Statistica Sinica Preprint No: SS-2022-0048	
Title	Homogeneity Tests for High-dimensional Mean Vectors and Covariance Matrices
Manuscript ID	SS-2022-0048
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202022.0048
Complete List of Authors	Wenwen Guo, Xinyuan Song and Hengjian Cui
Corresponding Authors	Xinyuan Song
E-mails	xysong@sta.cuhk.edu.hk

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

Statistica Sinica

Homogeneity Tests for High-dimensional Mean Vectors and Covariance Matrices

Wenwen Guo[1], Xinyuan Song[2*] and Hengjian Cui[1]

Capital Normal University[1] *and The Chinese University of Hong Kong*[2]

Abstract: This study aims to develop homogeneity tests for high-dimensional mean vectors and covariance matrices, in which the number of features may be greater than the sample size. We introduce two categorically weighted statistics to test the equality of means and of covariance matrices. We establish the asymptotic distributions of the proposed test statistics under certain mild conditions, and develop simplified algorithms to facilitate the implementation and application. Simulation studies demonstrate the satisfactory performance of the proposed tests in terms of the empirical size and power. We also apply the proposed test procedures to two microarray data sets.

Key words and phrases: Homogeneity, K-sample problem, High-dimension, Location and scale, MANOVA.

1. Introduction

Despite numerous studies on homogeneity tests for distributions or distribution features (mean vectors or covariance matrices) in different populations,

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

a crucial remaining problem is establishing whether gene expression levels differ among predefined patient populations in order to identify a disease's capital causal gene. However, in modern biological and financial studies, the data dimension is often much larger than the sample size. This "large p, small n" paradigm poses a considerable challenge to classical homogeneity tests, which were originally designed for fixed-dimensional problems.

This study focuses on homogeneity tests for high-dimensional mean vectors and covariance matrices. Assume that homogeneity tests for means, consider R groups. When $R = 2$, the traditional Hotelling T^2 test is optimal for normally distributed data when p is fixed. Several extensions of the Hotelling T^2 test have been proposed to accommodate high dimensionality; examples include those of Bai and Saranadasa (1996), Srivastava and Du (2008), Chen and Qin (2010), Cai et al. (2013), Feng et al. (2016), and Chang et al. (2017). When $R > 2$, researchers often use a multivariate analysis of variance (MANOVA) to investigate whether the population mean vectors are the same under the "large n, small p" paradigm. Cai and Xia (2014) test the equality of multiple high-dimensional sparse mean vectors under dependency. Recently, Hu et al. (2017) proposed a test for the equality of high-dimensional mean vectors based on the work of Chen and Qin (2010).

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

Several studies also test covariances based mostly on entropy or a quadratic loss function. Studies that examine the case of $R = 2$ include Wolf (2002), Bai et al. (2009), Chen et al. (2010), Li and Chen (2012), Cai and Ma (2013), Jiang and Yang (2013), Cai and Liu (2016), and Chang et al. (2017). For $R > 2$, Zhang et al. (2018) extend the two-sample test for covariances presented by Li and Chen (2012) , and obtain the asymptotic distribution of the statistic in a high-dimension case. Zheng et al. (2020) propose a homogeneity test for high-dimensional covariances, and enhance its power by comparing covariance matrices. Liu et al. (2017) also propose a two-sample homogeneity test for means and covariances.

In this study, we consider this kind of homogeneity test from a different perspective. Assume that Y is a categorical variable with R categories, and $\mathbf{X}$ is a p-dimensional random vector. Cui et al. (2015) propose a mean-variance index defined by $MV(\mathbf{X}|Y) = \mathrm{E}_{\mathbf{X}}[\mathrm{var}_Y F(\mathbf{x}|Y)]$, where $F(\mathbf{x}|Y)$ stands for the conditional distribution function of $\mathbf{X}$ given Y. $MV(\mathbf{X}|Y)$ indicates that $\mathbf{X}$ and Y are independent if and only if the conditional distributions $F_r = F(\mathbf{x}|Y = r)$, for $r = 1, \cdots, R$, are homogenous. Then, the homogeneity test for distributions can be regarded as an independence test between a categorical variable and a multivariate random vector. The mean-variance index takes advantage of the probabilities of the categorical

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

variable, which motivates us to introduce a categorically weighted index to measure the differences between the mean vectors and covariance matrices of different groups.

To accommodate high dimensionality, we correct the bias by adjusting the weights, and propose two statistics for testing the mean vectors and covariance matrices. Moreover, we obtain the asymptotic distributions of the proposed statistics under certain mild conditions. The proposed tests have four advantages. (1) They accommodate the high-dimensional setting. (2) No explicit distribution is imposed on the p-dimensional vectors; hence, our tests have high theoretical and practical value. (3) The proposed categorically weighted tests optimize the information of the categorical variable to improve the performance. (4) Simplified algorithms are proposed to calculate the associated statistics, thereby facilitating implementation and application.

The remainder of this paper is organized as follows. Sections 2 and 3 describe the methodology and asymptotic distributions of testing means and covariances, respectively. Section 4 introduces simplified algorithms to calculate the test statistics. Section 5 presents Monte Carlo simulations for assessing the performance of the proposed tests. Applications to gene expression data analysis are given in Section 6. Technical proofs are provided

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

in the Supplementary Material.

2. Homogeneity Test for Mean Vectors

We consider the homogeneity test for mean vectors, that is,

$$H_{10} : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_R = \boldsymbol{\mu}, \tag{2.1}$$

versus the composite alternative H_{11}: $\boldsymbol{\mu}_r \neq \boldsymbol{\mu}_s$, for $1 \leq r < s \leq R$, where $\boldsymbol{\mu}_r = E(\mathbf{X}|Y = r)$, $\boldsymbol{\mu} = E(\mathbf{X}) = \sum_{r=1}^{R} p_r \boldsymbol{\mu}_r$, and p_r is the probability that $\mathbf{X}$ comes from the rth population.

2.1 Measuring the difference between mean vectors

Similarly to the analysis of Cui et al. (2015), we use the variance of the conditional means of $\mathbf{X}$ given Y, $\mathrm{var}_Y\{\mathrm{E}(\mathbf{X}|Y)\}$, to measure the difference between the mean vectors, as expressed in Definition 1.

Definition 1. The variance of the conditional expectations of $\mathbf{X}$ given $Y = r$, for $r = 1, \cdots, R$, can be defined by

$$\mathcal{U}(\mathbf{X}|Y) = \mathrm{E}(\mathbf{X}_1^{\mathrm{T}}\mathbf{X}_2)\left\{\sum_{r=1}^{R} \frac{I(Y_1 = r)I(Y_2 = r)}{p_r} - 1\right\},$$

where $(\mathbf{X}_1, Y_1)$ and $(\mathbf{X}_2, Y_2)$ are independent copies of $(\mathbf{X}, Y)$, and $I(\cdot)$ is the indicator function.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

The following lemma shows that Definition 1 is reasonable.

Lemma 2.1. If $\mathbf{X}$ has a finite first moment, then $\mathcal{U}(\mathbf{X}|Y) = \mathrm{var}_Y\{\mathrm{E}(\mathbf{X}|Y)\} \geq 0$, and the equality holds if and only if the null hypothesis (2.1) is true.

Section S1 of the Supplementary Material shows the proof of Lemma 2.1. For observed random samples $\{(\mathbf{X}_k, Y_k) : k = 1, 2, \cdots, n\}$, we define

$$M_{n,p} = \sum_{(i,j)}^{*} \mathbf{X}_i^{\mathrm{T}} \mathbf{X}_j \left\{ \sum_{r=1}^{R} \frac{I(Y_i = r) I(Y_j = r)}{\widehat{p}_r} - 1 \right\},$$

where $\sum_{(i,j)}^{*}$ denotes summations over distinct indices, and $\widehat{p}_r = (N_r - 1)/(n-1)$, with $N_r = \sum_{i=1}^{n} I(Y_i = r)$. Notably, $\widehat{p}_r$ is a consistent estimator of p_r, and more importantly, it enables $\sum_{i \neq j} c_i \left\{ \sum_{r=1}^{R} I(Y_i = r) I(Y_j = r)/\widehat{p}_r - 1 \right\} = 0$, where c_i is any function of the ith sample. The good properties of the estimator $\widehat{p}_r$ make our test applicable to high-dimensional data.

Remark 1. Using an elemetary calculation, we obtain

$$M_{n,p} = \sum_{r>s}^{R} N_r N_s \left\{ \frac{\sum_{i \neq j}^{N_r} \mathbf{X}_{ri}^{\mathrm{T}} \mathbf{X}_{rj}}{N_r(N_r - 1)} + \frac{\sum_{i \neq j}^{N_s} \mathbf{X}_{si}^{\mathrm{T}} \mathbf{X}_{sj}}{N_s(N_s - 1)} - 2 \frac{\sum_{i=1}^{N_r} \sum_{j=1}^{N_s} \mathbf{X}_{ri}^{\mathrm{T}} \mathbf{X}_{sj}}{N_r N_s} \right\}, \tag{2.2}$$

where $\mathbf{X}_{ri}$ denotes the ith sample of the rth group, that is, $Y_i = r$. We show the proof in Section S2 of the Supplementary Material. When $R = 2$, Equation (2.2) indicates that $M_{n,p}$ is proportional to the statistic proposed by Chen and Qin (2010), which they use to measure the distance between

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

two sample means, that is, $\|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\|^2$. Therefore, our proposed statistic can be regarded as a weighted summation of the distances between the means in two different categories.

2.2 Main results for the homogeneity test for means

To establish the limiting distribution of $M_{n,p}$, we assume the following conditions.

Condition 1. Suppose that R is fixed, and there exist two positive constants c_1 and c_2, such that $c_1/R \leq \min_{1\leq r\leq R} p_r \leq \max_{1\leq r\leq R} p_r \leq c_2/R$.

Condition 2. Suppose that the random expression of $\mathbf{X}_i$ given $Y_i = r$ is $\mathbf{X}_i|(Y_i = r) = \boldsymbol{\mu}_r + \boldsymbol{\Gamma}_r\mathbf{Z}_i$, where $\boldsymbol{\mu}_r$ is the conditional mean vector, $\boldsymbol{\Gamma}_r$ is a $p\times p$ matrix, $\mathbf{Z}_i$ is independent of Y_i, and the coordinates of $\mathbf{Z}_i$ are assumed to be independent and identically distributed (i.i.d.); the first coordinate, denoted as Z_{i1}, satisfies $E(Z_{i1}) = 0$, $E(Z_{i1}^2) = 1$ and $E(Z_{i1}^4) = 3+ \triangle< \infty$.

Condition 3. $p = p(n) \to \infty$ as $n \to \infty$; $\mathrm{tr}(\boldsymbol{\Sigma}_r\boldsymbol{\Sigma}_s\boldsymbol{\Sigma}_k\boldsymbol{\Sigma}_t) = o\{\mathrm{tr}^2(\boldsymbol{\Sigma}^2)\}$, for $r, s, k, t \in \{1, 2, \cdots, R\}$.

Condition 4. $(\boldsymbol{\mu}_r-\boldsymbol{\mu}_s)^{\mathrm{T}}\boldsymbol{\Sigma}_k(\boldsymbol{\mu}_r-\boldsymbol{\mu}_s) = o\{n^{-1}\mathrm{tr}(\boldsymbol{\Sigma}^2)\}$, for $r, s, k \in \{1, 2, \cdots, R\}$.

Condition 1 imposes that p_r, for $r = 1, 2, \cdots, R$, must not degenerate; a similar condition appears in the study of Cui et al. (2015). Instead of im-

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

posing a specific parametric distribution of $\mathbf{X}|Y$, the pseudo-independence assumption is required in Condition 2. The pseudo-independence model was proposed by Bai and Saranadasa (1996), and is widely used in high-dimensional theoretical models; see Chen and Qin (2010), Li and Chen (2012), and Zhang et al. (2018). The eigenvalues of the conditional variance of $(\mathbf{X}|Y)$ are assumed to satisfy Condition 3, which holds naturally when the conditional covariances are bounded away from above and zero. We explore the asymptotic properties of the statistic $M_{n,p}$ under high dimensionality and local alternatives in Condition 4. This work does not impose any explicit relationships between p and n, and our test applies to high-dimensional data.

Theorem 2.1. Under Conditions 1, 2, 3, and either H_{10} or Condition 4, we have

$$\frac{M_{n,p} - \sum_{r>s}^{R} N_r N_s \|\boldsymbol{\mu}_r - \boldsymbol{\mu}_s\|^2}{\sqrt{d_{n,p}}} \xrightarrow{d} \mathcal{N}(0,1)$$

as $n, p \longrightarrow \infty$, where $d_{n,p} = 2n(n-1)\{\sum_{r=1}^{R}(1-p_r)^2\mathrm{tr}(\boldsymbol{\Sigma}_r^2)+\sum_{(r,s)}^{*} p_r p_s \mathrm{tr}(\boldsymbol{\Sigma}_r\boldsymbol{\Sigma}_s)\}$, where $\xrightarrow{d}$ denotes convergence in distribution.

Theorem 2.1 establishes the asymptotic normality of $M_{n,p}$, without imposing explicit conditions on the relationship between n and p. Under Condition 3, $d_{n,p} = O(n^2 p)$. Furthermore, if the conditional covariances of $(\mathbf{X}|Y = r)$ are equal, that is, $\boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_R = \boldsymbol{\Sigma}$, then

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

$d_{n,p} = 2n(n-1)(R-1)\mathrm{tr}(\mathbf{\Sigma}^2)$. Under H_{10} in (2.1),

$$\frac{M_{n,p}}{\sqrt{d_{n,p}}} \xrightarrow{d} \mathcal{N}(0,1). \tag{2.3}$$

We use (2.3) to formulate a test procedure based on Theorem 2.1; thus, estimating $d_{n,p}$ is required. Here, we choose the estimators of $\mathrm{tr}(\mathbf{\Sigma}_r^2)$ and $\mathrm{tr}(\mathbf{\Sigma}_r\mathbf{\Sigma}_s)$ proposed by Li and Chen (2012), and use $\widehat{p}_r = (N_r - 1)/(n-1)$ to estimate p_r. As $n \to \infty$, $\hat{p}_r$ is consistent, by the law of large numbers, and $\widehat{\mathrm{tr}(\mathbf{\Sigma}_r^2)}$ and $\widehat{\mathrm{tr}(\mathbf{\Sigma}_r\mathbf{\Sigma}_s)}$ are consistent under Conditions 1, 2, and 3 by Theorem 2 in Li and Chen (2012). Additional details about the algorithm for calculating $\widehat{\mathrm{tr}(\mathbf{\Sigma}_r^2)}$ and $\widehat{\mathrm{tr}(\mathbf{\Sigma}_r\mathbf{\Sigma}_s)}$ are discussed in Section 4. The proposed test rejects H_{10} at significance level α if $M_{n,p} \geq \widehat{d_{n,p}}^{1/2} z_\alpha$, where z_α is the upper-α quantile of $\mathcal{N}(0,1)$. Theorem 2.1 also implies that the proposed test has the asymptotic local power

$$\Psi_{1,n}^{New}(\boldsymbol{\mu}_1, \cdots, \boldsymbol{\mu}_R; \alpha) = \Phi\left(-z_\alpha + \frac{n\sum_{r>s} p_r p_s \|\boldsymbol{\mu}_r - \boldsymbol{\mu}_s\|^2}{\sqrt{d_{n,p}/n^2}}\right).$$

When $\sum_{r>s} p_r p_s \|\boldsymbol{\mu}_r - \boldsymbol{\mu}_s\|^2$ has a higher order of $\sqrt{p}/n$, the power converges to one.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

3. Homogeneity Test for Covariance Matrices

In this section, we consider the homogeneity test for covariance matrices, that is,

$$H_{20} : \mathbf{\Sigma}_1 = \cdots = \mathbf{\Sigma}_R = \mathbf{\Sigma}, \tag{3.1}$$

versus the composite alternative H_{21}: $\mathbf{\Sigma}_r \neq \mathbf{\Sigma}_s$, for $1 \leq r < s \leq R$. Here, $\mathbf{\Sigma}_r = \mathrm{var}(\mathbf{X}|Y = r)$ and $\mathbf{\Sigma} = \sum_{r=1}^{R} p_r \mathbf{\Sigma}_r$.

3.1 Measuring the difference between covariance matrices

Similarly to the analysis in Section 2, we propose an index to measure the difference between $\mathbf{\Sigma}_r$, for $r = 1, 2, \cdots, R$. The expression of this index is relatively complex compared with that of $\mathcal{U}(\mathbf{X}|Y)$.

Definition 2. The distance between the covariances of R categories is defined by

$$\mathcal{V}(\mathbf{X}|Y) = \frac{1}{4}\mathrm{E}\left\{(\mathbf{X}_1 - \mathbf{X}_2)^{\mathrm{T}}(\mathbf{X}_3 - \mathbf{X}_4)\right\}^2 f_{1234},$$

where $(\mathbf{X}_i, Y_i)$, for $i = 1, \cdots, 4$, are independent copies of $(\mathbf{X}, Y)$, and

$$f_{1234} = \sum_{r=1}^{R} I(Y_1 = Y_2 = Y_3 = Y_4 = r)\frac{(1 - p_r)}{p_r^3} - \sum_{(r,s)}^{*} \frac{I(Y_1 = Y_2 = r)I(Y_3 = Y_4 = s)}{p_r p_s}.$$

The following lemma ensures that Definition 2 is reasonable.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

Lemma 3.1. If $\mathbf{X}$ has a finite second moment, then $\mathcal{V}(\mathbf{X}|Y) \geq 0$, and the equality holds if and only if the null hypothesis (3.1) is true.

Similarly to the analysis for testing means, we define

$$T_{n,p} = \sum_{(i_1,i_2,i_3,i_4)}^{*} \frac{1}{4} \left\{ (\mathbf{X}_{i_1} - \mathbf{X}_{i_2})^{\mathrm{T}} (\mathbf{X}_{i_3} - \mathbf{X}_{i_4}) \right\}^2 \hat{f}_{i_1 i_2 i_3 i_4},$$

where $\sum^{*}_{(i_1,i_2,i_3,i_4)}$ denotes summations over distinct indices, and

$$\begin{aligned} \hat{f}_{i_1 i_2 i_3 i_4} = & \sum_{r=1}^{R} I(Y_{i_1} = Y_{i_2} = Y_{i_3} = Y_{i_4} = r) \frac{(1 - \hat{p}_r)}{\widehat{p_r^3}} \\ & - \sum_{(r,s)}^{*} \frac{I(Y_{i_1} = Y_{i_2} = r) I(Y_{i_3} = Y_{i_4} = s)}{\hat{p}_r \hat{p}_s}, \end{aligned}$$

with $\hat{p}_r = (N_r - 1)/(n - 1)$ and $\widehat{p_r^3} = (N_r - 3)(N_r - 2)(N_r - 1)/(n - 1)^3$.

3.2 Main results for the homogeneity test for covariance matrices

Theorem 3.1. Suppose that Conditions 1, 2, and 3 hold. Then, we have

$$\frac{T_{n,p} - (n-1)^2 \sum_{r>s}^{R} N_r N_s \mathrm{tr} \left\{ (\mathbf{\Sigma}_r - \mathbf{\Sigma}_s)^2 \right\}}{\sqrt{\delta_{n,p}}} \xrightarrow{d} \mathcal{N}(0,1)$$

as $p \to \infty$ and $n \to \infty$, where

$$\begin{aligned} \delta_{n,p} = & \ 4n^6 \left\{ \sum_{r=1}^{R} (1 - p_r)^2 \mathrm{tr}^2(\mathbf{\Sigma}_r^2) + \sum_{(r,s)}^{*} p_r p_s \mathrm{tr}^2(\mathbf{\Sigma}_r \mathbf{\Sigma}_s) \right\} \\ & + 8n^7 \sum_{r=1}^{R} p_r \mathrm{tr} \left\{ (\mathbf{\Sigma}_r^2 - \mathbf{\Sigma}_r \mathbf{\Sigma})^2 \right\} \\ & + 4 \triangle n^7 \sum_{r=1}^{R} p_r \mathrm{tr} \left\{ \mathbf{\Gamma}_r^{\mathrm{T}} (\mathbf{\Sigma}_r - \mathbf{\Sigma}) \mathbf{\Gamma}_r \circ \mathbf{\Gamma}_r^{\mathrm{T}} (\mathbf{\Sigma}_r - \mathbf{\Sigma}) \mathbf{\Gamma}_r \right\}. \end{aligned}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

Theorem 3.1 establishes the asymptotic normality of $T_{n,p}$. Under H_{20} and Condition 3, $\delta_{n,p} = 4n^6(R-1)\{\mathrm{tr}(\mathbf{\Sigma}^2)\}^2 = O(n^6p^2)$. We define

$$\delta_{0,n,p} = 4n^6\left\{\sum_{r=1}^{R}(1-p_r)^2\mathrm{tr}^2(\mathbf{\Sigma}_r^2) + \sum_{r\neq s}p_rp_s\mathrm{tr}^2(\mathbf{\Sigma}_r\mathbf{\Sigma}_s)\right\}.$$

From Theorem 3.1, we obtain

$$\frac{T_{n,p}}{\sqrt{\delta_{0,n,p}}} \xrightarrow{d} \mathcal{N}(0,1),$$

under H_{20}. To formulate a test procedure based on Theorem 3.1, we choose

$$\widehat{\delta_{0,n,p}} = 4n^6\left[\sum_{r=1}^{R}(1-\hat{p}_r)^2\left\{\widehat{\mathrm{tr}(\mathbf{\Sigma}_r^2)}\right\}^2 + \sum_{r\neq s}\hat{p}_r\hat{p}_s\left\{\widehat{\mathrm{tr}(\mathbf{\Sigma}_r\mathbf{\Sigma}_s)}\right\}^2\right].$$

The proposed test rejects H_{20} at significance level α if $T_{n,p} \geq \widehat{\delta_{0,n,p}}^{1/2}z_\alpha$. Theorem 3.1 also implies that the proposed test has asymptotic power

$$\Psi_{2,n}^{New}(\boldsymbol{\mu}_1,\cdots,\boldsymbol{\mu}_R;\alpha) = \Phi\left[-\sqrt{\frac{\delta_{0,n,p}}{\delta_{n,p}}}z_\alpha + \frac{\sum_{r>s}p_rp_s\mathrm{tr}\left\{(\mathbf{\Sigma}_r-\mathbf{\Sigma}_s)^2\right\}}{\sqrt{\delta_{n,p}/n^8}}\right].$$

When $\sum_{r>s}^{R}p_rp_s\mathrm{tr}\left\{(\mathbf{\Sigma}_r-\mathbf{\Sigma}_s)^2\right\}$ is of order p/n, the power converges to one.

4. Implementation

In this section, we introduce two efficient algorithms for our proposed tests for mean vectors and covariance matrices.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

4.1 Testing for mean vectors

When calculating the statistics of two tests, we need to introduce an efficient algorithm to estimate $\mathrm{tr}(\mathbf{\Sigma}_r^2)$ and $\mathrm{tr}(\mathbf{\Sigma}_r\mathbf{\Sigma}_s)$. We use the estimators of $\mathrm{tr}(\mathbf{\Sigma}_r^2)$ and $\mathrm{tr}(\mathbf{\Sigma}_r\mathbf{\Sigma}_s)$ proposed by Li and Chen (2012). That is,

$$\begin{aligned}
\widehat{\mathrm{tr}(\mathbf{\Sigma}_r^2)} &= \frac{1}{N_r(N_{r-1})}\sum_{(i,j)}^{*}(\mathbf{X}_{ri}^{\mathrm{T}}\mathbf{X}_{rj})^2 - \frac{2}{N_r(N_r-1)(N_r-2)}\sum_{(i,j,k)}^{*}\mathbf{X}_{ri}^{\mathrm{T}}\mathbf{X}_{rj}\mathbf{X}_{rj}^{\mathrm{T}}\mathbf{X}_{rk} \\
&\quad + \frac{1}{N_r(N_r-1)(N_r-2)(N_r-3)}\sum_{(i,j,k,l)}^{*}\mathbf{X}_{ri}^{\mathrm{T}}\mathbf{X}_{rj}\mathbf{X}_{rk}^{\mathrm{T}}\mathbf{X}_{rl}, \\
\widehat{\mathrm{tr}(\mathbf{\Sigma}_r\mathbf{\Sigma}_s)} &= \frac{1}{N_rN_s}\sum_{i}\sum_{j}(\mathbf{X}_{ri}^{\mathrm{T}}\mathbf{X}_{sj})^2 - \frac{1}{N_rN_s(N_r-1)}\sum_{(i,k)}^{*}\sum_{j}\mathbf{X}_{ri}^{\mathrm{T}}\mathbf{X}_{sj}\mathbf{X}_{sj}^{\mathrm{T}}\mathbf{X}_{rk} \\
&\quad - \frac{1}{N_rN_s(N_s-1)}\sum_{(i,k)}^{*}\sum_{j}\mathbf{X}_{si}^{\mathrm{T}}\mathbf{X}_{rj}\mathbf{X}_{rj}^{\mathrm{T}}\mathbf{X}_{sk} \\
&\quad + \frac{1}{N_r(N_r-1)N_s(N_s-1)}\sum_{(i,j,k,l)}^{*}\mathbf{X}_{ri}^{\mathrm{T}}\mathbf{X}_{sj}\mathbf{X}_{rk}^{\mathrm{T}}\mathbf{X}_{sl}.
\end{aligned}$$

Then, we obtain

$$\widehat{\mathrm{tr}(\mathbf{\Sigma}_r^2)} = \frac{1}{N_r(N_r-3)}\sum_{i\neq j}A_{ij}^rA_{ij}^r, \tag{4.1}$$

where $A_{ij}^r = a_{ij}^r - a_i^r/(N_r-2) - a_j^r/(N_r-2) + a^r/(N_r-1)/(N_r-2)$, with $a_{ij}^r = \|\mathbf{X}_{ri}-\mathbf{X}_{rj}\|^2/2$, $a_i^r = \sum_{k=1}^{N_r}a_{ik}^r$, and $a^r = \sum_{k=1}^{N_r}\sum_{l=1}^{N_r}a_{kl}^r$.

Similarly,

$$\widehat{\mathrm{tr}(\mathbf{\Sigma}_r\mathbf{\Sigma}_s)} = \frac{1}{(N_r-1)(N_s-1)}\sum_{i=1}^{N_r}\sum_{j=1}^{N_s}\left\{(\mathbf{X}_{ri}-\overline{\mathbf{X}}_r)^{\mathrm{T}}(\mathbf{X}_{sj}-\overline{\mathbf{X}}_s)\right\}^2, \tag{4.2}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

where $\overline{\mathbf{X}}_t = \sum_{i=1}^{N_t} \mathbf{X}_{ti}/N_t$, for $t = 1, 2, \cdots, R$. Because the proofs of Equations (4.1) and (4.2) require complicated calculations, we omit them here. Interested readers can derive them through numerical calculations.

4.2 Testing for covariance matrices

As indicated in Subsection 4.1, $\widehat{\delta_{0,n,p}}$ can be calculated straightforwardly. Hence, we discuss only the calculation of $T_{n,p}$ in the following. We write

$$\begin{aligned} D_r &= \frac{1}{4} \sum_{(i,j,k,l)}^{*} \{(\mathbf{X}_{ri} - \mathbf{X}_{rj})^{\mathrm{T}}(\mathbf{X}_{rk} - \mathbf{X}_{rl})\}^2, \\ D_{rs} &= \frac{1}{4} \sum_{(i,j)}^{*} \sum_{(k,l)}^{*} \{(\mathbf{X}_{ri} - \mathbf{X}_{rj})^{\mathrm{T}}(\mathbf{X}_{sk} - \mathbf{X}_{sl})\}^2. \end{aligned}$$

Then, $T_{n,p} = \sum_{r=1}^{R} D_r(1 - \hat{p}_r)/\hat{p}_r^3 - \sum_{(r,s)}^{*} D_{rs}/(\hat{p}_r\hat{p}_s)$. Similarly to the analysis for Equations (4.1) and (4.2), we obtain

$$\begin{aligned} D_r &= (N_r - 1)(N_r - 2) \sum_{(i,j)}^{*} A_{ij}^r A_{ij}^r, \\ D_{rs} &= N_r N_s \sum_{i=1}^{N_r} \sum_{j=1}^{N_s} \left\{(\mathbf{X}_{ri} - \overline{\mathbf{X}}_r)^{\mathrm{T}}(\mathbf{X}_{sj} - \overline{\mathbf{X}}_s)\right\}^2. \end{aligned}$$

Using the derivations, the two statistics and the associated parameters are expressed in the form of order two. Hence, these statistics are easy to calculate.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

5. Simulation Study

We design several simulation experiments to evaluate the performance of the two proposed tests by comparing them with other tests. Here, R is designed to be three or four, with probabilities $P_1 = (0.4, 0.4, 0.2)$ or $P_2 = (0.3, 0.3, 0.2, 0.2)$, respectively. We choose $n = 100$ or 200, and p ranges from 50 to 400.

Example 1 Test for means

We compare the proposed test for means (NEW.mean) with the distance covariance (dCov) test developed by Székely et al. (2007), the rank of distance test (HHG) proposed by Heller et al. (2013), and the HBWW test suggested by Hu et al. (2017). The distances of Y_i and Y_j when applying the dCov and HHG tests are defined as one if they are different, and zero otherwise. We randomly generate a categorical random variable Y from R classes. Then, for each given $Y_i = r$, the ith predictor $\mathbf{X}_i$ is generated by letting $\mathbf{X}_i = \boldsymbol{\mu}_r + \boldsymbol{\xi}_i$, where $\boldsymbol{\xi}_i$, for $i = 1, \cdots, n$, are random errors following $\mathcal{N}(\mathbf{0}, \mathbf{I}_p)$ or $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma} = (0.5^{|i-j|})$. We set $\boldsymbol{\mu}_1 = \text{signal} * (1, 2, 3, 0, \cdots, 0)^{\mathrm{T}}/\sqrt{14}$, $\boldsymbol{\mu}_2 = \text{signal} * (1, \cdots, 1, 0, \cdots, 0)^{\mathrm{T}}/\sqrt{p/2}$, and $\boldsymbol{\mu}_r = \mathbf{0}$, for $r \neq 1, 2$. The tests are repeated 1000 times to simulate the power.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

Table 1 shows the empirical sizes of the proposed test (NEW.mean) and the related tests (dCov, HHG, and HBWW). As shown in Table 1, the empirical sizes in all tests maintain the 5% nominal level. Figures 1 and 2 depict the empirical power of the tests when $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_p)$. As the "signal" increases, the proposed test outperforms the three other tests, and dCov and HBWW tests exhibit similar performance. The HHG test is the least effective in terms of detecting difference between the means of the R groups, implying that considering only the rank of a distance leads to a severe loss of information on distance. For example, when $(n, p) = (100, 200)$, $R = 3$, and signal $= 1.4$, the empirical power of the proposed test reaches as high as 67.3%. In contrast, the dCov and HBWW tests have power of 56.2% and 51.0%, respectively, and the HHG test has power of only 7.0%. Figure 3 displays the empirical power as p increases. The proposed test consistently outperforms the other tests.

Example 2 Test for covariance matrices

We compare our proposed test for covariances (NEW.cov) with the distance covariance (dCov) test developed by Székely et al. (2004) and Székely et al. (2007), the rank of distance test (HHG) proposed by Heller et al. (2013), the ZBHW test suggested by Zhang et al. (2018), and the ZLGY

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

Table 1: Empirical sizes of the NEW.mean, dCov, HHG, and HBWW tests for means at a significance level of 5% in Example 1.

n	p	$R=3$				$R=4$			
		NEW.mean	dCov	HHG	HBWW	NEW.mean	dCov	HHG	HBWW
		Case 1: $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_p)$							
100	50	0.043	0.033	0.042	0.036	0.060	0.043	0.047	0.057
	100	0.056	0.049	0.045	0.051	0.064	0.049	0.031	0.060
	150	0.047	0.043	0.040	0.039	0.063	0.043	0.061	0.053
	200	0.062	0.051	0.035	0.057	0.045	0.038	0.072	0.040
200	50	0.065	0.041	0.059	0.062	0.056	0.044	0.058	0.057
	100	0.058	0.048	0.057	0.059	0.063	0.043	0.047	0.066
	150	0.056	0.042	0.055	0.058	0.047	0.048	0.058	0.048
	200	0.058	0.047	0.051	0.062	0.060	0.047	0.045	0.060
		Case 2: $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$							
100	50	0.067	0.047	0.046	0.065	0.055	0.044	0.046	0.049
	100	0.059	0.041	0.061	0.050	0.055	0.041	0.051	0.053
	150	0.061	0.041	0.051	0.056	0.065	0.042	0.049	0.056
	200	0.058	0.044	0.044	0.055	0.054	0.051	0.048	0.055
200	50	0.051	0.048	0.055	0.054	0.066	0.044	0.040	0.068
	100	0.063	0.057	0.048	0.068	0.056	0.048	0.049	0.059
	150	0.050	0.047	0.042	0.046	0.048	0.041	0.047	0.046
	200	0.061	0.057	0.042	0.059	0.052	0.041	0.044	0.055

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

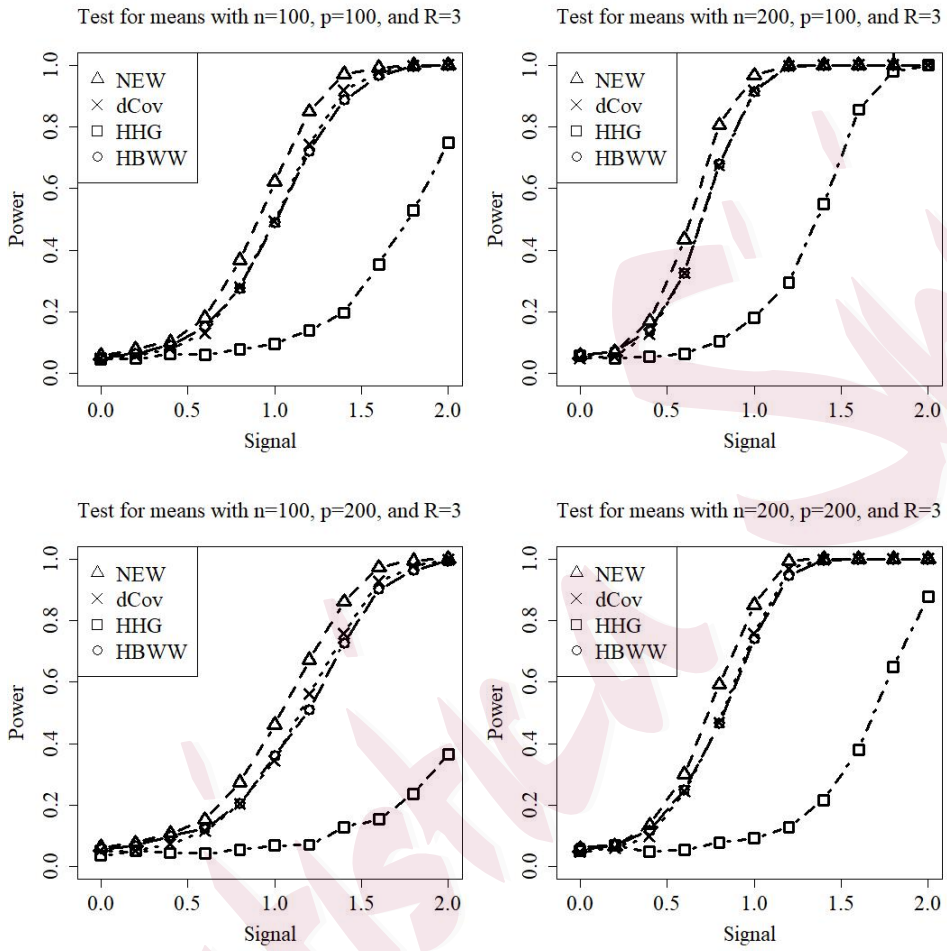

Figure 1: Performance of tests for means with different (n, p) and $R = 3$.

test introduced by Zheng et al. (2020). We randomly generate a categorical random variable Y from R classes. Then, for each given $Y_i = r$, the ith predictor $\mathbf{X}_i$ is generated by letting $\mathbf{X}_i = \mathbf{\Sigma}_r^{1/2}\mathbf{Z}_i$, where $\mathbf{Z}_i$, for $i = 1, \cdots, n$, are random errors following $\mathcal{N}(\mathbf{0}, \mathbf{I}_p)$. Set $\mathbf{\Sigma}_1 = 3\mathbf{I}_p + \text{signal} * \boldsymbol{\eta}_1\boldsymbol{\eta}_1^{\mathrm{T}}$, $\mathbf{\Sigma}_2 = 3\mathbf{I}_p + \text{signal} * \text{diag}(w_1, \cdots, w_p)$, and $\mathbf{\Sigma}_r = 3\mathbf{I}_p$, for $r \neq 1, 2$, where $\boldsymbol{\eta}_1 = (3, 3, 3, 0, \cdots, 0)^{\mathrm{T}}$ and $w_i \overset{iid.}{\sim} \text{Unif}(-3, 3)$.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

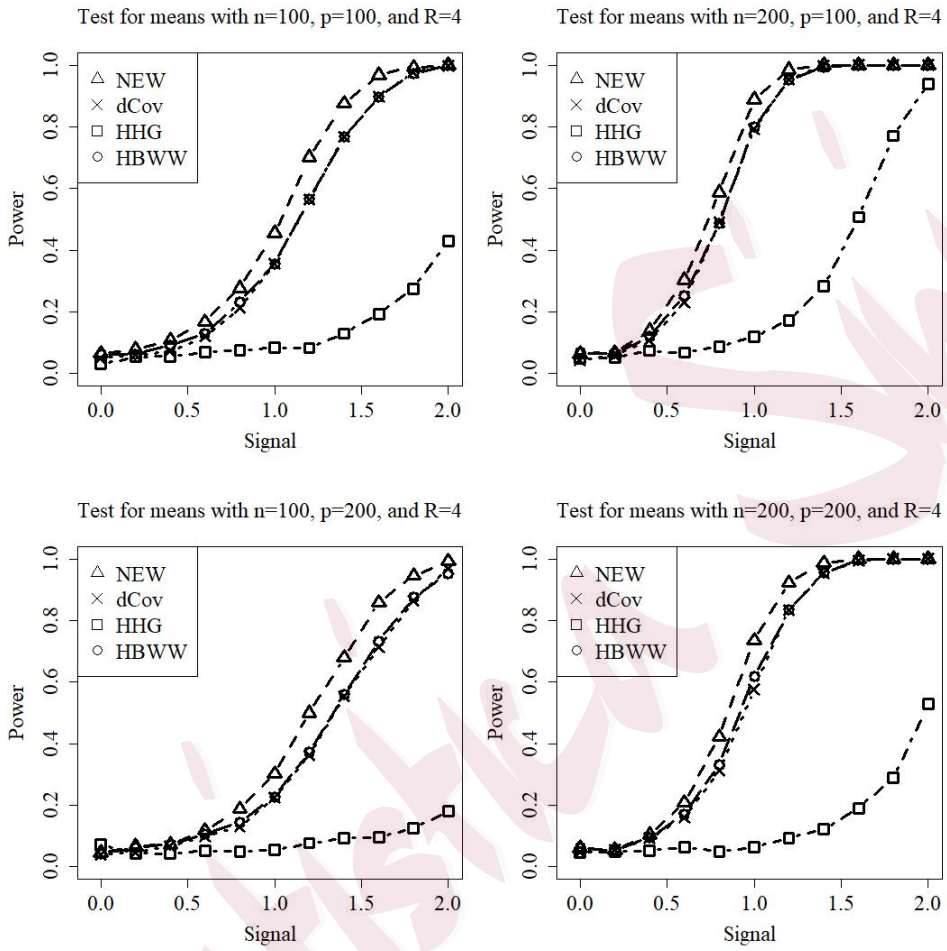

Figure 2: Performance of tests for means with different (n, p) and $R = 4$.

Table 2 presents the empirical sizes of the tests. As n and p approach infinity, the sizes of the five tests are close to the 5% nominal level. Figures 4 and 5 show the empirical power of the tests. As the "signal" increases, the proposed test outperforms the four other tests. Unlike in the test for means, the HHG test for covariances performs much better than the dCov test, which has power of around 5%. For example, when $(n, p) = (100, 200)$,

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

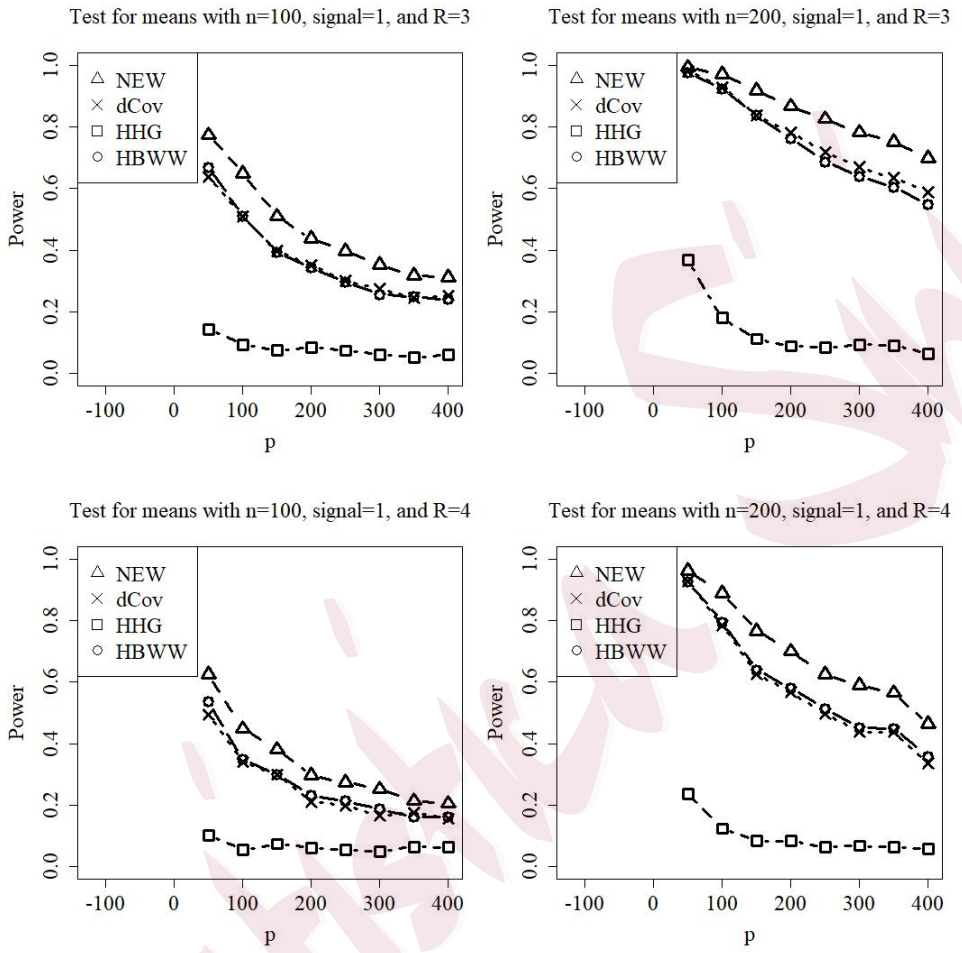

Figure 3: Performance of tests for means with different p values.

$R = 3$, and signal $= 0.7$, the empirical power of the proposed test reaches as high as 80.1%. In contrast, the ZBHW test has power of around 67.4%, the ZLGY test has power of around 62.8%, and the HHG test has power of only around 36.8%. Figure 6 displays the empirical power as p increases. Again, the proposed test consistently outperforms the other tests.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

Table 2: Empirical sizes of the NEW.cov, dCov, HHG, ZBHW, and ZLGY tests for covariances at a significance level of 5% in Example 2.

n	p	$R=3$					$R=4$				
		NEW.cov	dCov	HHG	ZBHW	ZLGY	NEW.cov	dCov	HHG	ZBHW	ZLGY
100	50	0.050	0.065	0.062	0.041	0.039	0.058	0.055	0.056	0.058	0.054
	100	0.064	0.054	0.041	0.065	0.047	0.055	0.054	0.047	0.052	0.047
	150	0.051	0.053	0.049	0.039	0.031	0.040	0.047	0.062	0.033	0.031
	200	0.062	0.038	0.062	0.043	0.033	0.054	0.059	0.053	0.050	0.039
200	50	0.060	0.054	0.047	0.046	0.043	0.060	0.068	0.057	0.060	0.050
	100	0.051	0.062	0.043	0.055	0.046	0.061	0.056	0.027	0.042	0.037
	150	0.060	0.049	0.039	0.048	0.039	0.056	0.045	0.054	0.056	0.041
	200	0.032	0.069	0.035	0.027	0.027	0.051	0.050	0.045	0.050	0.039

6. Application

6.1 Application 1

We apply the proposed test to a gene expression data set collected by Koh et al. (2014) to identify gene sets with significant differences in their mean vectors and covariances over time. The data set contains data on 11 pregnant women at four stages, namely, three stages during pregnancy (i.e., the first, second, and third trimesters) and one stage after delivery (i.e., postpartum). The microarray gene expression data in this data set were measured repeatedly, using 33,297 genes for each pregnant woman

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

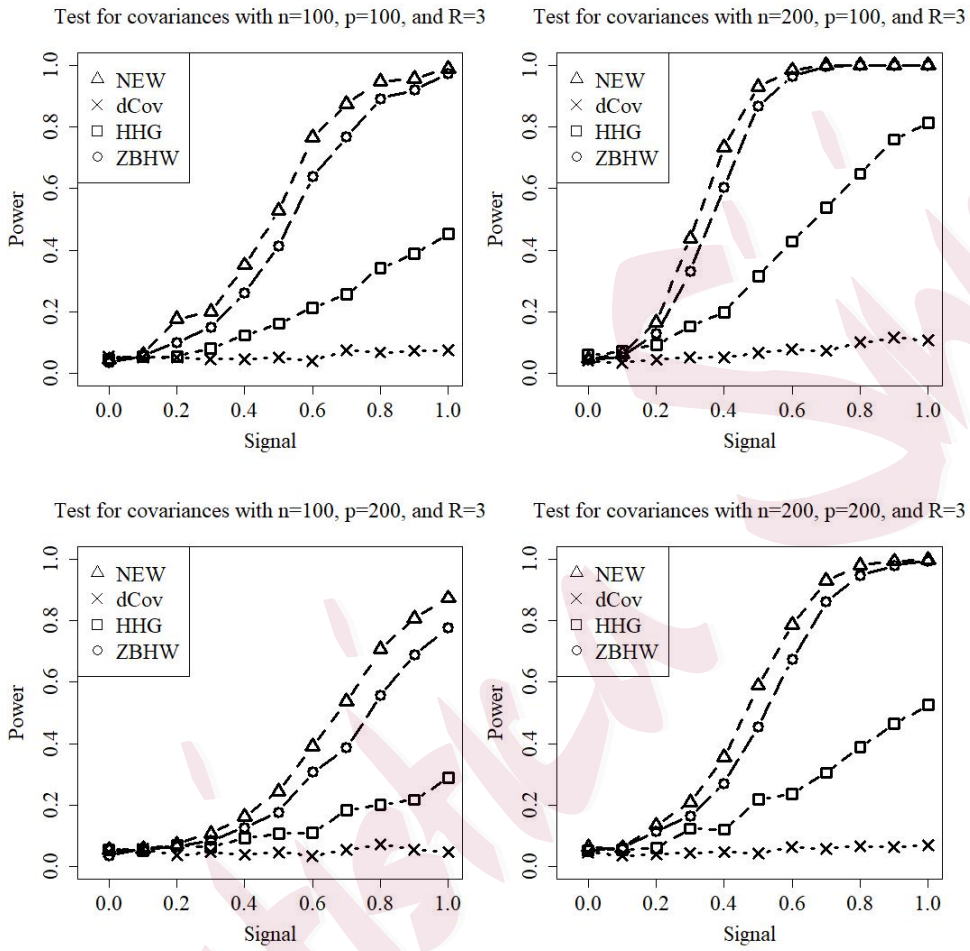

Figure 4: Performance of tests for covariances with different (n, p) and $R = 3$.

at the four stages. Based on their biological functions, the genes were defined using gene ontology (GO), yielding 3,910 GO terms. The data set is obtained from `http://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS5088`. For each GO term, our aim is to test whether the mean vectors and covariance matrices of the gene expression data are the same during the four stages. Table 3 shows the GO terms detected as significant

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

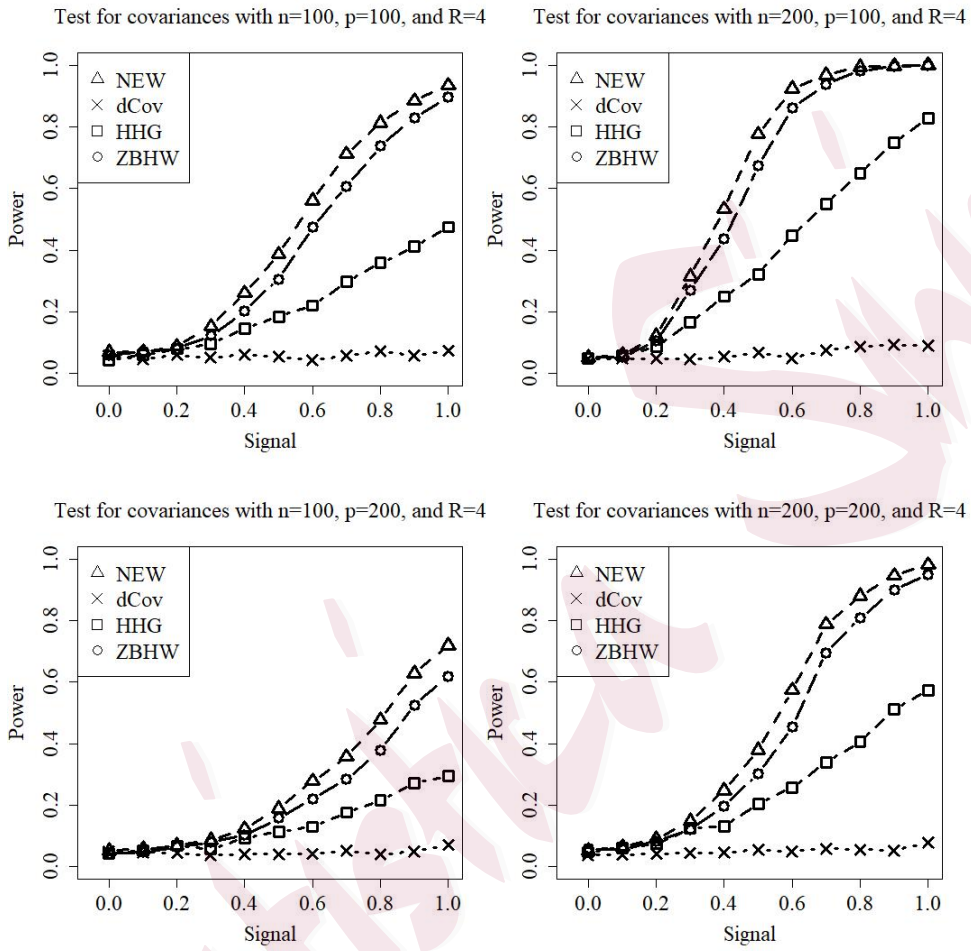

Figure 5: Performance of tests for covariances with different (n, p) and $R = 4$.

by the NEW.mean, dCov, HHG, HBWW, NEW.cov, ZBHW, and ZLGY tests. The gene set GO:0008499 is detected as significant only by the proposed NEW.mean test, and GO:0070513 and GO:0043008 are detected as significant only by the dCov test. A possible reason for this finding is that the proposed NEW.mean test is designed to detect the difference between mean vectors, whereas the dCov test focuses on identifying the variation

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

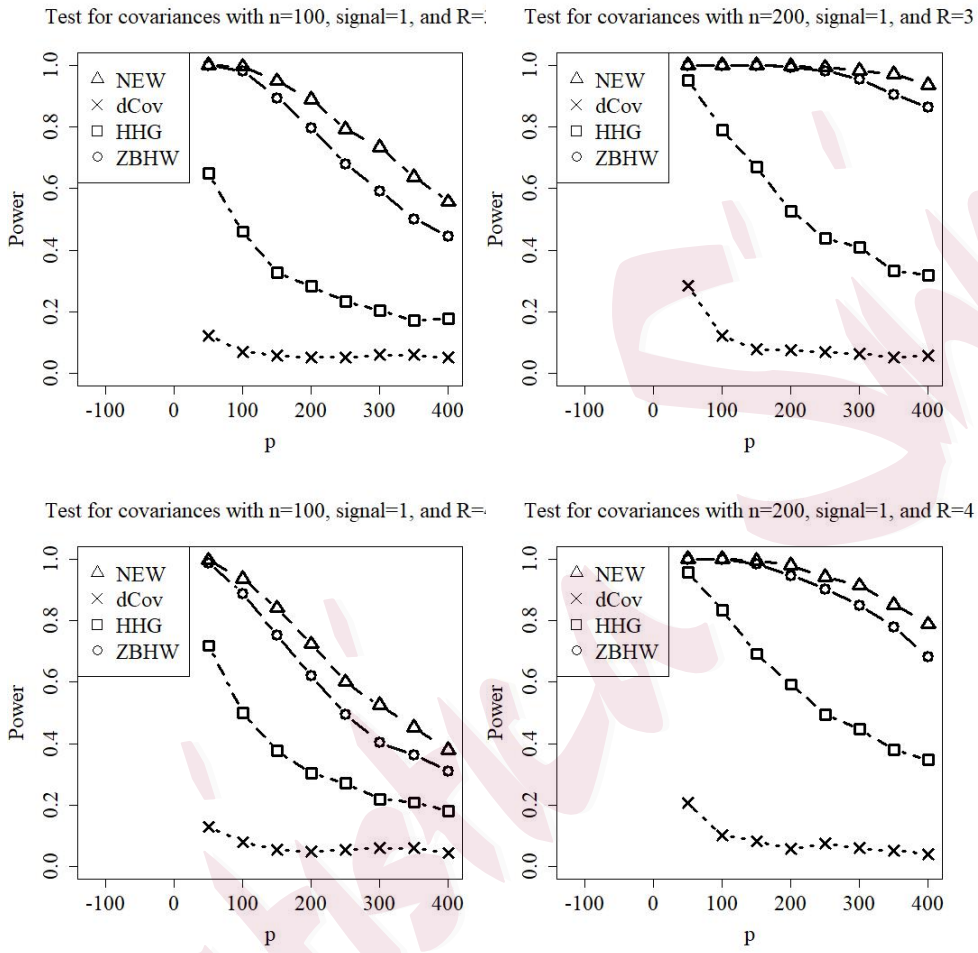

Figure 6: Performance of tests for covariances with different $p-$values.

of distribution functions. The NEW.mean and HBWW tests identify 12 GO terms as significant. Of these, GO:0050786 and GO:0005212 are also detected by the dCov and HHG tests, respectively, and GO:0005179 is identified as significant by the dCov and HHG tests.

In addition, our proposed NEW.cov test and the ZBHW test identify 12 other GO terms as significant gene sets for covariance matrices. However,

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

the dCov and HHG tests fail to identify any of them. Note that the ZLGY test identifies 219 significant GO terms, of which the trace-based term identifies nine, and the maximum norms detect the rest. This finding implies that using dimension reduction or feature screening methods can further enhance the power for detecting significant gene sets under sparse alternatives. It also motivates a promising extension of our tests to incorporate dimension reduction or feature screening.

Table 3: Significant GO terms obtained by the different tests at a significance level of 5%.

GO term	No. of genes	Satisfied test(s)	GO term	No. of genes	Satisfied test(s)
GO:0004869	36	NEW.mean	GO:0008200	11	NEW.cov/ZBHW/ZLGY
GO:0070513	17	dCov	GO:0008378	11	NEW.cov/ZBHW/ZLGY
GO:0043008	10	dCov	GO:0047617	14	NEW.cov/ZBHW
GO:0008499	15	NEW.mean/HBWW	GO:0015267	12	NEW.cov/ZBHW
GO:0008083	171	NEW.mean/HBWW	GO:0004012	15	NEW.cov/ZBHW/ZLGY
GO:0019864	13	NEW.mean/HBWW	GO:0032393	17	NEW.cov/ZBHW/ZLGY
GO:0015254	15	NEW.mean/HBWW	GO:0019870	10	NEW.cov/ZBHW/ZLGY
GO:0015204	10	NEW.mean/HBWW	GO:0070410	17	NEW.cov/ZBHW
GO:0015250	16	NEW.mean/HBWW	GO:0016712	10	NEW.cov/ZBHW/ZLGY
GO:0048037	18	NEW.mean/HBWW	GO:0033038	19	NEW.cov/ZBHW/ZLGY
GO:0005524	13	NEW.mean/HBWW	GO:0030275	10	NEW.cov/ZBHW/ZLGY
GO:0016594	14	NEW.mean/HBWW	GO:0030109	16	NEW.cov/ZBHW/ZLGY
GO:0050786	11	NEW.mean/dCov/HBWW			
GO:0005212	20	NEW.mean/HHG/HBWW			
GO:0005179	92	NEW.mean/dCov/HHG/HBWW			

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

6.2 Application 2

Here, we apply the proposed tests to a gene expression data set collected by Taylor et al. (2007) in a study to identify gene sets with significant differences in mean vectors and covariances over time. In this study, 69 patients with the hepatitis C virus were treated for up to 48 weeks using a specific clinical protocol. Their peripheral blood mononuclear cells were collected before treatment (day 0), and on days 1, 2, 7, 14, and 28 during treatment. The original data set is available at `https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7123`. The microarray gene expression data were measured using 22,283 genes for each patient repeatedly at six stages, defined using GO, based on the biological function of these genes. There are 1,218 GO terms, and a given gene can be a member of multiple GO terms. Further details about the data can be found in Taylor et al. (2007).

Before applying our tests, we preprocess the data by removing 11 individuals with an absent Microarray Suite 5.0 signal transcript, and keep 58 individuals with gene expression arrays at all six stages. We apply the NEW.mean test, dCov test, HHG test, HBWW test, NEW.cov test, ZBHW test, and ZLGY test to the 585 GO terms, with minimums of 10 genes. Let $X_{ri}^{(g)}|Y_i = r$ $(i = 1, 2, \cdots, 58, r = 1, 2, \cdots, 6, g = 1, 2, \cdots, 585)$ be the gene expression data for the gth GO term of the ith individual at the rth period,

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

where $r = 1, 2, \cdots, 6$ represents day 0, 1, 2, 7, 14, and 28, respectively. For each GO term, we test whether the means $\boldsymbol{\mu}_r^{(g)}$ and covariance matrices $\boldsymbol{\Sigma}_r^{(g)}$ are the same across $r = 1, 2, \cdots, 6$. Table 4 shows the various numbers of GO terms detected as significant by tests.

In all six stages, the NEW.mean and HBWW tests identify 525 and 524 GO terms, respectively, as significant, where the New.mean test detects GO:0005721, but the HBWW test does not. The dCov and HHG tests simultaneously identify only 459 of the 524 GO terms as significant. For the covariance matrices, the NEW.cov, ZBHW, and ZLGY tests identify 264, 263, and 297 GO terms, respectively, as significant, where the NEW.cov and ZLGY tests detect GO:0000792, but the ZBHW test fails to do so.

Table 4: Number of significant GO terms detected by different tests at a significance level of 5%

	NEW.mean	HBWW	dCov	HHG	NEW.cov	ZBHW	ZLGY
Day 0, 1, 2, 7, 14, and 28	525	524	543	475	264	263	297
Day 0 and 1	525	525	535	447	297	296	310
Day 1 and 2	138	137	149	78	42	42	39
Day 2 and 7	315	311	395	248	126	126	123
Day 7 and 14	41	41	48	21	157	157	145
Day 14 and 28	55	54	40	26	122	122	122

After identifying the significant GO terms, we apply the tests on binary

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

segmentation to identify the changes over time. As shown in Table 4, most of the identified changes in the mean vectors and the covariance matrices occurred within days zero and one. However, during the treatment, more GO terms are detected as having significant changes in means between days two and seven. In contrast, more significant changes are identified in the covariance matrices between days 7 and 14. These findings complement the results of Taylor et al. (2007), who observed that the majority of genes altered expression.

7. Conclusion

This study develops two categorically weighted tests for means and covariance matrices in high dimensions. Simulation studies and applications demonstrate the satisfactory performance of our tests. However, the present study has limitations, providing opportunities of future work in this area. While our proposed tests accommodate the high-dimensional setting, they are affected adversely by an increasing dimension, as shown in Figures 3 and 6. Therefore, they cannot deal with ultrahigh-dimensional problems. Moreover, the two tests are less powerful in detecting sparse signals of means and covariance matrices, which may be corrected using dimension reduction or feature screening.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

8. Supplementary Material

All technical proofs are provided in the Supplementary Material.

Acknowledgments

Guo's research was supported by grants from the State Key Program of the National Natural Science Foundation of China (12031016), the National Natural Science Foundation of China (11901406), and the Beijing Natural Science Foundation (Z210003). Song's research was supported by GRF (14302519, 14302220) from the Research Grant Council of Hong Kong Special Administration Region. Cui's research was supported by a grant from the National Natural Science Foundation of China (11971324). This work was also supported by the R&D Program of Beijing Municipal Education Commission (KM202110028017), and the Interdisciplinary Construction of Bioinformatics and Statistics.

References

Bai, Z., D. Jiang, and Y. S. Zheng (2009). Corrections to lrt on large dimensional covariance matrix by rmt. *Annals of Statistics 37*(6B), 3822–3840.

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

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

REFERENCES

Cai, T. T. and W. Liu (2016). Large-scale multiple testing of correlations. *Journal of the American Statistical Association 111*(513), 229–240.

Cai, T. T., W. Liu, and Y. Xia (2013). Two-sample test of high dimensional means under dependence. *Journal Of The Royal Statistical Society 76*(2).

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

Cai, T. T. and Y. Xia (2014). High-dimensional sparse manova. *Journal of Multivariate Analysis 131*, 174–196.

Chang, J., C. Zheng, W.-X. Zhou, and W. Zhou (2017). Simulation-based hypothesis testing of high dimensional means under covariance heterogeneity. *Biometrics 73*(4), 1300–1310.

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

Chen, S. X., L. X. Zhang, and P. S. Zhong (2010). Tests for high-dimensional covariance matrices. *Journal of the American Statistical Association 105*(490), 810–819.

Cui, H., R. Li, and W. Zhong (2015). Model-free feature screening for ultrahigh dimensional discriminant analysis. *Journal of the American Statistical Association 110*(510), 630–641.

Feng, L., C. Zou, and Z. Wang (2016). Multivariate-sign-based high-dimensional tests for the two-sample location problem. *Publications of the American Statistical Association 111*(514), 15.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

REFERENCES

Heller, R., Y. Heller, and M. Gorfine (2013). A consistent multivariate test of association based on ranks of distances. *Biometrika 100*(2), 503–510.

Hu, J., Z. Bai, C. Wang, and W. Wang (2017). On testing the equality of high dimensional mean vectors with unequal covariance matrices. *Annals of the Institute of Statistical Mathematics 69*(2), 365–387.

Jiang, T. and F. Yang (2013). Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions. *The Annals of Statistics 41*(4), 2029–2074.

Koh, W., W. Pan, C. Gawad, H. C. Fan, G. A. Kerchner, T. Wyss-Coray, Y. J. Blumenfeld, Y. Y. El-Sayed, and S. R. Quake (2014). Noninvasive in vivo monitoring of tissue-specific global gene expression in humans. *Proceedings of the National Academy of Sciences 111*(20), 7361–7366.

Li, J. and S. X. Chen (2012). Two sample tests for high-dimensional covariance matrices. *The Annals of Statistics 40*(2), 908–940.

Liu, Z., B. Liu, S. Zheng, and N.-Z. Shi (2017). Simultaneous testing of mean vector and covariance matrix for high-dimensional data. *Journal of Statistical Planning and Inference 188*, 82–93.

Srivastava, M. S. and M. Du (2008). A test for the mean vector with fewer observations than the dimension. *Journal of Multivariate Analysis 99*(3), 386–402.

Székely, G., M. L. Rizzo, and N. K. Bakirov (2007). Measuring and testing dependence by correlation of distances. *Annals of Statistics 35*(6), 2769–2794.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

REFERENCES

Székely, G. J., M. L. Rizzo, et al. (2004). Testing for equal distributions in high dimension. *InterStat 5*(16.10), 1249–1272.

Taylor, M. W., T. Tsukahara, L. Brodsky, J. Schaley, C. Sanda, M. J. Stephens, J. N. McClintick, H. J. Edenberg, L. Li, J. E. Tavis, et al. (2007). Changes in gene expression during pegylated interferon and ribavirin therapy of chronic hepatitis c virus distinguish responders from nonresponders to antiviral therapy. *Journal of virology 81*(7), 3391–3401.

Wolf, L. M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Annals of Statistics 30*(4), 1081–1102.

Zhang, C., Z. Bai, J. Hu, and C. Wang (2018). Multi-sample test for high-dimensional covariance matrices. *Communications in Statistics-Theory and Methods 47*(13), 3161–3177.

Zheng, S., R. Lin, J. Guo, and G. Yin (2020). Testing homogeneity of high-dimensional covariance matrices. *Statistica Sinica 30*(1), 35–53.

Wenwen Guo

School of Mathematical Sciences, Capital Normal University, Beijing 100048, China. E-mail: wwguo@cnu.edu.cn

Xinyuan Song

Department of Statistics, The Chinese University of Hong Kong, Hong Kong, China. E-mail: xysong@cuhk.edu.hk

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0048

REFERENCES

Hengjian Cui

School of Mathematical Sciences, Capital Normal University, Beijing 100048, China. E-mail: hjcui@bnu.edu.cn