Songklanakarin J. Sci. Technol.
38 (5), 521-535, Sep. - Oct. 2016

http://www.sjst.psu.ac.th

Original Article

A covariance matrix test for high-dimensional data

Saowapha Chaipitak and Samruam Chongcharoen*

School of Applied Statistics, National Institute of Development Administration (NIDA), Bang Kapi, Bangkok, 10240 Thailand.

Received: 25 October 2015; Accepted: 25 January 2016

Abstract

For the multivariate normally distributed data with the dimension larger than or equal to the number of observations, or the sample size, called high-dimensional normal data, we proposed a test for testing the null hypothesis that the covariance matrix of a normal population is proportional to a given matrix on some conditions when the dimension goes to infinity. We showed that this test statistic is consistent. The asymptotic null and non-null distribution of the test statistic is also given. The performance of the proposed test is evaluated via simulation study and its application.

Keywords: asymptotic distribution, high-dimensional data, null distribution, non-null distribution, multivariate normal, hypothesis testing

1. Introduction

Let $X_1,...,X_N$ be a set of independent observations from a multivariate normal distribution $N_p(\mu,\Sigma)$ where both the mean vector μ and covariance matrix Σ, Σ is a positive definite matrix, are unknown. In this paper, we are interested in the problem of testing the hypothesis that the covariance matrix of a normal population is proportional to a given matrix, that is, $H_0:\Sigma = t\Sigma_0$ against $H_1:\Sigma \neq t\Sigma_0$ where both $0<t<\infty$, Σ_0 are known. The likelihood ratio test (LRT),which is based on the sample covariance matrix, is the traditional technique to handle this hypothesis and requires $n \geq p$. But many applications in modern science and economics, e.g. the analysis of DNA microarrays, the dimension is usually in thousands of gene expressions whereas the sample size is small, which makes $n < p$, called high-dimensional data. For such data, the LRT is not applicable because the sample covariance matrix, *S*, is singular when $n < p$ (see, for examples, Muirhead, 1982, Sections 8.3 and 8.4; Anderson, 1984, Sections 10.7 and 10.8).

Recently, several authors have proposed methods for testing the related problems. Some of them are: John (1971); Nagao (1973); Ledoit and Wolf (2002); and Srivastava (2005), and Fisher *et al.* (2010). Those are given as follows.

John (1971) proposed a test statistic for testing that the covariance matrix of a normal population is proportional to an identity matrix, that is, $H_0':\Sigma = tI$, $0<t<\infty$ a known value which is the locally most powerful invariant test as

$$U = \frac{1}{p} tr\left[\left(\frac{S}{(1/p)tr(S)} - I\right)^2\right]$$

and Nagao (1973) proposed a test statistic for testing H_0'' : $\Sigma = I$ as

$$V = \frac{1}{p} tr\left[\left(S - I\right)^2\right]$$

Both U and V test statistics are consistent and have been studied under assuming that n goes to infinity while p remains fixed. So, Ledoit and Wolf (2002) demonstrated that the test statistic for testing H_0' based on U statistic is still consistent if n goes to infinity with p that is as $(n,p)\to\infty$ and $p/n \to c, c\in(0,\infty)$. The null hypothesis H_0' is rejected if

* Corresponding author.
Email address: samruam@as.nida.ac.th

$$U_j = \frac{npU}{2} \tag{1.1}$$

exceeds the appropriate quantile from the χ^2- distribution with $p(p+1)/2-1$ degree of freedom. For testing H''_0: $\Sigma = I$ if goes to infinity with p as $n < p$, Ledoit and Wolf (2002) showed that the statistic V is not consistent against every alternative and its n–limiting distribution differs from its $(n, p)-$ limiting distribution under the null hypothesis. Then they modified the statistic V as

$$W = \frac{1}{p}tr\left[(S-I)^2\right] - \frac{p}{n}\left(\frac{1}{p}trS\right)^2 + \frac{p}{n}.$$

They have shown that the statistic W is consistent as $(n, p) \to \infty$, including the case $n<p$. The test statistic based on W rejects the null hypothesis H''_0 if $npW/2$ exceeds the appropriate quantile from the χ^2-distribution with $p(p+1)/2$ degrees of freedom. Srivastava (2005) proposed a test statistic when $(n, p) \to \infty$, $n = O(p^\delta)$ $0 < \delta \le 1$, to reject the null hypothesis $H_0 : \Sigma = \sigma^2 I$, $\sigma^2 > 0$ unknown, with a test statistic which did not relate with unknown $\sigma^2 > 0$. Then we applied his test statistic for testing H'_0 and reject H'_0 if

$$T_{S1} = \frac{n}{2}(\hat{h}_2/\hat{h}_1^2 - 1) \tag{1.2}$$

exceeds the appropriate quantile of the standard normal distribution, where $\hat{h}_1 = (1/p)tr(S)$ and $\hat{h}_2 = \frac{n^2}{(n-1)(n+2)}\frac{1}{p}\left[trS^2 - \frac{1}{n}(trS)^2\right]$. The statistics $\hat{h}_1$ and $\hat{h}_2$ are $(n, p)-$consistent estimators of $(1/p)tr\Sigma$ and $(1/p)tr\Sigma^2$ respectively. Also he proposed a test to reject the null hypothesis H''_0 if

$$T_{S2} = \frac{n}{2}(\hat{h}_1 - 2\hat{h}_2 + 1) \tag{1.3}$$

exceeds the appropriate quantile of the standard normal distribution. Motivated by the result in Srivastava (2005), which requires, $p/n \to c$, $c \in (0,+\infty)$ Fisher *et al.* (2010) proposed the test for testing H'_0 based on unbiased and consistent estimators of the second and fourth arithmetic means of the sample eigenvalues. With the constants:

$$b = -\frac{4}{n},\ c^* = -\frac{2n^2+3n-6}{n(n^2+n+2)},\ d = \frac{2(5n+6)}{n(n^2+n+2)},$$

$$e = -\frac{5n+6}{n^2(n^2+n+2)}$$

and $$\tau = \frac{n^5(n^2+n+2)}{(n+1)(n+2)(n+4)(n+6)(n-1)(n-2)(n-3)},$$

they proposed the test statistic to reject the null hypothesis H'_0 if

$$T_F = \frac{n(\hat{h}_4/\hat{h}_2^2 - 1)}{\sqrt{8(8+12c+c^2)}}$$

exceeds the appropriate quantile of the standard normal distribution, where

$$\hat{h}_4 = \frac{\tau}{p}\left[trS^4 + btrS^3trS + c^*(trS^2)^2 + dtrS^2(trS)^2 + e(trS)^4\right]$$

is $(n, p)-$consistent estimator of $(1/p)tr\Sigma^4$.

The remainder of this paper is organized as follows. Section 2 provides the proposed test statistic and its asymptotic distribution under both the null and alternative hypotheses as (n, p) go to infinity even if $n < p$. Section 3 shows the performance of the proposed test statistic through simulation technique. Section 4 applies the test statistic to real data. Section 5 contains the conclusions. The theoretical derivations are given in the Appendix.

2. Description of the Proposed Test

Suppose $X_1,...,X_{n+1} \sim N_p(\mu,\Sigma)$ and we are interested in testing that the covariance matrix of a normal population is proportional to a given matrix, that is, H_0 : $\Sigma = t\Sigma_0$ against $H_1 : \Sigma \neq t\Sigma_0$ where $0 < t < \infty$ is known value and Σ_0 is a given known positive definite matrix. We proposed the test statistic by considering a measure of a distance between the two matrices

$$\psi = \frac{1}{p}(tr(\Sigma_0^{-1}\Sigma - tI)^2 = \frac{1}{p}tr(\Sigma_0^{-1}\Sigma)^2 - \frac{2t}{p}tr(\Sigma_0^{-1}\Sigma) + t^2 \tag{2.1}$$

where tr denotes the trace of matrix and if and $\psi = 0$ only if the null hypothesis holds. Thus, we may consider testing $H_0 : \psi = 0$ against $H_1 : \psi > 0$.

We shall make the following assumptions:

(A) $\lim_{p\to\infty} a_i = a_i^0$, $a_i^0 \in (0,\infty)$, $i = 1,...,8$

(B) $\lim_{(n,p)\to\infty} p/n = c$, $c \in (0,\infty)$

where $a_i = (1/p)tr(\Sigma_0^{-1}\Sigma)^i = (1/p)\sum_{j=1}^{p}(\lambda_j/d_j)^i$. The λ_j's are the eigenvalues of the covariance matrix Σ and d_j's are the eigenvalues of a known positive definite matrix Σ_0. We need estimators of a_1 and a_2 to be consistent estimators for large p and n even if $n < p$. The following theorem provides these consistent estimators.

Theorem 2.1 The unbiased and consistent estimators of $a_1 = (1/p)tr(\Sigma_0^{-1}\Sigma)$ and $a_2 = (1/p)tr(\Sigma_0^{-1}\Sigma)^2$ are respectively given by

$$\hat{a}_1 = (1/p)tr(\Sigma_0^{-1}S) \quad (2.2)$$

and

$$\hat{a}_2 = \frac{n^2}{(n-1)(n+2)}\frac{1}{p}\left[tr(\Sigma_0^{-1}S)^2 - \frac{1}{n}(tr(\Sigma_0^{-1}S))^2\right] \quad (2.3)$$

Thus we use estimators in Theorem 2.1 to define the unbiased and consistent estimator of ψ in (2.1) as

$$\hat{\psi} = \hat{a}_2 - 2t\hat{a}_1 + t^2 \quad (2.4)$$

The following theorem gives the asymptotic distribution of the estimators $\hat{a}_1$ and $\hat{a}_2$ in (2.4).

Theorem 2.2 Under the assumption (A), and (B), as $(n, p) \to \infty$

$$\begin{pmatrix}\hat{a}_1\\ \hat{a}_2\end{pmatrix} \xrightarrow{D} N_2\left[\begin{pmatrix}a_1\\ a_2\end{pmatrix}, \begin{pmatrix}2a_2/np & 4a_3/np\\ 4a_3/np & 4(2a_4 + ca_2^2)/np\end{pmatrix}\right]$$

where $x \xrightarrow{D} y$ denotes x converges in distribution to y.

The following theorem and corollary provide the asymptotic distribution of $\hat{\psi}$ under the alternative and null hypothesis by applying the delta method of a function of two random variables.

Theorem 2.3 Under the assumption (A), and (B), as $(n, p) \to \infty$

$$\hat{\psi} - \psi \xrightarrow{D} N\left(0, \delta^2\right) \quad (2.5)$$

with $\delta^2 = \frac{4}{np}(2ta_2 - 4ta_3 + 2a_4 + ca_2^2)$.

Corollary 2.1 Under the null hypothesis $H_0 : \Sigma = t\Sigma_0$ then $\psi = 0$ and under the assumption (A), and (B), as $(n, p) \to \infty$

$$T = \frac{\sqrt{np\hat{\psi}}}{2\sqrt{ct^4}} = \frac{n\hat{\psi}}{2t^2} \xrightarrow{D} N(0,1) \quad (2.6)$$

Remark If $t = 1$ and $\Sigma_0 = I$ where I is identity matrix, then the proposed statistic T is the test statistic T_{S2} in (1.3) given by Srivastava (2005).

3. Simulation Study

For studying the performance of the proposed test statistic T, we compute the attained significance level (ASL) of the proposed test by simulation technique. Based on 10,000 replications of the data set simulated under the null hypothesis $H_0 : \Sigma = t\Sigma_0$, test statistic T is computed and then we obtain the attained significance level (ASL) of the test by recording the proportion rejection of test statistic for the null hypothesis with the nominal significance level at 0.05. We simulate the ASL for different four null hypotheses as

1) $H_0^1 : \Sigma = t\Sigma_0 = C_{01}$ where $C_{01} = (c_{i,j})_{p\times p} = (c_{i-j})_{p\times p}$ is a Toeplitz matrix with elements $c_0 = 1,\ c_{-1} = c_1 = -0.5$ and the rest elements are equal to zero

2) $H_0^2 : \Sigma = t\Sigma_0 = C_{02} = 0.5I_p + 0.51_p1_p'$, where I_p denotes the $p\times p$ identity matrix, and 1_p denotes the $p\times 1$ vector having each element equal to 1

3) $H_0^3 : \Sigma = t\Sigma_0 = C_{03}$ where $C_{03} = (c_{i,j})_{p\times p} = (c_{j,i})_{p\times p}$ with $c_{i,j} = (-1)^{i+j}(i/2j)\ \forall i < j = 1,..,p$ and $c_{i,i} = 1.0\ \forall i = 1,...,p$

4) $H_0^4 : \Sigma = t\Sigma_0 = C_{04}$ where $C_{04} = (c_{i,j})_{p\times p} = (c_{j,i})_{p\times p}$ with $c_{i,j} = (-1)^{i+j}0.9^{|i-j|^{1/5}}\ \forall i, j = 1,..,p$

For each null hypothesis, we simulate the empirical power of the proposed test T under the alternative hypothesis for each of four null hypotheses as

1) $H_0^1 : \Sigma = C_{01}$ against $H_1^1 : \Sigma = C_1$ where $C_1 = (c_{i,j})_{p\times p} = (c_{i-j})_{p\times p}$ is a Toeplitz matrix with elements $c_0 = 1, c_{-1} = c_1 = -0.49$ and the rest elements are equal to zero

2) $H_0^2 : \Sigma = C_{02}$ against $H_1^2 : \Sigma = C_2 = 0.9I_p + (0.1)1_p1_p'$

3) $H_0^3 : \Sigma = C_{03}$ against $H_1^3 : \Sigma = C_3$ where $C_3 = (c_{i,j})_{p\times p} = (c_{j,i})_{p\times p}$ with $c_{i,j} = (-1)^{i+j}(i/4j)\ \forall i < j = 1,..,p$ and $c_{i,i} = 1.0\ \forall i = 1,...,p$

4) $H_0^4 : \Sigma = C_{04}$ against $H_1^4 : \Sigma = C_4$ where $C_4 = (c_{i,j})_{p\times p} = (c_{j,i})_{p\times p}$ with $c_{i,j} = (-1)^{i+j}0.9^{|i-j|^{2/5}}$ $\forall i, j = 1,..,p$

3.1 Simulation results

The ASL is provided in Table 1 corresponding to the null hypotheses. As expected, the ASL of the test statistic T

Table 1. The ASL of test statistic T under four null hypotheses at Nominal Significance Level $\alpha = 0.05$.

p	$n = N-1$	The ASL of T			
		$H_0^1 : \Sigma = C_{01}$	$H_0^2 : \Sigma = C_{02}$	$H_0^3 : \Sigma = C_{03}$	$H_0^4 : \Sigma = C_{04}$
10	9	0.059	0.058	0.059	0.059
40	9	0.055	0.055	0.055	0.055
	39	0.056	0.056	0.056	0.057
80	9	0.057	0.056	0.057	0.057
	39	0.052	0.052	0.052	0.052
	79	0.053	0.052	0.052	0.052
160	9	0.053	0.053	0.054	0.054
	39	0.056	0.055	0.055	0.056
	79	0.056	0.056	0.056	0.055
	159	0.053	0.053	0.053	0.053
320	9	0.052	0.052	0.052	0.052
	39	0.052	0.051	0.052	0.052
	79	0.051	0.051	0.050	0.050
	159	0.051	0.050	0.051	0.051
	319	0.053	0.051	0.053	0.053

is reasonably close to the nominal significant level 0.05 and gets better when p and n get large. We found that four sets of the ASL are almost the same that means the consistency of our test statistic is not affected by varying the null covariance matrices.

The empirical powers are shown in Table 2. It shows that four sets of the empirical power of test statistic T rapidly converge to one and stay high as n and p get large for $n < p$.

We also compute the ASL in a special case of the null covariance matrix with setting $t = 2$ and $\Sigma_0 = I$, that is, the test with the null hypothesis as $H_0' : \Sigma = 2I$ (spherecity). We compare the performance of the proposed test statistic T with the test statistics defined in Ledoit and Wolf (2002), denoted U_j in (1.1) and Srivastava (2005), denoted T_{S1} in (1.2). We compare them under the alternative hypothesis $H_1' : \Sigma = 2D$ where $D = diag(d_1,...,d_p)$; $d_i \sim Unif(0,1)$, $i = 1,2,...,p$. The ASL and the empirical powers are provided in Table 3. Table 3 reports that the ASL of the proposed test statistic T is similar to those provided in Table 1 and closed to those from the test statistic T_{S1} and U_j. But the test statistic T gives the best performance for all of the setting (n,p) and has substantially higher powers than those of U_j and T_{S1} for almost every n and p considered. These results suggest that the proposed test may more appropriate to use than U_j test and T_{S1} test, especially when is small.

4. A Real Example

In this section, the microarray dataset is collected from Notterman *et al.* (2001) is available at http://genomics-pubs.princeton.edu/oncology/Data/CarcinomaNormal datasetCancerResearch.xls (last accessed: 9 October 2011). There are 18 colon adenocarcinomas and their paired normal colon tissues and they are obtained on oligonucleotide arrays. The expression levels of 6500 human genes are measured on each. For simplicity, we will restrict attention to 18 colon adenocarcinomas with only first 256 measurements each. We examine whether the covariance matrix is the sphericity. The data gives the observed test statistic values as $T = 8.500$, $U_j = 284.567$ and $T_{S1} = 270.582$ with *p-value* ≈ 0 each, thus the hypothesis of being sphericity is rejected at any reasonable significance level.

5. Conclusions

For testing the covariance matrix in high-dimensional data, our test statistic is proposed under normality assumption. The test statistic is approximated by normal distribution. Numerical simulations indicate that our test statistic T in (2.6) constructed from the consistent estimators with accurately control size of test and their powers get better when (n,p) get large for $n < p$. Moreover, the test statistic gives higher power than, for testing being sphericity of the covariance matrix, those of the tests in Ledoit and Wolf (2002) and Srivastava (2005).

Acknowledgements

The authors would like to express their gratitude to the Commission on Higher Education (CHE) of Thailand for their financial support.

Table 2. The empirical power of T under four alternative hypotheses.

p	$n = N-1$	The empirical power of T			
		$H_1^1 = \Sigma = C_1$	$H_1^2 : \Sigma = C_2$	$H_1^3 : \Sigma = C_3$	$H_1^4 : \Sigma = C_4$
10	9	0.480	0.560	0.174	0.159
40	9	0.996	0.617	0.265	0.286
	39	1.000	1.000	0.772	0.877
80	9	1.000	0.624	0.300	0.330
	39	1.000	1.000	0.837	0.939
	79	1.000	1.000	0.998	1.000
160	9	1.000	0.625	0.319	0.346
	39	1.000	1.000	0.866	0.966
	79	1.000	1.000	0.999	1.000
	159	1.000	1.000	1.000	1.000
320	9	1.000	0.629	0.342	0.361
	39	1.000	1.000	0.891	0.977
	79	1.000	1.000	1.000	1.000
	159	1.000	1.000	1.000	1.000
	319	1.000	1.000	1.000	1.000

Table 3. The ASL (under $H_0' : \Sigma = 2I$) and the empirical power (under $H_1' : \Sigma = 2D$) of T, U_j and T_{S1} at Nominal Significance Level $\alpha = 0.05$.

p	$n = N-1$	ASL			Empirical Power		
		T	U_j	T_{S1}	T	U_j	T_{S1}
10	9	0.059	0.049	0.048	1.000	0.412	0.405
40	9	0.055	0.054	0.051	1.000	0.368	0.360
	39	0.056	0.056	0.053	1.000	0.999	0.999
80	9	0.057	0.057	0.053	1.000	0.356	0.348
	39	0.052	0.052	0.050	1.000	0.999	0.999
	79	0.052	0.051	0.050	1.000	1.000	1.000
160	9	0.053	0.056	0.054	1.000	0.354	0.346
	39	0.055	0.056	0.055	1.000	0.999	0.999
	79	0.055	0.057	0.055	1.000	1.000	1.000
	159	0.053	0.052	0.052	1.000	1.000	1.000
320	9	0.052	0.055	0.052	1.000	0.352	0.343
	39	0.052	0.054	0.052	1.000	0.999	0.999
	79	0.050	0.050	0.050	1.000	1.000	1.000
	159	0.051	0.050	0.050	1.000	1.000	1.000
	319	0.053	0.053	0.053	1.000	1.000	1.000

References

Anderson, T.W. 1984. An introduction to multivariate statistical analysis, 2nd edition, Wiley, New York, U.S.A.

Billingsley, P. 1995. Probability and measure, 3rd edtion, Wiley, New York, U.S.A.

Fisher, T., Sun, X., and Gallagher, C. M. 2010. A new test for sphericity of the covariance matrix for high dimensional data. Journal of Multivariate Analysis. 101, 2554-2570.

John, S. 1971. Some optimal multivariate tests. Biometrika. 58, 123-127.

Ledoit, O., and Wolf, M. 2002. Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. Annual Statistics. 30, 1081-1102.

Lehmann, E.L., and Romano, J.P., 2005. Testing statistical hypotheses, 3rd edition, Springer, New York, U.S.A.

Nagao, H., 1973. On some test criteria for covariance matrix. Annual Statistics. 1, 700-709.

Muirhead, R.J. 1982. Aspects of Multivariate Statistical Theory, Wiley, New York, U.S.A.
Notterman, D.A., Alon, U., Sierk, A.J., and Levine, A.J. 2001. Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. Cancer Research. 61, 3124-3130.
Rao, C.R. 1973. Linear statistical inference and its Application, 2nd edition, Wiley, New York, U.S.A.
Rencher, A.C. 2000. Linear models in statistics, Wiley, New York, U.S.A.
Srivastava, M.S., 2005. Some tests concerning the covariance matrix in high dimensional data. Journal of Japan Statistical Society. 35, 251-272.

Appendix

Before proving Theorem 2.1, we need the following information and lemma:

For positive symmetric definite matrix Σ and by spectral decomposition, we have $\Sigma = \Gamma\Lambda\Gamma'$ where $\Lambda = diag(\lambda_1, \lambda_2 ..., \lambda_p)$ with λ_i being the i^{th} eigenvalue of Σ and Γ is an orthogonal matrix with each column as normalized corresponding eigenvectors $\gamma_1, \gamma_2, ..., \gamma_p$. Similarly, we also can write Σ_0 as $\Sigma_0 = RDR'$ where $D = diag(d_1, d_2 ..., d_p)$ with d_i being the i^{th} eigenvalue of Σ_0 and R is an orthogonal matrix with each column as normalized corresponding eigenvectors $r_1, r_2, ..., r_p$ (Rencher, 2003).

Let $nS = YY' \sim W_p(\Sigma, n)$ where $Y = (y_1, y_2, ... y_n)$ and each $y_j \sim N_p(\mathbf{0}, \Sigma)$ and independent (Anderson (1984), Section 3.3; Srivastava (2005); Fisher *et al.* (2010)). Let $U = (u_1, u_2, ... u_n)$ where u_j is independently and identically distributed (iid.) $N_p(\mathbf{0}, I)$ and we can write $Y = \Sigma^{\frac{1}{2}} U$ where $\Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}} = \Sigma$. Define $W' = U'\Gamma' = (w_1, w_2, ..., w_p)$ and each w_i are iid. $N_n(\mathbf{0}, I)$. Thus, define $v_{ii} = w_i' w_i$ are iid chi-squared random variables with n degree of freedom.

Lemma A.1. For $v_{ii} = w_i' w_i$ and $v_{ij} = w_i' w_j$ for any $i \neq j$

$$E(v_{ii}^r) = n(n+2)...(n+2r-2), \quad r = 1, 2, ... \qquad Var(v_{ii}) = 2n,$$

$$Var(v_{ii}^2) = 8n(n+2)(n+3), \qquad E(v_{ii} - n)^3 = 8n,$$

$$E(v_{ii} - n)^4 = 12n(n+4), \qquad E(v_{ii}^2 - n(n+2))^4 = 3n(n+2)[272n^4 + O(n^3)],$$

$$E(v_{ij}^2) = n, \qquad E(v_{ij}^4) = 3n(n+2),$$

$$E(v_{ii} v_{ij}^2) = n(n+2), \qquad E(v_{ii}^2 v_{ij}^2) = n(n+2)(n+4),$$

$$E(v_{ij}^2 v_{ii} v_{jj}) = n(n+2)^2.$$

Proof. The first 6 results can be found in Srivastava (2005) and the last 5 results can be found in Fisher *et al.* (2010).

As in similar proofs of Srivastava (2005), we can write $(1/p)tr(\Sigma_0^{-1}S)$ and $(1/p)tr(\Sigma_0^{-1}S)^2$ in terms of chi-squared random variables.

$$\hat{a}_1 = \frac{1}{p} tr(\Sigma_0^{-1}S) = \frac{1}{p} tr\left((RDR')^{-1} \frac{1}{n} YY' \right) = \frac{1}{np} \sum_{i=1}^{p} \frac{\lambda_i}{d_i} v_{ii}. \tag{A.1}$$

Similarly, we also have

$$\frac{1}{p}tr(\Sigma_0^{-1}S)^2 = \frac{1}{n^2 p}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}v_{ii}^2 + \frac{2}{n^2 p}\sum_{i<j}^{p}\frac{\lambda_i\lambda_j}{d_i d_j}v_{ij}^2, \tag{A.2}$$

where $v_{ij} = w_i' w_j$. We let

$$\hat{a}_2 = \frac{n^2}{(n-1)(n+2)}\frac{1}{p}\left[tr(\Sigma_0^{-1}S)^2 - \frac{1}{n}(tr(\Sigma_0^{-1}S))^2\right]. \tag{A.3}$$

Thus

$$\begin{aligned}\hat{a}_2 &= \frac{n^2}{(n-1)(n+2)}\frac{1}{p}\left[\frac{1}{n^2 p}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}v_{ii}^2 + \frac{2}{n^2 p}\sum_{i<j}^{p}\frac{\lambda_i\lambda_j}{d_i d_j}v_{ij}^2 - \left(\frac{1}{np}\sum_{i=1}^{p}\frac{\lambda_i}{d_i}v_{ii}\right)^2\right] \\ &= \frac{n^2}{(n-1)(n+2)}\left[\frac{n-1}{n^3 p}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}v_{ii}^2 + \frac{2}{n^2 p}\sum_{i<j}^{p}\frac{\lambda_i\lambda_j}{d_i d_j}(v_{ij}^2 - \frac{1}{n}v_{ii}v_{jj})\right] \\ &= \frac{n^2}{(n-1)(n+2)}[b_1 + b_2],\end{aligned} \tag{A.4}$$

where $b_1 = \frac{n-1}{n^3 p}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}v_{ii}^2$, $\qquad b_2 = \frac{2}{n^2 p}\sum_{i<j}^{p}\frac{\lambda_i\lambda_j}{d_i d_j}(v_{ij}^2 - \frac{1}{n}v_{ii}v_{jj})$.

Proof of Theorem 2.1.

Since

$$\begin{aligned}E(\hat{a}_1) &= E\left[\frac{1}{p}tr(\Sigma_0^{-1}S)\right] = E\left(\frac{1}{np}\sum_{i=1}^{p}\frac{\lambda_i}{d_i}v_{ii}\right) = \frac{1}{np}\sum_{i=1}^{p}\frac{\lambda_i}{d_i}E(v_{ii}) = \frac{1}{np}\sum_{i=1}^{p}\frac{\lambda_i}{d_i}n \\ &= \frac{1}{p}\sum_{i=1}^{p}\frac{\lambda_i}{d_i} = \frac{1}{p}tr(\Sigma_0^{-1}\Sigma) = a_1.\end{aligned}$$

And from Lemma A.1, we easily find that $E(v_{ij}^2 - \frac{1}{n}v_{ii}v_{jj}) = 0$ then $E(b_2) = 0$. Thus

$$\begin{aligned}E(\hat{a}_2) &= \frac{n^2}{(n-1)(n+2)}E\left(\frac{n-1}{n^3 p}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}v_{ii}^2\right) = \frac{n^2}{(n-1)(n+2)}\left(\frac{n-1}{n^3 p}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}E(v_{ii}^2)\right) \\ &= \frac{n^2}{(n-1)(n+2)}\frac{n-1}{n^3 p}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}n(n+2) = \frac{1}{p}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2} = \frac{1}{p}tr(\Sigma_0^{-1}\Sigma)^2 = a_2\end{aligned}$$

This is shown that both $\hat{a}_1$ and $\hat{a}_2$ are unbiased estimators of a_1 and a_2 respectively. To show that $\hat{a}_1$ and $\hat{a}_2$ are consistent estimators considered by

$$Var(\hat{a}_1) = Var\left(\frac{1}{np}\sum_{i=1}^{p}\frac{\lambda_i}{d_i}v_{ii}\right) = \frac{1}{n^2p^2}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}Var(v_{ii}) = \frac{2}{np}\left(\frac{1}{p}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}\right) = \frac{2}{np}a_2. \tag{A.5}$$

And since $\hat{a}_2 = \dfrac{n^2}{(n-1)(n+2)}\left[b_1 + b_2\right]$ thus

$$Var(\hat{a}_2) = \left(\frac{n^2}{(n-1)(n+2)}\right)^2 \left\{Var(b_1) + Var(b_2) + 2COV(b_1, b_2)\right\}.$$

$$\begin{aligned} Var(b_1) &= \frac{(n-1)^2}{n^6 p^2} Var\left(\sum_{i=1}^{p} \frac{\lambda_i^2}{d_i^2} v_{ii}^2\right) = \frac{(n-1)^2}{n^6 p^2} \sum_{i=1}^{p} \frac{\lambda_i^4}{d_i^4} Var(v_{ii}^2) \\ &= \frac{8(n-1)^2(n+2)(n+3)}{n^5 p} a_4. \end{aligned} \tag{A.6}$$

$$\begin{aligned} Var(b_2) &= \frac{4}{n^4 p^2}\left(\sum_{i<j}^{p} \frac{\lambda_i \lambda_j}{d_i d_j} Var\left(v_{ij}^2 - \frac{1}{n} v_{ii} v_{jj}\right)\right) \\ &= \frac{4}{n^4 p^2} \sum_{i<j}^{p} \frac{\lambda_i^2 \lambda_i^2}{d_i^2 d_i^2}\left(E\left(v_{ij}^2 - \frac{1}{n} v_{ii} v_{jj}\right)^2 - \left(E\left(v_{ij}^2 - \frac{1}{n} v_{ii} v_{jj}\right)\right)^2\right) \\ &= \frac{4(n-1)(n+2)}{n^4}\left(a_2^2 - \frac{1}{p} a_4\right). \end{aligned} \tag{A.7}$$

And since $E(b_2) = 0$ then

$$\begin{aligned} COV(b_1, b_2) &= E(b_1 b_2) \\ &= \frac{2(n-1)}{n^5 p} E\left[\sum_{i=1}^{p} \frac{\lambda_i^2}{d_i^2} v_{ii}^2 \sum_{i<j}^{p} \frac{\lambda_j \lambda_k}{d_i d_j}\left(v_{jk}^2 - \frac{1}{n} v_{jj} v_{kk}\right)\right] \\ &= \frac{2(n-1)}{n^5 p}\left\{\begin{aligned} &E\left[\frac{\lambda_1^2}{d_1^2} v_{11}^2 \sum_{i<j}^{p} \frac{\lambda_j \lambda_k}{d_i d_j}\left(v_{jk}^2 - \frac{1}{n} v_{jj} v_{kk}\right)\right] + E\left[\frac{\lambda_2^2}{d_2^2} v_{22}^2 \sum_{i<j}^{p} \frac{\lambda_j \lambda_k}{d_i d_j}\left(v_{jk}^2 - \frac{1}{n} v_{jj} v_{kk}\right)\right] \\ &+ \ldots + E\left[\frac{\lambda_p^2}{d_p^2} v_{pp}^2 \sum_{i<j}^{p} \frac{\lambda_j \lambda_k}{d_i d_j}\left(v_{jk}^2 - \frac{1}{n} v_{jj} v_{kk}\right)\right] \end{aligned}\right\} \\ &= 0 \end{aligned} \tag{A.8}$$

because

$$E\left[v_{ii}^2 \sum_{i<j}^{p} \frac{\lambda_j\lambda_k}{d_i d_j}\left(v_{jk}^2 - \frac{1}{n} v_{jj} v_{kk} \right) \right]$$

$$= \sum_{i<j}^{p} \frac{\lambda_j\lambda_k}{d_i d_j} E\left(v_{ii}^2 v_{jk}^2 - \frac{1}{n} v_{ii}^2 v_{jj} v_{kk} \right) = \sum_{i<j}^{p} \frac{\lambda_j\lambda_k}{d_i d_j}\left(n(n+2)(n) - \frac{1}{n} n(n+2)(n)(n) \right) = 0, \quad for \;\; i \neq j \neq k,$$

$$= \sum_{i<j}^{p} \frac{\lambda_j\lambda_k}{d_i d_j} E\left(v_{ii}^2 v_{ik}^2 - \frac{1}{n} v_{ii}^3 v_{kk} \right) = \sum_{i<j}^{p} \frac{\lambda_j\lambda_k}{d_i d_j}\left(n(n+2)(n+4) - \frac{1}{n} n(n+2)(n+4)(n) \right) = 0, \quad for \;\; i = j \neq k, \tag{A.9}$$

$$= \sum_{i<j}^{p} \frac{\lambda_j\lambda_k}{d_i d_j} E\left(v_{ii}^2 v_{ij}^2 - \frac{1}{n} v_{ii}^3 v_{jj} \right) = \sum_{i<j}^{p} \frac{\lambda_j\lambda_k}{d_i d_j}\left(n(n+2)(n+4) - \frac{1}{n} n(n+2)(n+4)(n) \right) = 0, \quad for \;\; i \neq j = k.$$

By (A.6) – (A.8), then we have

$$Var(\hat{a}_2) = \left(\frac{n^4}{(n-1)^2 (n+2)^2} \right)\left\{ Var(b_1) + Var(b_2) + 2COV(b_1, b_2) \right\}$$

$$= \left(\frac{n^4}{(n-1)^2 (n+2)^2} \right)\left\{ \frac{8(n-1)^2 (n+2)(n+3)}{n^5 p} a_4 + \frac{4(n-1)(n+2)}{n^4}\left(a_2^2 - \frac{1}{p} a_4 \right) \right\} \tag{A.10}$$

$$= \frac{4(2n^2 + 3n - 6)}{n(n-1)(n+2)p} a_4 + \frac{4}{(n-1)(n+2)} a_2^2.$$

Since $\hat{a}_1$ and $\hat{a}_2$ are unbiased estimators of a_1 and a_2, respectively and from (A.5), (A.9), and by applying the Chebyshev's inequality, for any $\varepsilon > 0$ as $(n, p) \to \infty$,

$$P\left[\left| \hat{a}_1 - a_1 \right| > \varepsilon \right] \leq \frac{1}{\varepsilon^2} Var(\hat{a}_1) = \frac{1}{\varepsilon^2} \frac{2a_2}{np} \to 0 \quad \text{and}$$

$$P\left[\left| \hat{a}_2 - a_2 \right| > \varepsilon \right] \leq \frac{1}{\varepsilon^2} Var(\hat{a}_2) = \frac{1}{\varepsilon^2}\left\{ \frac{4(2n^2 + 3n - 6)}{n(n-1)(n+2)p} a_4 + \frac{4}{(n-1)(n+2)} a_2^2 \right\} \approx \left(\frac{8}{np} a_4 + \frac{4}{n^2} a_2^2 \right) \to 0.$$

Hence $\hat{a}_1$ and $\hat{a}_2$ are unbiased and consistent estimators of a_1 and a_2, respectively. The proof is completed.

Proof of Theorem 2.2.

From Theorem 2.1, we have

$$E(\hat{a}_1) = a_1, \qquad E(\hat{a}_2) = a_2 \tag{A.11}$$

By Lemma A.1., with simple calculations and in similar proofs of Srivastava (2005) under assumption (A), and (B), and as $(n, p) \to \infty$, we obtain

$$Var(\hat{a}_1) = 2a_2 / np, \tag{A.12}$$

$$Var(b_1) = \frac{8(n-1)^2 (n+2)(n+3)}{n^5 p} a_4 \approx 8a_4 / np, \tag{A.13}$$

$$Var(b_2) = \frac{4(n-1)(n+2)}{n^4}\left(a_2^2 - \frac{1}{p} a_4 \right) \approx 4c(a_2^2 - a_4 / p) / np, \tag{A.14}$$

$$Var(\hat{a}_2) = \frac{4(2n^2+3n-6)}{n(n-1)(n+2)p}a_4 + \frac{4}{(n-1)(n+2)}a_2^2 \approx 4(2a_4 + ca_2^2)/np, \tag{A.15}$$

$$\begin{aligned}
COV(\hat{a}_1, b_1) &= E(\hat{a}_1 b_1) - E(\hat{a}_1)E(b_1) \\
&= \frac{n-1}{n^4 p}E\left[\sum_{i=1}^{p}\frac{\lambda_i}{d_i}v_{ii}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}v_{ii}^2\right] - E\left[\frac{1}{np}\sum_{i=1}^{p}\frac{\lambda_i}{d_i}v_{ii}\right]E\left[\frac{n-1}{n^3 p}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}v_{ii}^2\right] \\
&= \frac{n-1}{n^4 p}E\left[\sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3}v_{ii}^3 + \sum_{i\neq j}^{p}\frac{\lambda_i\lambda_j^2}{d_i d_j^2}v_{ii}v_{jj}^2\right] - \frac{(n+1)(n+2)}{n^2p^2}\sum_{i=1}^{p}\frac{\lambda_i}{d_i}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}
\end{aligned}$$

$$\begin{aligned}
&= \frac{n-1}{n^4p^2}\left[\sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3}n(n+2)(n+4) + \sum_{i\neq j}^{p}\frac{\lambda_i\lambda_j^2}{d_i d_j^2}n^2(n+2)\right] \\
&\quad - \frac{(n+1)(n+2)}{n^2p^2}\left[\sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3} + \sum_{i\neq j}^{p}\frac{\lambda_i\lambda_j^2}{d_i d_j^2}\right] \\
&= \frac{(n-1)(n)(n+2)(n+4)}{n^4p^2}\sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3} + \frac{(n-1)n^2(n+2)}{n^4p^2}\sum_{i\neq j}^{p}\frac{\lambda_i\lambda_j^2}{d_i d_j^2} \\
&\quad - \frac{(n-1)(n+2)}{n^2p^2}\sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3} - \frac{(n-1)(n+2)}{n^2p^2}\sum_{i\neq j}^{p}\frac{\lambda_i\lambda_j^2}{d_i d_j^2} \\
&= \frac{(n-1)(n+2)(n+4)}{n^3p^2}\sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3} - \frac{(n-1)(n+2)}{n^2p^2}\sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3} \\
&= \frac{4(n-1)(n+2)a_3}{n^3p^2} \approx \frac{4a_3}{np}.
\end{aligned} \tag{A.16}$$

From the fact that $E(b_2) = 0$ and similar to the proof for $E(b_1 b_2)$

$$COV(\hat{a}_1, b_2) = E(\hat{a}_1 b_2) = \frac{2}{n^3p^2}E\left[\left(\sum_{i=1}^{p}\frac{\lambda_i}{d_i}v_{ii}\right)\left(\sum_{i<j}^{p}\frac{\lambda_i\lambda_j}{d_i d_j}\left(v_{ij}^2 - \frac{1}{n}v_{ii}v_{jj}\right)\right)\right] = 0. \tag{A.17}$$

Note that

$$E(\hat{a}_1\hat{a}_2) = E\left[\frac{1}{np}\sum_{i=1}^{p}\frac{\lambda_i}{d_i}v_{ii}\frac{n^2}{(n-1)(n+2)}\left(\frac{n-1}{n^3 p}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}v_{ii}^2 + \frac{2}{n^2 p}\sum_{i<j}^{p}\frac{\lambda_i\lambda_j}{d_i d_j}\left(v_{ij}^2 - \frac{1}{n}v_{ii}v_{jj}\right)\right)\right]$$

$$= E\left[\frac{1}{(n+2)n^2p^2}\sum_{i=1}^{p}\frac{\lambda_i}{d_i}v_{ii}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}v_{ii}^2 + \frac{2}{(n-1)(n+2)np^2}\sum_{i=1}^{p}\frac{\lambda_i}{d_i}v_{ii}\sum_{i<j}^{p}\frac{\lambda_i\lambda_j}{d_i d_j}\left(v_{ij}^2 - \frac{1}{n}v_{ii}v_{jj}\right)\right]$$

(A.18)

By similar proof to $E(b_1b_2)$ we have $E\left\{\sum_{i=1}^{p}\frac{\lambda_i}{d_i}v_{ii}\sum_{i<j}^{p}\frac{\lambda_i\lambda_j}{d_i d_j}\left(v_{ij}^2 - \frac{1}{n}v_{ii}v_{jj}\right)\right\} = 0$ then the expectation of the second term in (A.18) equals to zero. Thus, we obtain that

$$E(\hat{a}_1\hat{a}_2) = \frac{1}{(n+2)n^2p^2}E\left[\sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3}v_{ii}^3 + \sum_{i\neq j}^{p}\frac{\lambda_i\lambda_j^2}{d_i d_j^2}v_{ii}v_{jj}^2\right]$$

$$= \frac{1}{(n+2)n^2p^2}\left[\sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3}n(n+2)(n+4) + \sum_{i\neq j}^{p}\frac{\lambda_i\lambda_j^2}{d_i d_j^2}n^2(n+2)\right]$$

$$= \frac{(n+4)}{np^2}\sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3} + \frac{1}{p^2}\sum_{i\neq j}^{p}\frac{\lambda_i\lambda_j^2}{d_i d_j^2} \qquad (A.19)$$

$$= \frac{(n+4)}{np^2}\sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3} + \frac{1}{p^2}\left[\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}\sum_{i=1}^{p}\frac{\lambda_i}{d_i} - \sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3}\right]$$

$$= \frac{4}{np^2}\sum_{i=1}^{p}\frac{\lambda_i^3}{d_i^3} + \frac{1}{p^2}\sum_{i=1}^{p}\frac{\lambda_i^2}{d_i^2}\sum_{i=1}^{p}\frac{\lambda_i}{d_i}$$

$$= 4a_3 / np + a_2a_1.$$

By (A.11) and (A.19) as $(n, p) \to \infty$, we obtain

$$COV(\hat{a}_1,\hat{a}_2) = E(\hat{a}_1\hat{a}_2) - E(\hat{a}_1)E(\hat{a}_2) = 4a_3 / np \qquad (A.20)$$

To find the distribution of $\hat{a}_1$ and $\hat{a}_2$, we used Multivariate central limit theorem (Rao,1973,p.147) and Lindeberg Central Limit Theorem (Billingsley, 1995, p.359)

Since $\hat{a}_2 = \frac{n^2}{(n-1)(n+2)}[b_1 + b_2]$, so we need to find the distribution of $\hat{a}_1, b_1$ and b_2 which will distribute as Normal distribution, respectively. First, we find the distribution of $\hat{a}_1, b_1$ because both are functions of v_{ii} and the second is of b_2 because it is a function of $v_{ij}, i \neq j$. Finally, the distribution of $\hat{a}_2$ which is a distribution of a linear function of two normal random variables is obtained.

First, in order to find the distribution of $\hat{a}_1$ and b_1. Under λ_i and d_i as before, we let

$$u_{1i} = \frac{\lambda_i (v_{ii} - n)}{d_i \sqrt{n}} \text{ and } u_{2i} = \frac{\lambda_i^2 (v_{ii}^2 - n(n+2))}{d_i^2 \sqrt{n(n+2)(n+3)}}$$

where

$$E(u_{1i}) = 0,\ E(u_{2i}) = 0,\ Var(u_{1i}) = 2\lambda_i^2 / d_i^2,\ Var(u_{2i}) = 8\lambda_i^4 / d_i^4,\ COV(u_{1i}, u_{2i}) = 4\lambda_i^3 e_n / d_i^3$$ and

$e_n = \sqrt{n+2} / \sqrt{n+3} \approx 1$ as $n \to \infty$ Since v_{ii} s are independent, thus $\mathbf{u}_i = (u_{1i}, u_{2i})'$ are independently distributed random vectors, $i = 1, ..., p$ with $E(\mathbf{u}_i) = \mathbf{0}$ and covariance matrices Ω_{in} given by

$$\Omega_{in} = \begin{bmatrix} 2\lambda_i^2 / d_i^2 & 4\lambda_i^3 e_n / d_i^3 \\ 4\lambda_i^3 e_n / d_i^3 & 8\lambda_i^4 / d_i^4 \end{bmatrix}, \qquad i = 1, ..., p .$$

For any n as $p \to \infty$

$$\Omega_n = (\Omega_{1n} + ... + \Omega_{pn}) / p$$

$$= \begin{bmatrix} \frac{2}{p} \sum_{i=1}^{p} \frac{\lambda_i^2}{d_i^2} & \frac{4e_n}{p} \sum_{i=1}^{p} \frac{\lambda_i^3}{d_i^3} \\ \frac{4e_n}{p} \sum_{i=1}^{p} \frac{\lambda_i^3}{d_i^3} & \frac{8}{P} \sum_{i=1}^{p} \frac{\lambda_i^4}{d_i^4} \end{bmatrix} = \begin{bmatrix} 2a_2 & 4e_n a_3 \\ 4e_n a_3 & 8a_4 \end{bmatrix} \to \Omega_n^0 \neq \mathbf{0}$$

where $\Omega_n^0 = \begin{bmatrix} 2a_2^0 & 4e_n a_3^0 \\ 4e_n a_3^0 & 8a_4^0 \end{bmatrix}$.

If F_i is the distribution function of $\mathbf{u}_i$ then

$$\frac{1}{p} \sum_{i=1}^{p} \int_{\sqrt{\mathbf{u}_i' \mathbf{u}_i} > \varepsilon \sqrt{p}} \mathbf{u}_i' \mathbf{u}_i \, dF_i \le \frac{1}{p} \sum_{i=1}^{p} \frac{1}{\varepsilon^2 p} \int (\mathbf{u}_i' \mathbf{u}_i)^2 dF_i = \frac{1}{\varepsilon^2 p^2} \sum_{i=1}^{p} E(u_{1i}^2 + u_{2i}^2)^2 \le \frac{2}{\varepsilon^2 p^2} \sum_{i=1}^{p} E(u_{1i}^4 + u_{2i}^4),$$

from $C_r -$ inequality in Rao (1973, p.149). Since as $p \to \infty$ and from Lemma A.1.,

$$\frac{2}{\varepsilon^2 p^2} \sum_{i=1}^{p} E(u_{1i}^4) = \frac{2}{\varepsilon^2 p^2} \sum_{i=1}^{p} \frac{\lambda_i^4}{d_i^4 n^2} E(v_{ii} - n)^4 = \frac{2}{\varepsilon^2 p^2} \sum_{i=1}^{p} \frac{\lambda_i^4}{d_i^4 n^2} 12n(n+4) \to 0$$

and by an analogous derivation as $p \to \infty$,

Hence $\frac{2}{\varepsilon^2 p^2} \sum_{i=1}^{p} E(u_{1i}^4 + u_{2i}^4) \to 0$ as $p \to \infty$. By applying the multivariate central limit theorem, as $p \to \infty$ for any n

$$\frac{1}{\sqrt{p}}(\mathbf{u}_1+\mathbf{u}_2+\ldots+\mathbf{u}_p)=\begin{pmatrix}\frac{1}{\sqrt{np}}\sum_{i=1}^{p}\frac{\lambda_i(v_{ii}-n)}{d_i}\\ \frac{1}{\sqrt{np}}\sum_{i=1}^{p}\frac{\lambda_i^2(v_{ii}^2-n(n+2))}{d_i^2\sqrt{(n+2)(n+3)}}\end{pmatrix}\xrightarrow{D}N_2(\mathbf{0},\Omega_n^0)$$

Note that as $n\to\infty$, $e_n\approx 1$, $\Omega_n^0=\begin{bmatrix}2a_2^0 & 4e_na_3^0\\ 4e_na_3^0 & 8a_4^0\end{bmatrix}\to\Omega^0$, where $\Omega^0=\begin{bmatrix}2a_2^0 & 4a_3^0\\ 4a_3^0 & 8a_4^0\end{bmatrix}$. Thus, it follows that as $(n,p)\to\infty$,

$$\begin{pmatrix}\frac{1}{\sqrt{np}}\sum_{i=1}^{p}\frac{\lambda_i(v_{ii}-n)}{d_i}\\ \frac{1}{\sqrt{np}}\sum_{i=1}^{p}\frac{\lambda_i^2(v_{ii}^2-n(n+2))}{d_i^2\sqrt{(n+2)(n+3)}}\end{pmatrix}\xrightarrow{D}N_2(\mathbf{0},\Omega^0).$$

And under assumption (A) which leads to assuming that $\Omega\to\Omega^0$, where $\Omega=\begin{bmatrix}2a_2 & 4a_3\\ 4a_3 & 8a_4\end{bmatrix}$ then we have that

$$\begin{pmatrix}\frac{1}{\sqrt{np}}\sum_{i=1}^{p}\frac{\lambda_i(v_{ii}-n)}{d_i}\\ \frac{1}{\sqrt{np}}\sum_{i=1}^{p}\frac{\lambda_i^2(v_{ii}^2-n(n+2))}{d_i^2\sqrt{(n+2)(n+3)}}\end{pmatrix}\xrightarrow{D}N_2(\mathbf{0},\Omega).$$

For the first element in the previous random vector, since

$$\frac{1}{\sqrt{np}}\sum_{i=1}^{p}\frac{\lambda_i(v_{ii}-n)}{d_i}=\frac{1}{\sqrt{np}}\left[\sum_{i=1}^{p}\frac{\lambda_iv_{ii}}{d_i}-\sum_{i=1}^{p}\frac{n\lambda_i}{d_i}\right]=\frac{1}{\sqrt{np}}(np\hat{a}_1-npa_1)=\sqrt{np}(\hat{a}_1-a_1)\xrightarrow{D}N(0,2a_2),$$

then

$$\hat{a}_1\xrightarrow{D}N(a_1,2a_2/np). \tag{A.21}$$

For the second element, we have that

$$\frac{1}{\sqrt{np}}\sum_{i=1}^{p}\frac{\lambda_i^2(v_{ii}^2-n(n+2))}{d_i^2\sqrt{(n+2)(n+3)}}=\frac{1}{\sqrt{np}}\left[\sum_{i=1}^{p}\frac{\lambda_i^2v_{ii}^2}{d_i^2\sqrt{(n+2)(n+3)}}-\sum_{i=1}^{p}\frac{\lambda_i^2n(n+2)}{d_i^2\sqrt{(n+2)(n+3)}}\right]$$

$$=\frac{1}{\sqrt{np}}\left[\frac{n^3pb_1}{(n-1)\sqrt{(n+2)(n+3)}}-\frac{n(n+2)pa_2}{\sqrt{(n+2)(n+3)}}\right]\xrightarrow{D}N(0,8a_4).$$

Since as $n\to\infty$,

$$\frac{1}{\sqrt{np}}\left[\frac{n^3 pb_1}{(n-1)\sqrt{(n+2)(n+3)}}-\frac{n(n+2)pa_2}{\sqrt{(n+2)(n+3)}}\right]\approx\frac{1}{\sqrt{np}}(npb_1-npa_2)=\sqrt{np}(b_1-a_2)$$

then $\sqrt{np}(b_1-a_2)\xrightarrow{D}N(0,8a_4)$ also, and with a linear transformation we have the result that

$$b_1\xrightarrow{D}N(a_2,8a_4/np). \tag{A.22}$$

The next is to find the distribution of b_2. Srivastava (2005) gave the important results, which are used for the next proof, that $v_{ij}/\sqrt{n}\sim N(0,1)$ as $n\to\infty$ and $v_{ij}^2/n\sim\chi_1^2$ which are asymptotically independently distributed for all distinct i and j.

Note that b_2 defined in (A.4), now we let $\eta_{ij}=\frac{2\lambda_i\lambda_j}{n^2pd_id_j}(v_{ij}^2-\frac{1}{n}v_{ii}v_{ij})$. From Lemma A.1., we have $E(\eta_{ij})=0$ and let

$$\begin{aligned}
S_p^2&=\sum_{i<j}^{p}Var(\eta_{ij})=\sum_{i<j}^{p}Var\left\{\frac{2\lambda_i\lambda_j}{n^2pd_id_j}(v_{ij}^2-\tfrac{1}{n}v_{ii}v_{jj})\right\}\\
&=\frac{4}{n^4p^2}\sum_{i<j}^{p}\frac{\lambda_i\lambda_j}{d_id_j}Var\left\{(v_{ij}^2-\tfrac{1}{n}v_{ii}v_{jj})\right\}\\
&=Var(b_2)\\
&=\frac{4(n-1)(n+2)}{n^4}\left(a_2^2-\frac{a_4}{p}\right)\approx\frac{4c}{n^2}\left(a_2^2-\frac{a_4}{p}\right)
\end{aligned}$$

as $(n,p)\to\infty$.

Let $M_p=\sum_{i<j}^{p}\eta_{ij}=\frac{2}{n^2p}\sum_{i<j}^{p}\frac{\lambda_i\lambda_j}{d_id_j}(v_{ij}^2-\frac{1}{n}v_{ii}v_{ij})=b_2$. If P_{ij} is the distribution function of η_{ij}. Since, for $\varepsilon>0$

$$\begin{aligned}
\sum_{i<j}^{p}\frac{1}{S_p^2}\int_{|\eta_{ij}|>\varepsilon S_p}\eta_{ij}^2dP_{ij}&<\sum_{i<j}^{p}\frac{1}{\varepsilon^2S_p^2}\int\eta_{ij}^2dP_{ij}\\
&=\sum_{i<j}^{p}\frac{1}{\varepsilon^2S_p^2}E(\eta_{ij}^2)\\
&=\sum_{i<j}^{p}\frac{4}{n^4p^2\varepsilon^2S_p^2}E\left(\frac{\lambda_i\lambda_j}{d_id_j}\left(v_{ij}^2-\frac{1}{n}v_{ii}v_{jj}\right)\right)^2\\
&=\sum_{i<j}^{p}\frac{8(n-1)(n+2)\lambda_i^2\lambda_j^2}{n^4p^2\varepsilon^2S_p^2d_i^2d_j^2}\\
&\approx\sum_{i<j}^{p}\frac{8\lambda_i^2\lambda_j^2}{n^2p^2\varepsilon^2S_p^2d_i^2d_j^2}\to 0
\end{aligned}$$

as $p \to \infty$. Then, it follows from the Lindeberg Central Limit Theorem in Lemma A.3.,

$$\frac{M_p}{S_p} = \frac{\sqrt{np}\,b_2}{2\sqrt{c(a_2^2 - a_4 / p)}} \xrightarrow{D} N(0,1).$$

Then we have $b_2 \xrightarrow{D} N\left(0, \frac{4c}{np}(a_2^2 - a_4 / p)\right)$. (A.23)

By (A.8), then b_1 and b_2 are asymptotically independent. Note that $\hat{a}_2$ is a linear function of two random variables b_1 and b_2 that is, $\hat{a}_2 = \frac{n^2}{(n-1)(n+2)}[b_1 + b_2] \approx b_1 + b_2$ as $n \to \infty$. By (A.5), (A.15.), (A.22.), and (A.23.), then we have

$$\hat{a}_2 \xrightarrow{D} N\left(a_2, 4(2a_4 + ca_2^2) / np\right). \quad \text{(A.24)}$$

From (A.20), $COV(\hat{a}_1, \hat{a}_2) = 4a_3 / np$, (A.21), and (A.24), we have

$$\begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} \xrightarrow{D} N_2\left[\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} 2a_2 / np & 4a_3 / np \\ 4a_3 / np & 4(2a_4 + ca_2^2) / np \end{pmatrix}\right].$$

The proof is completed. □

Proof of Theorem 2.3. Note that our test statistic is $\hat{\psi} = \hat{a}_2 - 2t\hat{a}_1 + t^2$ and we have

$$\frac{\partial \hat{\psi}}{\partial \hat{a}_1} = -2t \text{ and } \frac{\partial \hat{\psi}}{\partial \hat{a}_2} = 1.$$

By applying the delta method (Lehmann and Romano, 2005, p.436), thus, $\hat{\psi} - \psi \xrightarrow{D} N(0, \delta^2)$ where

$$\delta^2 = \begin{pmatrix} -2t & 1 \end{pmatrix} \begin{pmatrix} 2a_2 / np & 4a_3 / np \\ 4a_3 / np & 4(2a_4 + ca_2^2) / np \end{pmatrix} \begin{pmatrix} -2t \\ 1 \end{pmatrix} = \frac{4}{np}(2t^2 a_2 - 4ta_3 + 2a_4 + ca_2^2)$$

The proof is completed. □

Proof of Corollary 2.1. Under H_0, $a_2 = t^2$, $a_3 = t^3$ and $a_4 = t^4$ Thus, $\delta^2 = 4ct^4 / np$. It follows from Theorem 2.3. that the null asymptotic distribution of T is $N(0,1)$. The proof is completed. □