A New Test on High-Dimensional Mean Vector Without Any Assumption on Population Covariance Matrix

SHOTA KATAYAMA AND YUTAKA KANO

Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan

Abstract

We propose a new test for the equality of the mean vectors between a two groups with the same number of the observations in high-dimensional data. The existing tests for this problem require a strong condition on the population covariance matrix. The proposed test in this paper does not require such conditions for it. This test will be obtained in a general model, that is, the data need not be normally distributed.

1 Introduction

Advances of information technology and database system make it possible to collect and store "high-dimensional data" which have fewer observations than or equal to the dimension. High-dimensional data appears in various fields, such as DNA microarray data analysis and marketing data analysis. In such high-dimensional data, however, traditional multivariate analysis is not often applicable.

In this paper, we propose a new test for the equality of two mean vectors of n observations which are given by $(\boldsymbol{X}_p^{(i1)}, \boldsymbol{X}_p^{(i2)}), i = 1, \ldots, n$, where $\boldsymbol{X}_p^{(ij)}$ are p-dimensional random vectors. Within the group $j = 1, 2$, $\boldsymbol{X}_p^{(ij)}$ $(i = 1, \ldots, n)$ are independent, while the two groups may be dependent. In a specific situation, we can regard $\boldsymbol{X}_p^{(i1)}$ and $\boldsymbol{X}_p^{(i2)}$ as the two measurements on the same subject i before and after a treatment. Let $\boldsymbol{X}_p^{(i)} = \boldsymbol{X}_p^{(i1)} - \boldsymbol{X}_p^{(i2)}$, $\boldsymbol{\mu}_p = E(\boldsymbol{X}_p^{(i)})$ and $\boldsymbol{\Sigma}_p = \mathrm{Var}(\boldsymbol{X}_p^{(i)})$, then the problem is a one-sample problem. We wish to test the following hypothesis:

$$H_0 : \boldsymbol{\mu}_p = \boldsymbol{0} \quad \text{versus} \quad H_1 : \boldsymbol{\mu}_p \neq \boldsymbol{0}. \tag{1.1}$$

A traditional method to test the null hypothesis is Hotelling's T^2 test. However, Hotelling's T^2 test is not applicable when $n \leq p$ since the inverse of the

sample covariance matrix does not exist. Bai and Saranadasa (1996) propose a test for the two sample problem with equal population covariance matrices. For the one sample problem in (1.1), the test is based on the statistic

$$F_{n,p} = \bar{\boldsymbol{X}}_{n,p}^T \bar{\boldsymbol{X}}_{n,p} - \frac{1}{n}\mathrm{tr}\boldsymbol{S}_{n,p}, \tag{1.2}$$

where $\bar{\boldsymbol{X}}_{n,p}$ is the sample mean vector and $\boldsymbol{S}_{n,p}$ is the sample covariance matrix. More recently, Chen and Qin (2010) propose a test based on the statistic

$$G_{n,p} = \frac{1}{n(n-1)} \sum_{i \neq j}^{n} \boldsymbol{X}_p^{(i)T} \boldsymbol{X}_p^{(j)} \tag{1.3}$$

for the one sample problem in (1.1). They also propose a test for the two sample problem with unequal population covariance matrices. The asymptotic normality of the test statistic is obtained under the condition $\mathrm{tr}\boldsymbol{\Sigma}_p^4 = o\{(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2\}$ and the null hypothesis H_0 in (1.1) when the sample size n and the dimension p tend to infinity.

We easily show that $F_{n,p} = G_{n,p}$. Hence, for the one sample problem in (1.1), the asymptotic normality of Bai and Saranadasa's test statistic is also obtained under the condition $\mathrm{tr}\boldsymbol{\Sigma}_p^4 = o\{(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2\}$ which is weaker than the original one. However, this condition is still restrictive. Indeed, the condition exclude some typical situation such as $\boldsymbol{\Sigma}_p = (1-\rho)\boldsymbol{I}_p + \rho \mathbf{1}_p \mathbf{1}_p^T, \rho \in (0,1)$, where $\boldsymbol{I}_p$ is the $p \times p$ identity matrix and $\mathbf{1}_p$ is the p-column vector with all entries one, or the case where the maximum eigenvalue of $\boldsymbol{\Sigma}_p$ has larger order of p than $1/2$. Moreover, it is hard to estimate whether the condition is satisfied or not from the data because of its high-dimensionality.

When the condition is not satisfied, the population covariance matrix makes an enormous effect on the asymptotic null distributions of the two test statistics. Katayama et al. (2010) show that the type of the asymptotic null distribution of Bai and Saranadasa's test statistic depends on the population covariance matrix, hence so is Chen and Qin's test statistic. To illustrate this phenomenon, we give a simple example. Suppose temporarily $\boldsymbol{X}_p^{(i)}$'s are independently and identically distributed (i.i.d.) as p-dimensional multivariate normal distribution with mean vector $\mathbf{0}_p$ and covariance matrix $\boldsymbol{\Sigma}_p = (1-\rho_p)\boldsymbol{I}_p + \rho_p \mathbf{1}_p \mathbf{1}_p^T, \rho_p = p^{-\alpha}, \alpha \in (0,1)$. Here, $\mathbf{0}_p$ denotes the p-column vector with all entries zero. Note that the condition is satisfied only when $1/2 < \alpha < 1$. Since the two test statistics are equivalent, we consider Chen

and Qin's test statistic only. We define the standardized version of the test statistic as follows:

$$\tilde{T}_{CQ} = \frac{G_{n,p}}{\sqrt{2\mathrm{tr}\boldsymbol{\Sigma}_p^2/n(n-1)}}.$$

The distributions of $\tilde{T}_{CQ}$ are plotted for $(n,p) = (80,300), (80,500)$ based on $10,000$ simulations in Figure 1. This shows that the type of the asymptotic null distributions of $\tilde{T}_{CQ}$ depends on α, hence on $\boldsymbol{\Sigma}_p$. When the condition is not satisfied, i.e., $0 < \alpha \leq 1/2$, the asymptotic distribution of $\tilde{T}_{CQ}$ is obviously not normal for the two cases considered. Indeed, the asymptotic distribution is standardized chi-square distribution with 1 degree of freedom for $0 < \alpha < 1/2$, and the convolution of normal and chi-square distributions for $\alpha = 1/2$. See Katayama et al. (2010) for more details. Hence the type I error of Chen and Qin's test would be asymptotically incorrect.

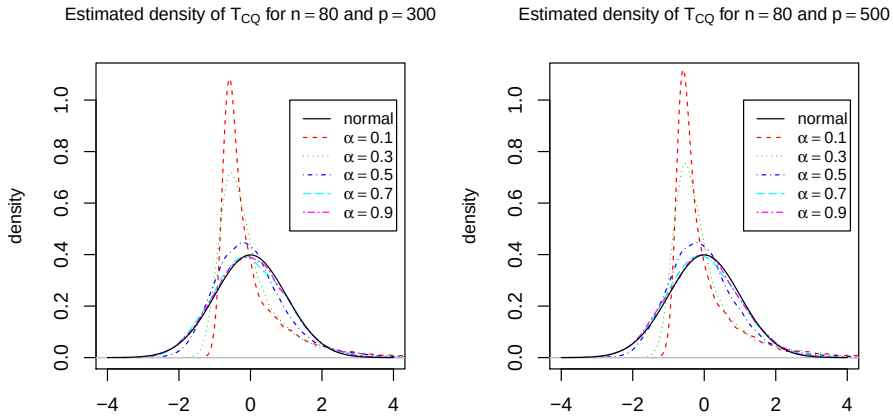

Figure 1: Distribution of $\tilde{T}_{CQ}$, when $n = 80$ and $p = 300, 500$, based on $10,000$ Monte Carlo simulations

Motivated by these findings, in this paper, we construct a test statistic which has the asymptotic normality under H_0 without any assumption on population covariance matrix. Hence, the type I error of the proposed test is always asymptotically correct when we use the percentile point of normal

distribution. The asymptotic normality is obtained without assuming that the random vectors are i.i.d. as multivariate normal distribution.

The paper is organized as follows. In Section 2, we introduce the proposed test and obtain the asymptotic normality. Several simulations and a real data example are given in Section 3. All the proofs are given in Section 4.

2 Proposed Test

In this section, we propose a test statistic for (1.1). The asymptotic normality of the test statistic is obtained without any assumption on the population covariance matrix under H_0. Like Srivastava (2009), we assume

$$\boldsymbol{X}_p^{(i)} = \boldsymbol{\mu}_p + \boldsymbol{C}_p \boldsymbol{Z}_p^{(i)}, \quad i = 1, \ldots, n, \tag{2.1}$$

where $\boldsymbol{C}_p$ is a $p \times p$ matrix with $\boldsymbol{\Sigma}_p = \boldsymbol{C}_p \boldsymbol{C}_p^T$ and $\boldsymbol{Z}_p^{(i)} = (z_{i1}, \ldots, z_{ip})^T$. Assume that z_{ij}'s are independent random variables which satisfy

$$E(z_{ij}) = 0, \quad E(z_{ij}^2) = 1, \quad E(z_{ij}^4) = \gamma < \infty.$$

For the high-dimensional inference, we assume the following condition:

Condition A. $n = n(p) \to \infty$ as $p \to \infty$.

Now we propose the test statistic as

$$T_{n,p} = \mathrm{tr} \boldsymbol{X}_{n,p}^T \boldsymbol{M}_n \boldsymbol{X}_{n,p}, \tag{2.2}$$

where $\boldsymbol{X}_{n,p} = (\boldsymbol{X}_p^{(1)}, \ldots, \boldsymbol{X}_p^{(n)})^T$ is the data matrix and $\boldsymbol{M}_n$ is a known $n \times n$ symmetric matrix which satisfies the following condition:

Condition B. $\boldsymbol{M}_n = (m_{ij})$ with $m_{ij} \geq 0$ and $m_{ii} = 0$, $\mathbf{1}_n^T \boldsymbol{M}_n \mathbf{1}_n \neq 0$, $\mathrm{tr} \boldsymbol{M}_n^4 = o\{(\mathrm{tr} \boldsymbol{M}_n^2)^2\}$.

The proposed test statistic is also defined as

$$T_{n,p} = \sum_{i \neq j}^{n} m_{ij} \boldsymbol{X}_p^{(i)T} \boldsymbol{X}_p^{(j)}.$$

The mean and variance of $T_{n,p}$ are given by

$$E(T_{n,p}) = (\mathbf{1}_n^T \boldsymbol{M}_n \mathbf{1}_n) \boldsymbol{\mu}_p^T \boldsymbol{\mu}_p, \tag{2.3}$$

$$\mathrm{Var}(T_{M_n}) = 2\mathrm{tr} \boldsymbol{M}_n^2 \mathrm{tr} \boldsymbol{\Sigma}_p^2 + 4(\mathbf{1}_n^T \boldsymbol{M}_n^2 \mathbf{1}_n) \boldsymbol{\mu}_p^T \boldsymbol{\Sigma}_p \boldsymbol{\mu}_p. \tag{2.4}$$

The derivation of (2.4) is a little complicated, while the equation (2.3) is easily obtained. We shall give the derivation of (2.4) in Section 4. Since $\mathbf{1}_n^T \boldsymbol{M}_n \mathbf{1}_n \neq 0$ under Condition B, we note that $T_{n,p}$ is a valid statistic for testing the null hypothesis in (1.1). In view of $\mathrm{tr}\boldsymbol{M}_n^2 \neq 0$, we also note that $\mathrm{Var}(T_{n,p}) \neq 0$ for any population mean vector $\boldsymbol{\mu}_p$.

For the asymptotic normality of $T_{n,p}$ under the null hypothesis H_0 in (1.1), we have the following theorem.

Theorem 2.1 *Under the null hypothesis H_0 and Conditions A and B, we have*

$$\tilde{T}_{n,p} = \frac{T_{n,p}}{\sqrt{2\mathrm{tr}\boldsymbol{M}_n^2 \mathrm{tr}\boldsymbol{\Sigma}_p^2}} \xrightarrow{d} N(0,1), \quad p \to \infty,$$

where $\xrightarrow{d}$ denotes the convergence in distribution and $N(0,1)$ denotes a standard normal distribution.

Since $\mathrm{tr}\boldsymbol{\Sigma}_p^2$ is unknown generally, we need to estimate it. Define

$$U_2 = \frac{1}{P_n^2} \sum_{i \neq j}^{n} \boldsymbol{X}_p^{(i)T} \boldsymbol{X}_p^{(j)} \boldsymbol{X}_p^{(i)T} \boldsymbol{X}_p^{(j)},$$

$$U_3 = \frac{1}{P_n^3} \sum_{i \neq j \neq k}^{n} \boldsymbol{X}_p^{(i)T} \boldsymbol{X}_p^{(j)} \boldsymbol{X}_p^{(i)T} \boldsymbol{X}_p^{(k)},$$

$$U_4 = \frac{1}{P_n^4} \sum_{i \neq j \neq k \neq \ell}^{n} \boldsymbol{X}_p^{(i)T} \boldsymbol{X}_p^{(j)} \boldsymbol{X}_p^{(k)T} \boldsymbol{X}_p^{(\ell)},$$

where $P_n^r = n!/(n-r)!$. Following Chen et al. (2010), an unbiased estimator $\widehat{\mathrm{tr}\boldsymbol{\Sigma}_p^2}$ of $\mathrm{tr}\boldsymbol{\Sigma}_p^2$ is given by

$$\widehat{\mathrm{tr}\boldsymbol{\Sigma}_p^2} = U_2 - 2U_3 + U_4. \tag{2.5}$$

In addition to the unbiasedness, the estimator $\widehat{\mathrm{tr}\boldsymbol{\Sigma}_p^2}$ is ratio-consistent which is given in the following theorem:

Theorem 2.2 *Under Condition A, we have*

$$\frac{\widehat{\mathrm{tr}\boldsymbol{\Sigma}_p^2}}{\mathrm{tr}\boldsymbol{\Sigma}_p^2} \xrightarrow{P} 1, \quad p \to \infty, \tag{2.6}$$

where $\xrightarrow{P} 1$ denotes the convergence in probability.

From Theorem 2.1 and 2.2, under H_0 and Conditions A and B, we have

$$\frac{T_{n,p}}{\sqrt{2\mathrm{tr}\boldsymbol{M}_n^2\widehat{\mathrm{tr}\boldsymbol{\Sigma}_p^2}}} \xrightarrow{d} N(0,1), \quad p \to \infty.$$

This result does not require any conditions on the population covariance matrix. The asymptotic normality is attained by using the known matrix $\boldsymbol{M}_n$ which satisfies Condition B. For instance, the matrix $\boldsymbol{M}_n$ is given by

(i) $\boldsymbol{M}_n = (\rho^{|i-j|}I(|i-j|>0))$ with $\rho \in (0,1)$;

(ii) $\boldsymbol{M}_n = (|i-j|^{-s}I(|i-j|>0))$ with $s>1$;

(iii) $\boldsymbol{M}_n = (m_{|i-j|}I(|i-j|>0))$ with $\lim_{n\to\infty}\sum_{k=1}^{n-1}|m_k| < \infty$,

where $I(E)$ for a event E equals to 1 when E is true, equals to 0 when E is false. The matrices (i) and (ii) are the special cases of (iii). We apply the Szegö theorem (see, e.g., Grenander and Szegö, 1958) to show that (iii) meets Condition B.

For the asymptotic power of $\tilde{T}_{n,p}$, we assume the following condition:

Condition C. $(\mathbf{1}_n^T\boldsymbol{M}_n^2\mathbf{1}_n)\boldsymbol{\mu}_p^T\boldsymbol{\Sigma}_p\boldsymbol{\mu}_p = o(\mathrm{tr}\boldsymbol{M}_n^2\mathrm{tr}\boldsymbol{\Sigma}_p^2)$.

The following theorem establishes the asymptotic power of $\tilde{T}_{n,p}$ under the local alternative defined by Condition C.

Theorem 2.3 *Under the alternative H_1 and Conditions A–C, we have*

$$P\left(\tilde{T}_{n,p} > z_{1-\alpha}|H_1\right) - \Phi\left(-z_{1-\alpha} + \frac{(\mathbf{1}_n^T\boldsymbol{M}_n\mathbf{1}_n)\boldsymbol{\mu}_p^T\boldsymbol{\mu}_p}{\sqrt{2\mathrm{tr}\boldsymbol{M}_n^2\mathrm{tr}\boldsymbol{\Sigma}_p^2}}\right) \to 0, \quad p \to \infty,$$

where Φ denotes the distribution function of a standard normal distribution, $z_{1-\alpha}$ is the upper α-percentile point of the standard normal distribution, and $P(\cdot\,|H_1)$ denotes the probability under the alternative H_1.

From Theorem 2.3, we note that the asymptotic power of $\tilde{T}_{n,p}$ depends on $\lambda_n = \mathbf{1}_n^T\boldsymbol{M}_n\mathbf{1}_n/(\mathrm{tr}\boldsymbol{M}_n^2)^{1/2}$. Hence, we suppose that selecting the matrix $\boldsymbol{M}_n$ which has larger λ_n leads to larger power of $\tilde{T}_{n,p}$. Obviously, the matrix $\boldsymbol{M}_n$ which maximizes λ_n is given by $c\mathbf{1}_n\mathbf{1}_n^T$ where c is any constant. However, this matrix does not satisfy $m_{ii}=0$ and $\mathrm{tr}\boldsymbol{M}_n^4 = o\{(\mathrm{tr}\boldsymbol{M}_n^2)^2\}$ in Condition B. We need to maximize λ_n over the matrices which satisfy Condition B. Unfortunately, it is hard to find such matrix even in the case (iii) since the

numbers of the parameters, we need to determine to maximize λ_n, increase as the sample size increases. For this reason, we shall compare $\tilde{T}_{n,p}$ using (i) and (ii) numerically later in the next Section.

It should be mentioned that when we use the matrix $\boldsymbol{M}_n$ which maximizes λ_n subject only to $m_{ii} = 0$ in Condition B, the proposed statistic $T_{n,p}$ is equivalent to the statistic $G_{n,p}$ given in (1.3), and hence, to $F_{n,p}$ given in (1.2). Thus we note that the power of the proposed test would be lower than the two tests. In this view, we should use the existing test when $\boldsymbol{\Sigma}_p$ is known and the asymptotic distribution is specified using the result of Katayama et.al. (2010). Nevertheless, the proposed test is useful when we have no information about $\boldsymbol{\Sigma}_p$ for the correct type I error.

In the rest of this section, we consider the consistency of the proposed test. Assume that the matrix $\boldsymbol{M}_n$ is given by (iii). Then we have

$$\mathbf{1}_n^T \boldsymbol{M}_n \mathbf{1}_n = 2\sum_{k=1}^{n-1}(n-k)m_k = O(n)$$

and

$$\mathrm{tr}\boldsymbol{M}_n^2 = 2\sum_{k=1}^{n-1}(n-k)m_k^2 = O(n),$$

since $\lim_{n\to\infty}\sum_{k=1}^{n-1}|m_k| < \infty$ and $\sum_{k=1}^{n-1} m_k^2 \leq (\sum_{k=1}^{n-1}|m_k|)^2$. Hence,

$$\frac{\mathbf{1}_n^T \boldsymbol{M}_n \mathbf{1}_n}{\sqrt{\mathrm{tr}\boldsymbol{M}_n^2}} = O(\sqrt{n}). \tag{2.7}$$

Theorem 2.3 and (2.7) yield the following theorem:

Theorem 2.4 *Assume Conditions A–C and that $\boldsymbol{M}_n$ has the form given by (iii). Let $\delta_{n,p} = n^{1/2}\boldsymbol{\mu}_p^T\boldsymbol{\mu}_p/(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^{1/2}$. Under the alternative H_1, if $\delta_{n,p} \to \infty$ as $p \to \infty$, then we have $P(\tilde{T}_{n,p} > z_{1-\alpha}\,|H_1) \to 1$ where $P(\cdot|H_1)$ denotes the probability under the alternative H_1.*

3 Numerical Results

In this section, we evaluate performance of the proposed test statistic $\tilde{T}_{n,p}$ with the existing one $\tilde{T}_{CQ}$ by computing the empirical significance level and power. The performance for a DNA microarray data is also evaluated.

3.1 Computational Complexity of $\widehat{\mathrm{tr}\boldsymbol{\Sigma}_p^2}$

Before conducting several simulations and a DNA microarray data analysis, we need to mention the computational complexity of $\widehat{\mathrm{tr}\boldsymbol{\Sigma}_p^2}$. Indeed, a large amounts of time is required to calculate $\widehat{\mathrm{tr}\boldsymbol{\Sigma}_p^2}$ since it includes the summation over the set of up to quadruplet, i.e., $\{(i,j,k,\ell)\ |1 \le i,j,k,\ell \le n\}$. Let $a_{ij} = \boldsymbol{X}_p^{(i)\,T}\boldsymbol{X}_p^{(j)}$ for the sake of simplicity. Then it follows that

$$\begin{aligned} P_n^3 U_3 &= \sum_{i,j,k=1}^{n} a_{ij}a_{ik} - \sum_{i\neq j}^{n} a_{ij}^2 - 2\sum_{i\neq j} a_{ii}a_{ij} - \sum_{i=1}^{n} a_{ii}^2 \\ &= \sum_{i=1}^{n}\left\{\left(\sum_{j=1}^{n} a_{ij}\right)^2\right\} - \sum_{i,j=1}^{n} a_{ij}^2 - 2\sum_{i=1}^{n} a_{ii}\left(\sum_{j=1}^{n} a_{ij}\right) + 2\sum_{i=1}^{n} a_{ii}^2 \quad (3.1) \end{aligned}$$

and

$$\begin{aligned} P_n^4 U_4 &= \sum_{i,j,k,\ell=1}^{n} a_{ij}a_{k\ell} - 2\sum_{i\neq j\neq k}^{n} a_{ii}a_{jk} - 4\sum_{i\neq j\neq k}^{n} a_{ij}a_{ik} \\ &\quad - \sum_{i\neq j}^{n} a_{ii}a_{jj} - 2\sum_{i\neq j}^{n} a_{ij}^2 - 4\sum_{i\neq j}^{n} a_{ii}a_{ij} - \sum_{i=1}^{n} a_{ii}^2 \\ &= \left(\sum_{i,j=1}^{n} a_{ij}\right)^2 - 2\left(\sum_{i=1}^{n} a_{ii}\right)\left(\sum_{j,k=1}^{n} a_{jk}\right) - 4\sum_{i=1}^{n}\left\{\left(\sum_{j=1}^{n} a_{ij}\right)^2\right\} \\ &\quad + \left(\sum_{i=1}^{n} a_{ii}\right)^2 + 2\sum_{i,j=1}^{n} a_{ij}^2 + 8\sum_{i=1}^{n} a_{ii}\left(\sum_{j=1}^{n} a_{ij}\right) - 6\sum_{i=1}^{n} a_{ii}^2. \quad (3.2) \end{aligned}$$

Thus, when a_{ij} $(i,j = 1,\ldots,p)$ are given, we only require $O(n^2)$ operations to calculate $\widehat{\mathrm{tr}\boldsymbol{\Sigma}_p^2}$ through the above equations (3.1) and (3.2), while it costs as many as $O(n^4)$ operations to calculate it directly.

3.2 Simulations

We consider three population distributions in the model (2.1): random variables z_{ij}'s are i.i.d. as standard normal distribution, standardized t distribution with 5 degrees of freedom and standardized gamma distribution with shape and scale parameters 2 and 2. For the population covariance matrix,

we choose

$$\mathbf{\Sigma}_1 = (\sigma_i\sigma_j\rho^{|i-j|}), \quad \mathbf{\Sigma}_2 = (\sigma_i\sigma_j\rho), \quad \mathbf{\Sigma}_3 = (\sigma_i\sigma_j u_{ij})$$

where $\rho = 0.5$, $\sigma_i^2 = 2 + (p - i + 1)/p$, $u_{ii} = 1$ and $u_{ij} = u_{ji}$ $(i \neq j)$ are i.i.d. as $U(0, 1)$, uniform distribution with the support $(0, 1)$. We note that $\tilde{T}_{n,p}$ and $\tilde{T}_{CQ}$ include the unknown parameter $\mathrm{tr}\mathbf{\Sigma}_p^2$. The estimator (2.5) is used in the following simulations. The empirical significance level and power are calculated based on $10,000$ Monte Carlo simulations.

			New Test (1)		New Test (2)		CQ	
Covariance	Population	p	.0500	.0100	.0500	.0100	.0500	.0100
$\mathbf{\Sigma}_1$	Normal	100	.0502	.0122	.0500	.0124	.0629	.0201
		300	.0462	.0097	.0466	.0092	.0594	.0180
		500	.0508	.0101	.0527	.0098	.0566	.0145
	t	100	.0537	.0133	.0535	.0130	.0602	.0185
		300	.0489	.0101	.0492	.0099	.0564	.0172
		500	.0500	.0114	.0510	.0109	.0548	.0156
	Gamma	100	.0491	.0107	.0487	.0108	.0612	.0193
		300	.0521	.0129	.0526	.0121	.0598	.0159
		500	.0539	.0111	.0557	.0109	.0513	.0143
$\mathbf{\Sigma}_2$	Normal	100	.0576	.0141	.0567	.0144	.0726	.0418
		300	.0614	.0176	.0612	.0174	.0724	.0430
		500	.0596	.0163	.0592	.0166	.0716	.0414
	t	100	.0553	.0169	.0560	.0171	.0747	.0424
		300	.0557	.0148	.0561	.0145	.0729	.0443
		500	.0554	.0153	.0554	.0153	.0722	.0441
	Gamma	100	.0603	.0155	.0593	.0146	.0736	.0449
		300	.0563	.0159	.0568	.0153	.0722	.0406
		500	.0564	.0154	.0563	.0152	.0681	.0391
$\mathbf{\Sigma}_3$	Normal	100	.0530	.0143	.0510	.0139	.0616	.0316
		300	.0501	.0129	.0487	.0125	.0663	.0344
		500	.0568	.0129	.0554	.0131	.0647	.0329
	t	100	.0606	.0159	.0622	.0157	.0707	.0363
		300	.0527	.0141	.0524	.0145	.0646	.0322
		500	.0553	.0142	.0548	.0144	.0646	.0336
	Gamma	100	.0551	.0150	.0542	.0152	.0663	.0348
		300	.0503	.0133	.0499	.0132	.0635	.0312
		500	.0551	.0131	.0548	.0134	.0609	.0314

Table 1: Empirical significance levels for $\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_2$ and $\mathbf{\Sigma}_3$ at 5% and 1% significance based on the normal approximation.

In Table 1, we evaluate empirical significance levels of the three statistics based on the normal approximation at 5% and 1% significance. We choose $n = 80$ and $p = 100, 300, 500$. In the table, New Test (1) and (2) denote the proposed test using (i) and (ii) as $\boldsymbol{M}_n$ respectively. Considering $\mathrm{tr}\boldsymbol{M}_n^4 = o\{(\mathrm{tr}\boldsymbol{M}_n^2)^2\}$ in Condition B for the asymptotic normality of $T_{n,p}$, we determine ρ and s in (i) and (ii) such that $\mathrm{tr}\boldsymbol{M}_n^4/(\mathrm{tr}\boldsymbol{M}_n^2)^2 = 0.03$. CQ denotes the test proposed by Chen and Qin (2010). It is found that the proposed test has better approximation for the significance level, particularly at 1%, for the all populations and population covariance matrices considered.

We denote $\eta = \boldsymbol{\mu}_p^T\boldsymbol{\mu}_p/(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^{1/2}$ for comparing the empirical powers of the three tests. This simulation is conducted at significance 5%. In Figure 2, we show the empirical power curves for $n = 80$, $p = 300$, $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$ and $\boldsymbol{\Sigma}_3$ with standard normal, t and gamma distributions. It is found that the power of the proposed test which uses (ii) as $\boldsymbol{M}_n$ is slightly higher than that using (i). As described in Section 2, the power of the proposed test is lower than that of Chen and Qin's test, theoretically. We also find this result in the figure. For improvement of the power of the proposed test, we may need to find a optimal matrix $\boldsymbol{M}_n$ in the matrices which satisfy Condition B. This is a future problem.

3.3 A Real Data Example

We evaluate performance of the proposed test using the prostate cancer data (Singh et al., 2002). This data contain 12,600 genes for 102 patients, 52 of which are prostate tumor patients and 50 of which are prostate normal patients. Following Dudoit et al. (2002), we preprocess the data by first truncating on the closed set $[1, 5000]$, second removing genes with $max/min \leq 5$ or $max - min \leq 150$ where max and min denote the maximum and minimum intensities over 102 patients, and finally transforming all intensities to base-10 logarithms. The number of genes is reduced to $p = 2,745$. We use $n = 50$ observations from each of the tumor and normal patients. Then this setup leads to the situation of (1.1).

In Figure 3, we plot the distributions of $T_{n,p}$ and T_{CQ} under H_0 using the parametric bootstrap method, that is, we draw 50 samples from $N_p(\mathbf{0}_p, \boldsymbol{S}_{n,p})$ ten thousand times, and then plot them each time. For the proposed test statistic $T_{n,p}$, the matrix $\boldsymbol{M}_n$ is given by (ii) of which parameter s is determined such that $\mathrm{tr}\boldsymbol{M}_n^4/(\mathrm{tr}\boldsymbol{M}_n^2)^2 = 0.05$. This shows that the distribution of the proposed test statistic is much closer to the standard normal distribution

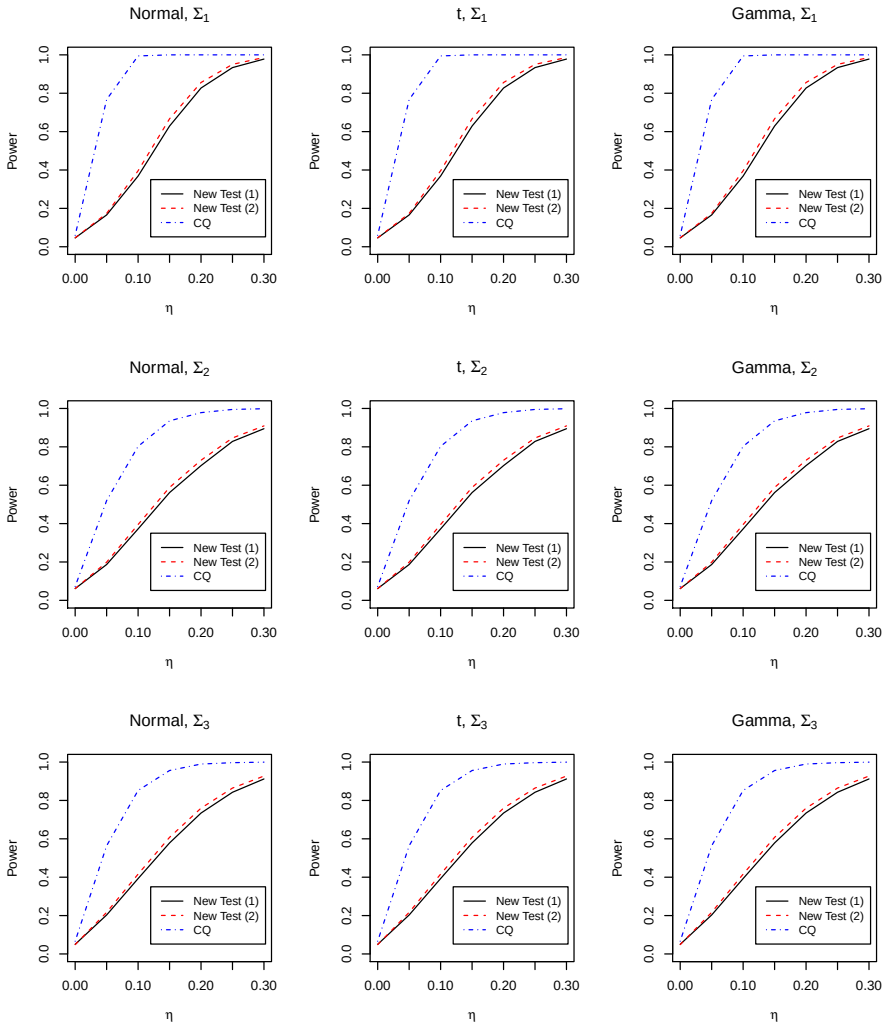

Figure 2: Power curves of the proposed test and Chen and Qin's test for $n = 80$, $p = 300$, $\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_2$ and $\mathbf{\Sigma}_3$ with normal, t and gamma distributions.

than the existing one.

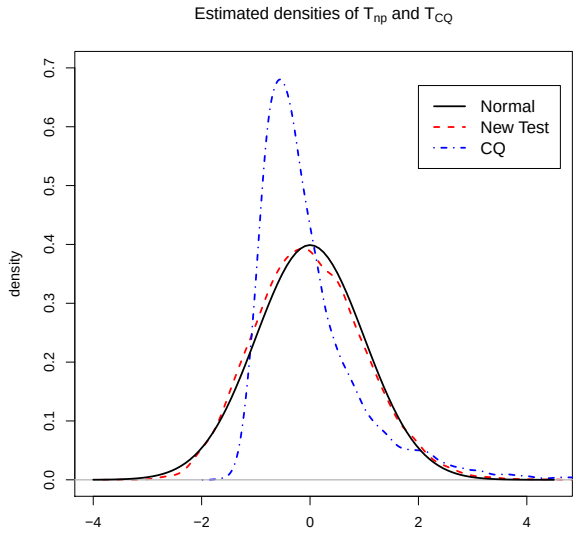

Figure 3: Distributions of $T_{n,p}$ and T_{CQ} under H_0 based on the parametric bootstrap method

4 Proofs

In this section, we give proofs of the equation (2.4) and Theorem 2.1–2.3 in Section 2. Throughout this section, we describe $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$ for $\boldsymbol{X}_p^{(i)}$ and $\boldsymbol{Z}_p^{(i)}$ respectively.

4.1 Proof of the equation (2.4)

Note that

$$\begin{aligned} T_{n,p} &= \sum_{i\neq j}^{n} m_{ij}(\boldsymbol{X}_i-\boldsymbol{\mu}_p)^T(\boldsymbol{X}_j-\boldsymbol{\mu}_p) + 2\sum_{i\neq j}^{n} m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{X}_j - \sum_{i\neq j}^{n} m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{\mu}_p \\ &= \sum_{i\neq j}^{n} m_{ij}\boldsymbol{Z}_i^T\boldsymbol{C}_p^T\boldsymbol{C}_p\boldsymbol{Z}_j + 2\sum_{i\neq j}^{n} m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{C}_p\boldsymbol{Z}_j + \sum_{i\neq j}^{n} m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{\mu}_p. \end{aligned}$$

Hence, we have

$$\begin{aligned} \mathrm{Var}(T_{n,p}) &= \mathrm{Var}\left(\sum_{i\neq j}^{n} m_{ij}\boldsymbol{Z}_i^T\boldsymbol{C}_p^T\boldsymbol{C}_p\boldsymbol{Z}_j\right) + 4\mathrm{Var}\left(\sum_{i\neq j}^{n} m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{C}_p\boldsymbol{Z}_j\right) \\ &\quad + 4\sum_{i\neq j}^{n}\sum_{k\neq \ell}^{n} m_{ij}m_{k\ell}\,\mathrm{Cov}\left(\boldsymbol{Z}_i^T\boldsymbol{C}_p^T\boldsymbol{C}_p\boldsymbol{Z}_j, \boldsymbol{\mu}_p^T\boldsymbol{C}_p\boldsymbol{Z}_\ell\right) \\ &= V_1 + 4V_2 + 4V_3, \quad \text{say.} \end{aligned}$$

Clearly, $V_3 = 0$ since $\boldsymbol{Z}_i$'s are independent and $E(\boldsymbol{Z}_i) = \boldsymbol{0}_p$. For the first term V_1, using the vec-operator $\mathrm{vec}(\cdot)$ and the Kronecker product $\otimes$ (see, e.g., Schott, 2005), it follows that

$$\sum_{i\neq j}^{n} m_{ij}\boldsymbol{Z}_i^T\boldsymbol{C}_p^T\boldsymbol{C}_p\boldsymbol{Z}_j = \mathrm{vec}(\boldsymbol{Z}_{n,p})^T(\boldsymbol{C}_p^T\boldsymbol{C}_p\otimes\boldsymbol{M}_n)\mathrm{vec}(\boldsymbol{Z}_{n,p}),$$

where $\boldsymbol{Z}_{n,p} = (\boldsymbol{Z}_1,\ldots,\boldsymbol{Z}_n)^T$. Then, from Lemma 2.1 in Srivastava (2009),

$$V_1 = 2\mathrm{tr}(\boldsymbol{C}_p^T\boldsymbol{C}_p\otimes\boldsymbol{M}_n)^2 = 2\mathrm{tr}\boldsymbol{M}_n^2\mathrm{tr}\boldsymbol{\Sigma}_p^2.$$

For the second term V_2, let $\boldsymbol{\Theta} = \boldsymbol{1}_p\boldsymbol{\mu}_p^T$. Note that

$$\sum_{i\neq j}^{n} m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{C}_p\boldsymbol{Z}_j = \mathrm{vec}(\boldsymbol{\Theta})^T(\boldsymbol{C}_p\otimes\boldsymbol{M}_n)\mathrm{vec}(\boldsymbol{Z}_{n,p}).$$

Then we obtain

$$\begin{aligned} V_2 &= \mathrm{vec}(\boldsymbol{\Theta})^T(\boldsymbol{C}_p\otimes\boldsymbol{M}_n)(\boldsymbol{C}_p\otimes\boldsymbol{M}_n)^T\mathrm{vec}(\boldsymbol{\Theta}) \\ &= (\boldsymbol{\mu}_p\otimes\boldsymbol{1}_p)(\boldsymbol{\Sigma}_p\otimes\boldsymbol{M}_n^2)(\boldsymbol{\mu}_p\otimes\boldsymbol{1}_p) \\ &= (\boldsymbol{1}_p^T\boldsymbol{M}_n^2\boldsymbol{1}_p)\boldsymbol{\mu}_p^T\boldsymbol{\Sigma}_p\boldsymbol{\mu}_p. \end{aligned}$$

Thus, the equation (2.4) is derived.

4.2 Proof of Theorem 2.1

This proof is based on the results in Srivastava (2009) and Chen and Qin (2010). Let $\tilde{m}_{ij} = m_{ij}/(2\mathrm{tr}\boldsymbol{M}_n^2\mathrm{tr}\boldsymbol{\Sigma}_p^2)^{1/2}$ and $\tilde{\boldsymbol{M}}_n = (\tilde{m}_{ij})$. Then we have

$$\tilde{T}_{n,p} = \frac{\sum_{i\neq j}^n m_{ij}\boldsymbol{X}_i^T\boldsymbol{X}_j}{\sqrt{2\mathrm{tr}\boldsymbol{M}_n^2\mathrm{tr}\boldsymbol{\Sigma}_p^2}} = \sum_{i\neq j}^n \tilde{m}_{ij}\boldsymbol{X}_i^T\boldsymbol{X}_j = 2\sum_{j=2}^n\sum_{i=1}^{j-1}\tilde{m}_{ij}\boldsymbol{X}_i^T\boldsymbol{X}_j = 2\sum_{j=2}^n Y_j,$$

where $Y_j = \sum_{i=1}^{j-1}\tilde{m}_{ij}\boldsymbol{X}_i^T\boldsymbol{X}_j$. We need to show that

$$Q_n = \sum_{j=2}^n Y_j \xrightarrow{d} N(0, 1/4), \quad p \to \infty. \tag{4.1}$$

We can apply Martingale Central Limit Theorem (Hall and Heyde, 1980) for Q_n to show (4.1). Let $\mathcal{F}_n$ be the σ algebra which is generated by $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n\}$. Then we have $\mathcal{F}_{j-1} \subset \mathcal{F}_j$ for any $j = 2, \ldots, n$. It is easily obtained that $E(Q_n) = 0$ and $E(Q_n^2) < \infty$. We note for $n > m$ that $Q_n = Q_m + \sum_{j=m+1}^n Y_j$ and $E(\sum_{j=m+1}^n Y_j|\mathcal{F}_m) = 0$ since when $m+1 \leq j \leq n$, $E(\boldsymbol{X}_i^T\boldsymbol{X}_j|\mathcal{F}_m) = \boldsymbol{X}_i^T E(\boldsymbol{X}_j) = 0$ for $1 \leq i \leq m$ and $E(\boldsymbol{X}_i^T\boldsymbol{X}_j|\mathcal{F}_m) = E(\boldsymbol{X}_i^T\boldsymbol{X}_j) = 0$ for $m+1 \leq i \leq n-1$. Hence we find that the sequence $\{Q_n, \mathcal{F}_n\}$ is a zero mean and square integrable martingale array. It suffices to show that under Conditions A and B

$$\sum_{j=2}^n E(Y_j^2|\mathcal{F}_{j-1}) \xrightarrow{P} \frac{1}{4}, \quad p \to \infty \tag{4.2}$$

$$\sum_{j=2}^n E(Y_j^4) \to 0, \quad p \to \infty. \tag{4.3}$$

We first show (4.2). Note that

$$\begin{aligned}
\sum_{j=2}^n E(Y_j^2|\mathcal{F}_{j-1}) &= \sum_{j=2}^n E\left[\left(\sum_{i=1}^{j-1}\tilde{m}_{ij}\boldsymbol{X}_i^T\boldsymbol{X}_j\right)^2\middle|\mathcal{F}_{j-1}\right] \\
&= \sum_{i<j}^n \tilde{m}_{ij}^2\boldsymbol{X}_i^T\boldsymbol{\Sigma}_p\boldsymbol{X}_i + \sum_{j=2}^n\sum_{i_1\neq i_2}^{j-1}\tilde{m}_{i_1j}\tilde{m}_{i_2j}\boldsymbol{X}_{i_1}^T\boldsymbol{\Sigma}_p\boldsymbol{X}_{i_2} \\
&= A_1 + A_2, \quad \text{say.}
\end{aligned}$$

Also we note that

$$E(A_1) = \sum_{i<j}^{n} \tilde{m}_{ij}^2 E(\boldsymbol{X}_i^T \boldsymbol{\Sigma}_p \boldsymbol{X}_i) = \sum_{i<j}^{n} \tilde{m}_{ij}^2 \mathrm{tr}\boldsymbol{\Sigma}_p^2 = \frac{1}{2}\mathrm{tr}\tilde{\boldsymbol{M}}_n^2 \mathrm{tr}\boldsymbol{\Sigma}_p^2 = \frac{1}{4},$$

and we have

$$\begin{aligned}
\mathrm{Var}(A_1) &= \mathrm{Var}\left(\sum_{i<j}^{n} \tilde{m}_{ij}^2 \boldsymbol{X}_i^T \boldsymbol{\Sigma}_p \boldsymbol{X}_i\right) \\
&= \sum_{i=1}^{n-1} \left(\sum_{j=i+1}^{n} \tilde{m}_{ij}\right)^2 \mathrm{Var}(\boldsymbol{X}_i^T \boldsymbol{\Sigma}_p \boldsymbol{X}_i) \\
&= \sum_{i=1}^{n-1} \left(\sum_{j=i+1}^{n} \tilde{m}_{ij}\right)^2 \left[(\gamma - 3)\sum_{k=1}^{p} v_{kk}^2 + \mathrm{tr}\boldsymbol{\Sigma}_p^4\right],
\end{aligned}$$

where v_{ii} is the diagonal elements of $\boldsymbol{C}_p^T \boldsymbol{\Sigma}_p \boldsymbol{C}_p = (\boldsymbol{C}_p^T \boldsymbol{C}_p)(\boldsymbol{C}_p^T \boldsymbol{C}_p)$. The last equation is derived from Lemma 2.1 in Srivastava (2009). Let $\boldsymbol{D}_p = \boldsymbol{C}_p^T \boldsymbol{C}_p = (d_{ij})$. Following Lemma 2.6 in Srivastava (2009), we note that

$$\sum_{k=1}^{p} v_{kk}^2 = \sum_{k=1}^{p} \left(\sum_{i=1}^{p} d_{ik}^2\right)^2 \leq \mathrm{tr}(\boldsymbol{D}_p^T \boldsymbol{D}_p)^2 = \mathrm{tr}\boldsymbol{\Sigma}_p^4. \tag{4.4}$$

Also we note that under Condition B

$$\sum_{i=1}^{n-1} \left(\sum_{j=i+1}^{n} \tilde{m}_{ij}^2\right)^2 \leq \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \tilde{m}_{ij}^2\right)^2 + \sum_{i_1 \neq i_2}^{n} \left(\sum_{j=1}^{n} \tilde{m}_{i_1 j} \tilde{m}_{i_2 j}\right)^2 = \mathrm{tr}\tilde{\boldsymbol{M}}_n^4.$$

Hence, we have

$$\mathrm{Var}(A_1) = O(\mathrm{tr}\tilde{\boldsymbol{M}}_n^4 \mathrm{tr}\boldsymbol{\Sigma}_p^4).$$

Since $\mathrm{tr}\boldsymbol{\Sigma}_p^4/(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2 \leq 1$, we obtain

$$\mathrm{tr}\tilde{\boldsymbol{M}}_n^4 \mathrm{tr}\boldsymbol{\Sigma}_p^4 = \frac{\mathrm{tr}\boldsymbol{M}_n^4 \mathrm{tr}\boldsymbol{\Sigma}_p^4}{4(\mathrm{tr}\boldsymbol{M}_n^2)^2 (\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2} \to 0, \quad p \to \infty \tag{4.5}$$

under Conditions A and B. Thus $A_1 \xrightarrow{P} 1/4$ from the Chebyshev's inequality. Next, we shall show $A_2 \xrightarrow{P} 0$. Since

$$A_2 = \sum_{i_1 \neq i_2}^{n-1} \boldsymbol{X}_{i_1}^T \boldsymbol{\Sigma}_p \boldsymbol{X}_{i_2} \left(\sum_{j=\min\{i_1,i_2\}+1}^{n} \tilde{m}_{i_1 j} \tilde{m}_{i_2 j} \right),$$

and $E(\boldsymbol{X}_{i_1}^T \boldsymbol{\Sigma}_p \boldsymbol{X}_{i_2} \boldsymbol{X}_{i_3}^T \boldsymbol{\Sigma}_p \boldsymbol{X}_{i_4}) = 0$ when at least one index is different between (i_1, i_2) and (i_3, i_4), we have $E(A_2) = 0$ and

$$E(A_2^2) = \sum_{i_1 \neq i_2}^{n} E(\boldsymbol{X}_{i_1}^T \boldsymbol{\Sigma}_p \boldsymbol{X}_{i_2})^2 \left(\sum_{j=\min\{i_1,i_2\}+1} \tilde{m}_{i_1 j} \tilde{m}_{i_2 j} \right)^2.$$

Note that $E(\boldsymbol{X}_{i_1}^T \boldsymbol{\Sigma}_p \boldsymbol{X}_{i_2})^2 = \mathrm{tr}\boldsymbol{\Sigma}_p^4$ for $i_1 \neq i_2$ and

$$\begin{aligned} \sum_{i_1 \neq i_2}^{n} \left(\sum_{j=\min\{i_1,i_2\}+1}^{n} \tilde{m}_{i_1 j} \tilde{m}_{i_2 j} \right)^2 &\leq \sum_{i_1 \neq i_2}^{n} \left(\sum_{j=1}^{n} \tilde{m}_{i_1 j} \tilde{m}_{i_2 j} \right)^2 + \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \tilde{m}_{ij}^2 \right)^2 \\ &= \mathrm{tr}\tilde{\boldsymbol{M}}_n^4, \end{aligned}$$

under Condition B. Then we have $E(A_2^2) = O(\mathrm{tr}\tilde{\boldsymbol{M}}_n^4 \mathrm{tr}\boldsymbol{\Sigma}_p^4)$. From (4.5), we obtain that $E(A_2^2) \to 0$ as $p \to \infty$ under Conditions A and B. The Chebyshev's inequality yields $A_2 \xrightarrow{P} 0$.

Next, we verify (4.3). Note that

$$\begin{aligned} E(Y_j^4) &= E\left(\sum_{i=1}^{j-1} \tilde{m}_{ij} \boldsymbol{X}_i^T \boldsymbol{X}_j \right)^4 \\ &= E\left[\sum_{i=1}^{j-1} \tilde{m}_{ij}^2 (\boldsymbol{X}_i^T \boldsymbol{X}_j)^2 + \sum_{i_1 \neq i_2}^{j-1} \tilde{m}_{i_1 j} \tilde{m}_{i_2 j} \boldsymbol{X}_{i_1}^T \boldsymbol{X}_j \boldsymbol{X}_{i_2}^T \boldsymbol{X}_j \right]^2 \\ &= \sum_{i=1}^{j-1} \tilde{m}_{ij}^4 E(\boldsymbol{X}_i^T \boldsymbol{X}_j)^4 + 3 \sum_{i_1 \neq i_2}^{j-1} \tilde{m}_{i_1 j}^2 \tilde{m}_{i_2 j}^2 E[(\boldsymbol{X}_{i_1}^T \boldsymbol{X}_j)^2 (\boldsymbol{X}_{i_2}^T \boldsymbol{X}_j)^2]. \end{aligned}$$

Following Chen and Qin (2010), we have

$$E(\boldsymbol{X}_i^T \boldsymbol{X}_j)^4 = O(\mathrm{tr}\boldsymbol{\Sigma}_p^4) + O\{(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2\}, \quad i \neq j.$$

Also we have from (4.4),

$$\begin{aligned} E[(\boldsymbol{X}_{i_1}^T\boldsymbol{X}_j)^2(\boldsymbol{X}_{i_2}^T\boldsymbol{X}_j)^2] &= E[(\boldsymbol{X}_j^T\boldsymbol{\Sigma}_p\boldsymbol{X}_j)^2] \\ &= (\gamma-3)\sum_{k=1}^{p} v_{kk}^2 + 2\mathrm{tr}\boldsymbol{\Sigma}_p^4 + (\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2 \\ &= O(\mathrm{tr}\boldsymbol{\Sigma}_p^4) + O\{(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2\}, \quad i_1 \neq i_2 \neq j. \end{aligned}$$

Following Corollary 2.7 in Srivastava (2009), we have

$$\sum_{j=2}^{n}\sum_{i=1}^{j-1}\tilde{m}_{ij}^4 \leq \sum_{i\neq j}^{n}\tilde{m}_{ij}^4 \leq \mathrm{tr}\tilde{\boldsymbol{M}}_n^4$$

and

$$\sum_{j=2}^{n}\sum_{i_1\neq i_2}^{j-1}\tilde{m}_{i_1j}^2\tilde{m}_{i_2j}^2 \leq \sum_{j=1}^{n}\sum_{i_1\neq i_2}^{n}\tilde{m}_{i_1j}^2\tilde{m}_{i_2j}^2 \leq \sum_{i=1}^{n}\left(\sum_{j=1}^{n}\tilde{m}_{ij}^2\right)^2 \leq \mathrm{tr}\tilde{\boldsymbol{M}}_n^4.$$

From (4.5) and the fact that

$$\mathrm{tr}\tilde{\boldsymbol{M}}_n^4(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2 = \frac{\mathrm{tr}\boldsymbol{M}_n^4}{4(\mathrm{tr}\boldsymbol{M}_n^2)^2} \to 0, \quad p\to\infty$$

under Conditions A and B, we obtain $\sum_{j=2}^{n}E(Y_j^4)\to 0,\ p\to\infty$. This ends the proof.

4.3 Proof of Theorem 2.2

Note that $\widehat{\mathrm{tr}\boldsymbol{\Sigma}_p^2}$ is invariant under the location transformations $\boldsymbol{X}_i \to \boldsymbol{X}_i + \boldsymbol{c}_p$ where $\boldsymbol{c}_p$ is an arbitrary p-column constant vector. Hence, we assume $\boldsymbol{\mu}_p = \boldsymbol{0}$ without loss of generality. Obviously, we have $E(U_2) = \mathrm{tr}\boldsymbol{\Sigma}_p^2$ and $E(U_3) = E(U_4) = 0$. It follows from Chen et al. (2010) that

$$\begin{aligned} \mathrm{Var}(U_2) = \frac{4}{n^2}(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2 + \frac{8}{n}\mathrm{tr}\boldsymbol{\Sigma}_p^4 + \frac{4(\gamma-3)}{n}\mathrm{tr}[(\boldsymbol{C}_p^T\boldsymbol{C}_p)^2\odot(\boldsymbol{C}_p^T\boldsymbol{C}_p)^2] \\ + O\left[\frac{1}{n^3}(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2 + \frac{1}{n^2}\mathrm{tr}\boldsymbol{\Sigma}_p^4\right], \end{aligned}$$

$$\mathrm{Var}(U_3) = \frac{2}{n^3}(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2 + \frac{2}{n^2}\mathrm{tr}\boldsymbol{\Sigma}_p^4 + O\left[\frac{1}{n^4}(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2 + \frac{1}{n^3}\mathrm{tr}\boldsymbol{\Sigma}_p^4\right]$$

and

$$\mathrm{Var}(U_4) = \frac{8}{n^4}(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2 + O\left[\frac{1}{n^5}(\mathrm{tr}\boldsymbol{\Sigma}_p^2)^2 + \frac{1}{n^4}\mathrm{tr}\boldsymbol{\Sigma}_p^4\right],$$

where the symbol $\odot$ denotes the Hadamard product. Using the fact that

$$\mathrm{tr}[(\boldsymbol{C}_p^T\boldsymbol{C}_p)^2 \odot (\boldsymbol{C}_p^T\boldsymbol{C}_p)^2] \leq \mathrm{tr}(\boldsymbol{C}_p^T\boldsymbol{C}_p)^4 = \mathrm{tr}\boldsymbol{\Sigma}_p^4,$$

we have $\mathrm{Var}(U_i/\mathrm{tr}\boldsymbol{\Sigma}_p^2) \to 0$ for $i = 2, 3, 4$ under Condition A. This and the Chebychev's inequality yield (2.6).

4.4 Proof of Theorem 2.3

Using Theorem 2.2, we have

$$\frac{\sum_{i\neq j}^n m_{ij}(\boldsymbol{X}_i - \boldsymbol{\mu}_p)^T(\boldsymbol{X}_j - \boldsymbol{\mu}_p)}{\sigma_{n,p}} \xrightarrow{d} N(0,1), \quad p \to \infty,$$

under Conditions A and B. Note that

$$T_{n,p} = \sum_{i\neq j}^n m_{ij}(\boldsymbol{X}_i - \boldsymbol{\mu}_p)^T(\boldsymbol{X}_j - \boldsymbol{\mu}_p) + 2\sum_{i\neq j}^n m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{X}_j - \sum_{i\neq j}^n m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{\mu}_p.$$

Hence, it remains to show that

$$\frac{\sum_{i\neq j}^n m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{X}_j - \sum_{i\neq j}^n m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{\mu}_p}{\sigma_{n,p}} \xrightarrow{P} 0, \quad p \to \infty. \tag{4.6}$$

We notice $E(\sum_{i\neq j}^n m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{X}_j) = \sum_{i\neq j}^n m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{\mu}_p$ and

$$\mathrm{Var}\left(\sum_{i\neq j}^n m_{ij}\boldsymbol{\mu}_p^T\boldsymbol{X}_j\right) = 4(\mathbf{1}_n^T\boldsymbol{M}_n^2\mathbf{1}_n)\boldsymbol{\mu}_p^T\boldsymbol{\Sigma}_p\boldsymbol{\mu}_p = o(\sigma_{n,p}^2),$$

under Condition C. The Chebyshev's inequality yields (4.6). The Slutsky's theorem and (4.6) complete the proof.

References

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

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

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

[4] Dudoit, S., Fridlyand, J. and Speed, T. P. (2002). Comparison of discrimination methods for the classification of tumors using gene expression data. *Jounal of the American Statistical Association* 97: 77–87.

[5] Grenander, U. and Szegö, G. (1958). *Toeplitz forms and their applications, 2nd edition*. New York: Chelsea Publishing Company.

[6] Hall, P and Heyde, C. (1980). *Martingale limit theory and applications*. New York: Academic Press.

[7] Katayama, S., Kano, Y. and Srivastava, M.S. (2010). Asymptotic distributions of some test criteria for the mean vector with fewer observations than the dimension. *Submitted*.

[8] Schott, J. R. (2005). *Matrix analysis for statistics, 2nd edition*. New York: Wiley.

[9] Singh, D., Febbo, P. G., Ross, K., Jackson, D. G., Manola, J., Ladd, C., Tamayo, P., Renshaw, A. A., D'Amico, A. V., Richie, J. P., Lander, E. S., Loda, M., Kantoff, P. W., Golub, T. R. and Sellers, W. R. (2002). Gene expression correlates of clinical prostate cancer behavior. *Cancer Cell* 1: 203–209

[10] Srivastava, M. S. (2009). A test for the mean vector with fewer observations than the dimension under non-normality. *Journal of Multivariate Analysis*. 100: 518–532.