where $t = t_n := \frac{s}{n\sigma_n}$.

Fix s with $|s| \le 1/4$. By the second conclusion of Lemma 3.5,

$$|\frac{s}{n\sigma_n}| \le \frac{1}{4n\sigma_{n,0}} \le 1 - \frac{p-1}{n-1}.$$

Let $\Lambda_{n,0}^*$ be a random variable of the distribution of Λ_n^* under H_0 in (1.1). Taking $\mathbf{\Sigma} = \mathbf{I}$ in Lemma 3.3, we are able to write

$$E[(\Lambda_n^*)^t] = E[(\Lambda_{n,0}^*)^t] \cdot M_{n,1} \tag{3.12}$$

where

$$\begin{aligned} E[(\Lambda_{n,0}^*)^t] &:= \Big(\frac{2e}{q}\Big)^{pqt/2} \frac{1}{(1+t)^{pq(1+t)/2}} \frac{\Gamma_p\big(\frac{q(1+t)}{2}\big)}{\Gamma_p\big(\frac{q}{2}\big)}, \\ M_{n,1} &= |\mathbf{\Sigma}|^{qt/2}\Big(\frac{(1+t)^p}{|\mathbf{I}+t\mathbf{\Sigma}|}\Big)^{q(1+t)/2} \end{aligned} \tag{3.13}$$

and $q = n-1$. From Lemma 3.5 and (3.7), we know $n\sigma_n \ge n\sigma_{n,0} \ge p/2$. Then

$$|t| = \frac{|s|}{n\sigma_n} \le \frac{1}{p}$$

for all $n \ge 3$ by (3.10). With assumption $A := \sup_{n\ge 3}\{\|\mathbf{\Sigma}_n\|\} < \infty$, we see $|t| \cdot (2\|\mathbf{\Sigma}\| + 3) \le 1$ as n is sufficiently large. Apply Lemma 3.6 to have

$$\begin{aligned} & \log\Big(\frac{(1+t)^p}{|\mathbf{I}+t\mathbf{\Sigma}|}\Big)^{q(1+t)/2} \\ = \; & -\frac{q}{2}[\mathrm{tr}(\mathbf{\Sigma}) - p]t + \frac{q}{4}\mathrm{tr}[(\mathbf{\Sigma}-\mathbf{I})^2]t^2 + D_n \cdot (nt^3) \cdot \mathrm{tr}[(\mathbf{\Sigma}-\mathbf{I})^2] \end{aligned}$$

as $n \to \infty$ with $|D_n| \le 4A+5$. For $t = \frac{s}{n\sigma_n}$, by the definition of σ_n^2, we have

$$\frac{1}{n\sigma_n^2}\mathrm{tr}[(\mathbf{\Sigma}-\mathbf{I})^2] \le 2 \tag{3.14}$$

for $n \ge 3$. It follows that

$$n|t|^3 \cdot \mathrm{tr}[(\mathbf{\Sigma}-\mathbf{I})^2] \le \frac{2|s|^3}{n\sigma_n} \le \frac{1}{p}$$

by Lemma 3.5. These together with (3.12) and (3.13) imply that

$$\begin{aligned} \log EM_{n,1} \; = \; & \log E[(\Lambda_{n,0}^*)^t] + \\ & \frac{q}{2}\big(p - \mathrm{tr}(\mathbf{\Sigma}) + \log|\mathbf{\Sigma}|\big)t + \frac{q}{4}\,\mathrm{tr}\big[(\mathbf{\Sigma}-\mathbf{I})^2\big]\cdot t^2 + o(1) \end{aligned} \tag{3.15}$$

as $n \to \infty$.

Step 2. We now analyze $\log E[(\Lambda_{n,0}^*)^t]$. By Lemma 3.4, $(\log \Lambda_{n,0} - \mu_{n,0})/(n\sigma_{n,0}) \to N(0,1)$ as $n \to \infty$. Hence, by (3.9),

$$\frac{\log \Lambda_{n,0}^* - \mu_{n,0}}{n\sigma_n} s = \frac{\log \Lambda_{n,0} - \mu_{n,0}}{n\sigma_{n,0}} \cdot \frac{\sigma_{n,0}}{\sigma_n} s \to sy_0 \cdot N(0,1) \tag{3.16}$$

in distribution as $n \to \infty$ for all $s \in \mathbb{R}$. Further, by (3.3), $E\exp\big\{s(\log \Lambda_{n,0}^* - \mu_{n,0})/(n\sigma_{n,0})\big\} \to e^{s^2/2}$ as $n \to \infty$ for all s with $|s| \le \delta_0$, where $\delta_0 > 0$ is a constant. Use the fact $\sigma_{n,0} \le \sigma_n$ for all n and the Hölder inequality to obtain

$$E\exp\Big\{\frac{\log \Lambda_{n,0}^* - \mu_{n,0}}{n\sigma_n} s\Big\} \le \Big[E\exp\Big\{\frac{\log \Lambda_{n,0}^* - \mu_{n,0}}{n\sigma_{n,0}} s\Big\}\Big]^{\frac{\sigma_{n,0}}{\sigma_n}} < \infty$$

for each $|s| \le \delta_0$ and $n \ge 3$, and the last term has limit $e^{y_0 s^2/2}$ as $n \to \infty$. This says that

$$\sup_{n\ge 3} E\Big[\exp\Big\{\frac{\log \Lambda_{n,0}^* - \mu_{n,0}}{n\sigma_n} s\Big\}\Big]^2 < \infty$$

for each $|s| \le \delta_0/2$. Therefore, $\{\exp(\frac{\log \Lambda_{n,0}^* - \mu_{n,0}}{n\sigma_n} s); n \ge 3\}$ is uniformly integrable for each $|s| \le \delta_0/2$. This fact together with (3.16) concludes that

$$\lim_{n\to\infty} \log E\exp\Big\{\frac{\log \Lambda_{n,0}^* - \mu_{n,0}}{n\sigma_n} s\Big\} = \log E e^{sy_0 N(0,1)} = \frac{s^2 y_0^2}{2}.$$

Or equivalently,

$$\log E[(\Lambda_{n,0}^*)^t] - \mu_{n,0} t = \frac{s^2 y_0^2}{2} + o(1) \tag{3.17}$$

as $n \to \infty$ for all $|s| \le \delta_0/2$.

Step 3. Evidently, from (3.7) and (3.14),

$$\begin{aligned} \sigma_{n,0}^2 + \frac{n-1}{2n^2} \cdot \mathrm{tr}\big[(\mathbf{\Sigma} - \mathbf{I})^2\big] &= \sigma_{n,0}^2 + \frac{1}{2n} \cdot \mathrm{tr}\big[(\mathbf{\Sigma} - \mathbf{I})^2\big] - \frac{1}{2n^2} \cdot \mathrm{tr}\big[(\mathbf{\Sigma} - \mathbf{I})^2\big] \\ &= \sigma_n^2 (1 + o(1)). \end{aligned} \tag{3.18}$$

Further, it is easy to read that

$$\mu_n = \mu_{n,0} + \frac{q}{2}\big(p - \mathrm{tr}(\mathbf{\Sigma}) + \log|\mathbf{\Sigma}|\big).$$

This together with (3.15), (3.17) and the notation $q = n - 1$ implies

$$\begin{aligned} \log E[(\Lambda_n^*)^t] &= \Big(\mu_{n,0} t + \frac{s^2 y_0^2}{2}\Big) + \frac{q}{2}\big(p - \mathrm{tr}(\mathbf{\Sigma}) + \log|\mathbf{\Sigma}|\big)t + \frac{q}{4}\,\mathrm{tr}\big[(\mathbf{\Sigma} - \mathbf{I})^2\big] \cdot t^2 + o(1) \\ &= \mu_n t + \frac{s^2 y_0^2}{2} + \frac{1}{4}(n-1) \cdot \mathrm{tr}\big[(\mathbf{\Sigma} - \mathbf{I})^2\big] \cdot t^2 + o(1). \end{aligned}$$

Recalling (3.9), the above is equal to

$$\begin{aligned} & \mu_n t + \frac{1}{2}\frac{\sigma_{n,0}^2}{\sigma_n^2} s^2 + \frac{n-1}{4n^2} \cdot \mathrm{tr}\big[(\mathbf{\Sigma} - \mathbf{I})^2\big] \cdot \frac{s^2}{\sigma_n^2} + o(1) \\ =\ & \mu_n t + \frac{s^2}{2\sigma_n^2} \cdot \Big\{\sigma_{n,0}^2 + \frac{n-1}{2n^2} \cdot \mathrm{tr}\big[(\mathbf{\Sigma} - \mathbf{I})^2\big]\Big\} + o(1) \\ =\ & \mu_n t + \frac{s^2}{2} + o(1) \end{aligned}$$

as $n \to \infty$ for all $|s| \le \frac{1}{4} \wedge \frac{\delta_0}{2}$, where the last identity comes from (3.18). This gives (3.11). ■

3.3 Proof of Theorem 2

The following lemma is needed.

Lemma 3.7 *(Theorems 8.5.3 from Muirhead (1982)) Assume $n > p$. Let Λ_n be as in (1.9). Then*

$$E(\Lambda_n^t) = \Big(\frac{2e}{n}\Big)^{npt/2}\frac{\Gamma_p(\frac{n(1+t)-1}{2})}{\Gamma_p(\frac{n-1}{2})}\frac{|\mathbf{\Sigma}|^{nt/2}}{|\mathbf{I}+t\mathbf{\Sigma}|^{n(1+t)/2}}\cdot\exp\Big\{-\frac{1}{2}nt\mu'\big[\mathbf{I}-t(\mathbf{\Sigma}^{-1}+t\mathbf{I})^{-1}\big]\mu\Big\}$$

for any $t > \frac{p}{n}-1$.

The next result reveals a detail of the exponent appearing in Lemma 3.7.

Lemma 3.8 *Let $\mu \in \mathbb{R}^p$ and $\mathbf{\Sigma}$ be a $p\times p$ positive definite matrix. Write*

$$\mu'\big[\mathbf{I}-t(\mathbf{\Sigma}^{-1}+t\mathbf{I})^{-1}\big]\mu = \|\mu\|^2-(\mu'\mathbf{\Sigma}\mu)\,t+\gamma\cdot t^2.$$

Then, $|\gamma|\le 2\|\mathbf{\Sigma}\|^2\|\mu\|^2$ as $|t|\le(2\|\mathbf{\Sigma}\|)^{-1}$.

Proof. Without loss of generality, we assume $\mu\neq\mathbf{0}$. Let $\mathbf{O}$ be an orthogonal matrix such that $\mathbf{\Sigma}=\mathbf{O}'\mathrm{diag}(\lambda_i;\,1\le i\le p)\mathbf{O}$, where $\lambda_1,\cdots,\lambda_p$ are the eigenvalues of $\mathbf{\Sigma}$. Set $\mathbf{Z}=\mathbf{O}\mu=(z_1,\cdots,z_p)'$. Then

$$\mu'\mathbf{\Sigma}\mu = \mathbf{Z}'\,\mathrm{diag}(\lambda_i;\,1\le i\le p)\,\mathbf{Z}=\sum_{i=1}^p\lambda_i z_i^2; \tag{3.19}$$
$$\mathbf{I}-t(\mathbf{\Sigma}^{-1}+t\mathbf{I})^{-1}=\mathbf{O}'\,\mathrm{diag}\Big(\frac{1}{1+\lambda_i t};\,1\le i\le p\Big)\,\mathbf{O}.$$

Hence, by (3.19)

$$\mu'\big[\mathbf{I}-t(\mathbf{\Sigma}^{-1}+t\mathbf{I})^{-1}\big]\mu=\sum_{i=1}^p\frac{z_i^2}{1+\lambda_i t}.$$

Write $\frac{1}{1+x}=1-x+\eta(x)$. Then $|\eta(x)|=\frac{x^2}{1+x}\le 2x^2$ for $|x|\le\frac{1}{2}$. Now,

$$\begin{aligned}\sum_{i=1}^p\frac{z_i^2}{1+\lambda_i t} &= \sum_{i=1}^p z_i^2 - t\sum_{i=1}^p\lambda_i z_i^2+\sum_{i=1}^p\eta(\lambda_i t)z_i^2\\ &= \|\mu\|^2-t(\mu'\mathbf{\Sigma}\mu)+\epsilon_n(t)\end{aligned}$$

where $\epsilon_n(t):=\sum_{i=1}^p\eta(\lambda_i t)z_i^2$. Notice $\max_{1\le i\le p}|\lambda_i|=\|\mathbf{\Sigma}\|$ and $\sum_{i=1}^p z_i^2=\|\mu\|^2$. We then obtain

$$|\epsilon_n(t)|\le 2t^2\sum_{i=1}^p\lambda_i^2 z_i^2\le 2t^2\|\mathbf{\Sigma}\|^2\sum_{i=1}^p z_i^2=(2\|\mathbf{\Sigma}\|^2\|\mu\|^2)t^2$$

for all t such that $|\lambda_i t|\le\frac{1}{2}$ for all $1\le i\le p$, which is equivalent to that $|t|\le(2\|\mathbf{\Sigma}\|)^{-1}$. ■

We are now ready to prove Theorem 2. The main idea is studying the m.g.f. of the LRT statistic as before. Lemmas 3.6 and 3.8 are new elements relative to the technical parts appeared in Jiang and Yang (2013) and Jiang and Qi (2015).

Proof of Theorem 2. Define

$$\mu_{n,0} = -\frac{1}{4}\Big[n(2n-2p-3)\log(1-\frac{p}{n-1}) + 2(n+1)p\Big],$$
$$\sigma_{n,0}^2 = -\frac{1}{2}\Big(\frac{p}{n-1} + \log(1-\frac{p}{n-1})\Big) > 0,$$

and

$$\mu_{n,1} = \frac{1}{2}n\cdot\big(p - \mathrm{tr}(\mathbf{\Sigma}) + \log|\mathbf{\Sigma}|\big) - \frac{1}{2}n\|\mu\|^2,$$
$$\sigma_{n,1}^2 = \frac{1}{2n}\cdot\mathrm{tr}\big[(\mathbf{\Sigma}-\mathbf{I})^2\big] + \frac{1}{n}(\mu'\mathbf{\Sigma}\mu).$$

Then,

$$\mu_n = \mu_{n,0} + \mu_{n,1} \quad \text{and} \quad \sigma_n^2 = \sigma_{n,0}^2 + \sigma_{n,1}^2 \tag{3.20}$$

for all n. We need to show

$$Q_n := \frac{\log\Lambda_n - \mu_n}{n\sigma_n} \text{ converges to } N(0,1) \tag{3.21}$$

in distribution as $n\to\infty$. By the same argument as in the corresponding part from the proof of Theorem 1, without loss of generality, it suffices to prove (3.21) under the condition that

$$\lim_{n\to\infty}\frac{\sigma_{n,0}}{\sigma_n} = y_0 \in [0,1] \quad \text{and} \quad \lim_{n\to\infty}\frac{p_n}{n} = y \in [0,1]. \tag{3.22}$$

To prove (3.21), similar to the argument as in (3.3), it only needs to show

$$E\exp\Big\{\frac{\log\Lambda_n - \mu_n}{n\sigma_n}s\Big\} \to e^{s^2/2}$$

as $n\to\infty$ for all s such that

$$|s| < \frac{1}{8}, \tag{3.23}$$

or equivalently,

$$\log E[(\Lambda_n)^t] - \mu_n t \to \frac{s^2}{2} \tag{3.24}$$

as $n\to\infty$, where $t = t_n := \frac{s}{n\sigma_n}$. Trivially, $1-\frac{p}{n} = \frac{n-p}{n} \geq \frac{n-p}{2(n-1)} = \frac{1}{2}(1-\frac{p-1}{n-1})$. By (3.20) and Lemma 3.5,

$$n\big(1-\frac{p}{n}\big)\sigma_n \geq \frac{1}{2}\cdot n\big(1-\frac{p-1}{n-1}\big)\sigma_{n,0} \geq \frac{1}{8}$$

for all $n\geq 3$. Thus, for any $|s|\leq\frac{1}{8}$, we see $|t| = \frac{|s|}{n\sigma_n} \leq 1-\frac{p}{n}$. Hence, $t > \frac{p}{n}-1$. Now, fix s satisfying (3.23). From Lemma 3.7 we are able to write

$$E(\Lambda_n^t) = E(\Lambda_{n,0}^t)\cdot A_n\cdot B_n \tag{3.25}$$

where

$$\Lambda_{n,0} = \Lambda_n\big|_{\mu=0,\mathbf{\Sigma}=\mathbf{I}}; \qquad A_n = \frac{|\mathbf{\Sigma}|^{nt/2}}{|\mathbf{I}+t\mathbf{\Sigma}|^{n(1+t)/2}}(1+t)^{np(1+t)/2};$$
$$B_n = \exp\Big\{-\frac{1}{2}nt\mu'\big[\mathbf{I} - t(\mathbf{\Sigma}^{-1}+t\mathbf{I})^{-1}\big]\mu\Big\}.$$

We next study the three terms one by one.

Step 1. The analysis of $E(\Lambda_{n,0}^t)$ in (3.25). This is the t-th moment of Λ_n under the null hypothesis, which is investigated by Jiang and Qi (2015). By Theorem 5 from the paper, $(\log \Lambda_{n,0} - \mu_{n,0})/(n\sigma_{n,0}) \to N(0,1)$ as $n \to \infty$. Reviewing the proof of Theorem 5 from Jiang and Qi (2015), by the same argument as obtaining (3.17), we get

$$\log E(\Lambda_{n,0}^t) = \mu_{n,0}t + \frac{s^2 y_0^2}{2} + o(1) \tag{3.26}$$

as $n \to \infty$.

Step 2. The analysis of A_n in (3.25). By the first conclusion of Lemma 3.5, $|t| \leq \frac{1}{n\sigma_{n,0}} \to 0$ as $n \to \infty$. With assumption $A := \sup_{n\geq 3}\{\|\mathbf{\Sigma}_n\|\} < \infty$, we see $|t| \cdot (2\|\mathbf{\Sigma}\| + 3) \leq 1$ as n is sufficiently large. From Lemma 3.6 we get

$$\log A_n = \frac{n}{2}\big(p - \operatorname{tr}(\mathbf{\Sigma}) + \log|\mathbf{\Sigma}|\big)t + \frac{n}{4}\operatorname{tr}\big[(\mathbf{\Sigma}-\mathbf{I})^2\big]\cdot t^2 + D_n \cdot (nt^3)\cdot \operatorname{tr}[(\mathbf{\Sigma}-\mathbf{I})^2]$$

as $n \to \infty$ with $|D_n| \leq 4A+5$. For $t = \frac{s}{n\sigma_n}$, by the definition of σ_n^2, we have $\frac{1}{n\sigma_n^2}\operatorname{tr}[(\mathbf{\Sigma}-\mathbf{I})^2] \leq 2$ for $n \geq 3$. It follows that

$$n|t|^3 \cdot \operatorname{tr}[(\mathbf{\Sigma}-\mathbf{I})^2] \leq \frac{2|s|^3}{n\sigma_n} \leq \frac{1}{2n\sigma_{n,0}} \leq \frac{1}{p}$$

by the first conclusion of Lemma 3.5. Therefore,

$$\log A_n = \frac{n}{2}\big(p - \operatorname{tr}(\mathbf{\Sigma}) + \log|\mathbf{\Sigma}|\big)t + \frac{n}{4}\operatorname{tr}\big[(\mathbf{\Sigma}-\mathbf{I})^2\big]\cdot t^2 + o(1) \tag{3.27}$$

as $n \to \infty$.

Step 3. The analysis of B_n in (3.25). Denote $\tau_n = -\log B_n = \frac{1}{2}nt\mu'\big[\mathbf{I} - t(\mathbf{\Sigma}^{-1} + t\mathbf{I})^{-1}\big]\mu$. By Lemma 3.8,

$$\mu'\big[\mathbf{I} - t(\mathbf{\Sigma}^{-1} + t\mathbf{I})^{-1}\big]\mu = \|\mu\|^2 - (\mu'\mathbf{\Sigma}\mu)\,t + \gamma\cdot t^2$$

with $|\gamma| \leq 2C^2\|\mu\|^2$ for all $|t| \leq (2C)^{-1}$. Therefore,

$$\tau_n = \frac{n\|\mu\|^2}{2}t - \frac{n(\mu'\mathbf{\Sigma}\mu)}{2}t^2 + \epsilon_n(t) \tag{3.28}$$

where $\epsilon_n(t) = \frac{1}{2}n\gamma t^3$. Easily,

$$|\epsilon_n(t)| \leq nC^2\|\mu\|^2 t^3 \leq C^2 \cdot \frac{\|\mu\|^2}{n^2\sigma_n^3} \tag{3.29}$$

as n is sufficiently large since $t = \frac{s}{n\sigma_n}$. Now we estimate this term in two cases.

Case (a): $\|\mu\| = o\big((p^3/n)^{1/2}\big)$. In this situation, we have from (3.29),

$$|\epsilon_n(t)| \leq C^2 \cdot o\Big(\frac{p^3}{n^3\sigma_n^3}\Big) = o(1)$$

since $n\sigma_n \geq p/2$ by the first conclusion of Lemma 3.5.

Case (b): $\|\mathbf{\Sigma}_n^{-1}\| = o(p)$. Notice that

$$\mu'\mathbf{\Sigma}\mu \geq \lambda_{\min}\|\mu\|^2 = \frac{\|\mu\|^2}{\|\mathbf{\Sigma}^{-1}\|},$$

where $\lambda_{\min}$ is the smallest eigenvalue of $\mathbf{\Sigma}$. From (3.29)

$$|\epsilon_n(t)| \leq \frac{C^2}{n\sigma_n} \cdot \frac{\|\mu\|^2}{n\sigma_n^2} \leq \frac{2C^2}{p} \cdot \frac{\|\mu\|^2}{\mu'\mathbf{\Sigma}\mu} \leq \frac{2C^2\|\mathbf{\Sigma}^{-1}\|}{p}$$

which is of the order $o(1)$, where we use the fact $n\sigma_n^2 \geq \mu'\mathbf{\Sigma}\mu$ in the second inequality by the definition of σ_n. Review (3.28). In both cases, we get

$$\log B_n = -\tau_n = -\frac{n\|\mu\|^2}{2}t + \frac{n(\mu'\mathbf{\Sigma}\mu)}{2}t^2 + o(1)$$

as $n \to \infty$. This combing with (3.25), (3.26) and (3.27) concludes

$$\begin{aligned}
\log E[(\Lambda_n)^t] &= \Big(\mu_{n,0}t + \frac{s^2y_0^2}{2}\Big) + \Big(\frac{n}{2}\big(p - \operatorname{tr}(\mathbf{\Sigma}) + \log|\mathbf{\Sigma}|\big)t + \frac{n}{4}\operatorname{tr}\big[(\mathbf{\Sigma}-\mathbf{I})^2\big]\cdot t^2\Big) \\
&\quad + \Big(-\frac{n\|\mu\|^2}{2}t + \frac{n(\mu'\mathbf{\Sigma}\mu)}{2}t^2\Big) + o(1) \\
&= \mu_n t + \frac{s^2}{2}\Big(\frac{\sigma_{n,0}^2}{\sigma_n^2} + \frac{1}{\sigma_n^2}\cdot\frac{1}{2n}\operatorname{tr}\big[(\mathbf{\Sigma}-\mathbf{I})^2\big] + \frac{1}{\sigma_n^2}\cdot\frac{1}{n}(\mu'\mathbf{\Sigma}\mu)\Big) + o(1) \\
&= \mu_n t + \frac{s^2}{2}\frac{\sigma_{n,0}^2 + \sigma_{n,1}^2}{\sigma_n^2} + o(1) = \mu_n t + \frac{s^2}{2} + o(1)
\end{aligned}$$

as $n \to \infty$, where the second equality comes from (3.20), (3.22) and the notation $t = \frac{s}{n\sigma_n}$; the third identity follows from the second assertion of (3.20). This leads to (3.24). ■

3.4 Verifications of (1.7), (1.8), (1.11) and (1.12)

In this section we will verify these four statements.

Verifications of (1.7) and (1.8). Recall

$$\begin{aligned}
\mu_{n,0} &= -\frac{1}{4}(n-1)\Big[2p + \big(2n-2p-3\big)\log\big(1-\frac{p}{n-1}\big)\Big]; \\
\sigma_{n,0}^2 &= -\frac{1}{2}\Big[\frac{p}{n-1} + \log\Big(1-\frac{p}{n-1}\Big)\Big].
\end{aligned}$$

Then

$$\begin{aligned}
\frac{c_\alpha - \mu_n}{n\sigma_n} &= \frac{\mu_{n,0} + n\sigma_{n,0}\Phi^{-1}(\alpha) - \mu_n}{n\sigma_n} \\
&= \frac{\sigma_{n,0}}{\sigma_n}\Phi^{-1}(\alpha) + \frac{n-1}{2n}\cdot\frac{\operatorname{tr}(\mathbf{\Sigma}_n) - \log|\mathbf{\Sigma}_n| - p}{\sigma_n}. \qquad (3.30)
\end{aligned}$$

Notice that

$$\sup_{n\geq 3}\Big|\frac{\sigma_{n,0}}{\sigma_n}\Phi^{-1}(\alpha)\Big| \leq |\Phi^{-1}(\alpha)| \qquad (3.31)$$

for any non-singular matrix $\mathbf{\Sigma}_n$. So (1.7) follows. Further, let $0 < C_1 < C_2 < 1$ be two constants such that $C_1 \leq \frac{p}{n} \leq C_2$ for all $n \geq 3$. Take $\mathbf{\Sigma}_n = a\mathbf{I}_p$ for any $a > 0$. Trivially

$$\operatorname{tr}(\mathbf{\Sigma}_n) - \log|\mathbf{\Sigma}_n| - p = (a - \log a - 1)p; \qquad (3.32)$$

$$\sigma_n^2 = -\frac{1}{2}\Big[\frac{p}{n-1} + \log\Big(1-\frac{p}{n-1}\Big)\Big] + \frac{p}{2n}\cdot(a-1)^2. \qquad (3.33)$$

Then

$$0 < \liminf_{n\to\infty} \sigma_n \le \limsup_{n\to\infty} \sigma_n < \infty \tag{3.34}$$

uniformly for all $a \in (0, C_3]$, where $C_3 > 0$ is a constant not depending on n. Let $a_n = 1 + \frac{c_n}{\sqrt{p}} > 0$, where $-\sqrt{p} < c_n \le A\sqrt{p}$ for all $n \ge 3$ and some constant $A > 0$ not depending on n. Let us prove (1.8).

Assume $\lim_{n\to\infty} c_n = 0$. Then $\lim_{n\to\infty} a_n = 1$. Consequently,

$$\lim_{n\to\infty} \sigma_n = \sigma_{n,0} \quad \text{and} \quad \mathrm{tr}(\mathbf{\Sigma}_n) - \log|\mathbf{\Sigma}_n| - p \sim \frac{(a_n-1)^2 p}{2} \to 0 \tag{3.35}$$

from (3.32) and (3.33). We also use the fact $x - \log x - 1 \sim \frac{1}{2}(x-1)^2$ as $x \to 1$. This, (3.30) and (3.34) conclude that $\lim_{n\to\infty} \frac{c_\alpha - \mu_n}{n\sigma_n} = \Phi^{-1}(\alpha)$. We get the second assertion in (1.8) from (1.6).

By using the fact that $x - 1 - \log x > 0$ for all $0 < x \ne 1$ and that $x - 1 - \log x \sim \frac{1}{2}(x-1)^2$ as $x \to 1$, we have

$$\inf_{0<x\le A+1, x\ne 1} \frac{x - 1 - \log x}{(x-1)^2} > 0.$$

Therefore, from (3.32),

$$\mathrm{tr}(\mathbf{\Sigma}_n) - \log|\mathbf{\Sigma}_n| - p = \frac{a_n - \log a_n - 1}{(a_n-1)^2} \cdot c_n^2 \to \infty \tag{3.36}$$

provided $\lim_{n\to\infty} |c_n| = \infty$. This together with (3.30), (3.31) and (3.34) yields that $\frac{c_\alpha - \mu_n}{n\sigma_n} \to \infty$. The first assertion in (1.8) then follows from (1.6). ■

Verifications of (1.11) and (1.12). Review

$$\mu_{n,0} = -\frac{1}{4}\Big[n(2n-2p-3)\log(1-\frac{p}{n-1}) + 2(n+1)p\Big];$$
$$\sigma_{n,0}^2 = -\frac{1}{2}\Big(\frac{p}{n-1} + \log(1-\frac{p}{n-1})\Big).$$

Then

$$\sup_{n\ge 3}\Big|\frac{\sigma_{n,0}}{\sigma_n}\Phi^{-1}(\alpha)\Big| \le |\Phi^{-1}(\alpha)| \tag{3.37}$$

for any non-singular matrix $\mathbf{\Sigma}_n$ and

$$\begin{aligned}\frac{c_\alpha - \mu_n}{n\sigma_n} &= \frac{\mu_{n,0} + n\sigma_{n,0}\Phi^{-1}(\alpha) - \mu_n}{n\sigma_n} \\ &= \frac{\sigma_{n,0}}{\sigma_n}\Phi^{-1}(\alpha) + \frac{1}{2\sigma_n}\cdot\big[\mathrm{tr}(\mathbf{\Sigma}_n) - \log|\mathbf{\Sigma}_n| - p + \|\mu\|^2\big].\end{aligned} \tag{3.38}$$

Consequently, the assertion (1.11) follows from (1.10) and (3.37). Next we show (1.12).

Suppose $0 < C_1 < C_2 < 1$ be two constants such that $C_1 \le \frac{p}{n} \le C_2$ for all $n \ge 3$. Take $\mathbf{\Sigma}_n = a\mathbf{I}_p$ for any $a > 0$. Trivially

$$\mathrm{tr}(\mathbf{\Sigma}_n) - \log|\mathbf{\Sigma}_n| - p + \|\mu\|^2 = (a - \log a - 1)p + \|\mu\|^2; \tag{3.39}$$

$$\sigma_n^2 = -\frac{1}{2}\Big[\frac{p}{n-1} + \log\Big(1 - \frac{p}{n-1}\Big)\Big] + \frac{p}{2n}\cdot(a-1)^2 + \frac{a}{n}\|\mu\|^2. \tag{3.40}$$

Evidently,

$$0 < \liminf_{n\to\infty} \sigma_n \le \limsup_{n\to\infty} \sigma_n < \infty \tag{3.41}$$

uniformly for all a and μ such that $a \in (0, C_3]$ and $\sup_{n\ge 3} \frac{1}{n}\|\mu\|^2 < \infty$, where $C_3 > 0$ is a constant not depending on n. By (3.39),

$$\mathrm{tr}(\boldsymbol{\Sigma}_n) - \log|\boldsymbol{\Sigma}_n| - p + \|\mu\|^2 \ge (a - \log a - 1)p. \tag{3.42}$$

Let $a_n = 1 + \frac{c_n}{\sqrt{p}} > 0$, where $-\sqrt{p} < c_n \le A\sqrt{p}$ for all $n \ge 3$ and some constant $A > 0$ not depending on n. We discuss the two cases in (1.12) next.

Assume $\lim_{n\to\infty} c_n = 0$ and $\|\mu\| = o(1)$. Then $\lim_{n\to\infty} a_n = 1$ and

$$\lim_{n\to\infty} \sigma_n = \sigma_{n,0} \quad \text{and} \quad \mathrm{tr}(\boldsymbol{\Sigma}_n) - \log|\boldsymbol{\Sigma}_n| - p + \|\mu\|^2 \to 0$$

from (3.39), (3.40) and the second assertion of (3.35). By (3.38) and (3.41), this implies $\frac{c_\alpha - \mu_n}{n\sigma_n} \to \Phi^{-1}(\alpha)$. Hence, the second limit in (1.12) is yielded. To obtain the first limit in (1.12), since $-\sqrt{p} < c_n \le A\sqrt{p}$, we know $\sup_{n\ge 3} a_n < \infty$. Further, $\sup_{n\ge 3} \frac{1}{n}\|\mu\|^2 < \infty$ if $\|\mu\| = O(\sqrt{p})$. To show the first conclusion in (1.12), by (1.11), (3.41) and (3.42), it is enough to check $(a_n - \log a_n - 1)p \to \infty$. This is assured by (3.36) under the condition $\lim_{n\to\infty} |c_n| = \infty$. ■

Acknowledgement. The participation of Lin Zhang in the early stage of this project is acknowledged.

References

[1] Anderson, T. (2003). An Introduction to Multivariate Statistical Analysis. *Wiley-Interscience*, 3rd Ed.

[2] Billingsley, P. (1986). Probability and Measure. *Wiley Series in Probability and Mathematical Statistics*, 2nd Ed.

[3] Bai, Z., Jiang, D., Yao, J. and Zheng, S. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *Ann. Stat.* 37, 3822-3840.

[4] Cai, T. and Ma, Z. (2013). Optimal hypothesis testing for high dimensional covariance matrices. *Bernoulli* 19(5B), 2359-2388.

[5] Chen, S., Zhang, L., and Zhong, P. (2010). Tests for high dimensional covariance matrices. *J. Amer. Stat. Assoc.* 105, 810-819.

[6] Eaton, M. (1983). Multivariate Statistics: A Vector Space Approach (Wiley Series in Probability and Statistics). *John Wiley & Sons Inc.*

[7] Jiang, D., Jiang, T. and Yang, F. (2012). Likelihood ratio tests for covariance matrices of high-dimensional normal distributions. *J. Stat. Plann. Inference* 142, 2241-2256.

[8] Jiang, T. and Qi, Y. (2015). Likelihood ratio tests for high-dimensional normal distributions. *Scand. J. of Stat. 42(4), 988-1009.*

[9] Jiang, T. and Yang, F. (2013). Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions. *Ann. Stat. 41(4), 2029-2074.*

[10] Ledoit, O. and Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Ann. Stat.* 30, 1081-1102.

[11] Muirhead, R. J. (1982). Aspects of Multivariate Statistical Theory. *Wiley, New York.*

[12] Onatski, A., Moreira, M. and Hallin, M. (2014). Signal detection in high dimension: The multispiked case. *Ann. Statist.* 42(1), 225-254.

[13] Schott, J. R. (2007). A test for the equality of covariance matrices when the dimension is large relative to the sample sizes. *Comput. Statist. Data Anal.* 51, 6535-6542.

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

[15] Schott, J. R. (2001). Some tests for the equality of covariance matrices. *J. Stat. Plann. Inference* 94, 25-36.

[16] Srivastava, M. S. (2005). Some tests concerning the covariance matrix in high-dimensional data. Journal of the Japan Statistical Society 35, 251-272.

[17] Zheng, S., Bai, Z. and Yao, J. (2015). Substitution principle for CLT of linear spectral statistics of high-dimensional sample covariance matrices with applications to hypothesis testing. *Ann. Stat.* 43, 546-591.