Performance Evaluation of Two Tests for the Equality of High-Dimensional Mean Vectors with Unequal Covariance Matrices

Dr.Knavoot Jiamwattanapong [a], Nisanad Ingadapa [b], Dr.Piyada Prueksawatnon [c], Dr.Bandhita Plubin [d]

[a] Assistant Professor, Department of General Education (Mathematics), Faculty of Liberal Arts, Rajamangala University of Technology Rattanakosin, Thailand
[b]Assistant Professor, Department of General Education (English), Faculty of Liberal Arts, Rajamangala University of Technology Rattanakosin, Thailand
[c]Head of Undergraduate Program in Statistics, School of Science, University of Phayao, Thailand
[d]Assistant Professor, Department of Statistics, Faculty of Science, Chiang Mai University, Thailand

Abstract: The rapid development of measurement technology in fields such as genetics, medicine, and economics has made high-dimensional data analysis an essential practice, as analysts and researchers need to collect and analyze data from a large number of variables. Analyzing high-dimensional data requires advanced statistical techniques as conventional methods like comparing mean vectors using Hotelling's T^2-test are not applicable. In this study, we compare the performance of two tests, T_{MCQ} and T_{SKK}, for testing the equality of high-dimensional mean vectors when population covariance matrices are not equal. The T_{MCQ} was initially proposed by Chen, S. X. and Qin, Y. L. in 2010 and modified by Srivastava, M. S., Katayama, S., & Kano, Y. in 2013, while the T_{SKK} was developed by Srivastava, M. S. et al. in 2013. A simulation study was conducted using two independent normal samples of equal size. The population covariance matrices were created under five structures: Sphericity, Compound Symmetry (CS), Heterogeneous Compound Symmetry (CSH), Toeplitz, and Block Diagonal (BD) matrix. The results showed that the choice of covariance structure had an impact on the performance of both tests. The T_{MCQ} test performed well for certain covariance structures including Sphericity, Toeplitz, and BD, while the T_{SKK} performed well for the CS covariance structure. The performance of T_{SKK} increased for large sample sizes of at least 60 under the covariance structures of Toeplitz and BD. Under the covariance structures of Sphericity, Toeplitz, and BD, the T_{MCQ} outperformed the T_{SKK}, while T_{SKK} was more efficient than T_{MCQ} when the covariance matrix structure was CS. Additionally, the performance of T_{SKK} improved when the sample size was at least 60. We also examined the performance of the two tests when the difference between two population covariance matrices increased and found that both tests continued to perform well for certain covariance structures. However, under the covariance structure of CSH, both tests underperformed for the case being studied.

Keywords: High-dimensional data, tests for mean vectors, unequal covariance matrices, covariance matrix structures

1. Introduction

Collecting data from a large number of variables, known as high-dimensional data, is common in various fields such as medicine, genetics, and economics. An example of high-dimensional data is DNA microarray data that stores information on thousands of human genes, with a sample size smaller than the number of variables. For instance, the **Alon, U., Barkai, N., Notterman, D. A., Gish, K., Ybarra, S., Mack, D., & Levine, A. J. (1999)** study shows that group 1 (normal tissue group) had a sample size of 22 while group 2 (tumor tissue group) had a sample size of 40, with 6,500 variables of interest. Although initial data analysis may group variables and reduce some of them, there may still be around 2,000 variables remaining. Another example can be seen in leading global business organizations, such as Illumina, which was ranked as the smartest organization in the world in 2014 **(MIT Technology Review, 2014)**. Illumina is a biotechnology company based in the United States that researches the sequencing of genes using the technology of optically reading DNA. This advanced technology can finish the process in hours instead of days, and the utilization of rapid data processing techniques leads to substantial reductions in operational expenses.

Analyzing high-dimensional data is distinct from analyzing data with few variables. As the number of variables increases relative to the sample size, the analysis becomes more complex. Traditional methods used for analyzing one or few variables, and even multivariate analysis such as comparing mean vectors using Hotelling's T^2-test **(Hotelling, H., 1931)**, are not applicable **(Bühlmann, P. & Van De Geer, S., 2011; Schott, J. R., 2005)**.

Testing the equality of two high-dimensional mean vectors can be considered under two scenarios: when population covariance matrices are equal and when they are not. The latter scenario, where population covariance matrices are not equal, is more challenging and requires more complex approaches, such as regularized estimators **(Tibshirani, R., 1996; Fan, J. & Li, R., 2001)**. Several testing procedures have been developed, such as the tests of **Bai, Z. & Saranadasa, H. (1996)**, **Srivastava, M. S. & Du, M. (2008)**, **Chen, S. X. & Qin, Y. L. (2010)**, **Srivastava, M. S., Katayama, S., & Kano, Y. (2013)**, and **Hu, J., Bai, Z., Wang, C., & Wang, W. (2017)**. However, the performance of these tests can be highly dependent on the conditions of the data, such as underlying covariance structures **(Gelper, S. et al., 2017; Schreiber, J. B., 2017)** .

The objectives of this research are to study and compare the performance of two tests - Modified Chen and Qin's test which was initially proposed by **Chen, S. X. & Qin, Y. L. (2010)** and then modified by **Srivastava, M. S. et al.(2013)**, and **Srivastava, M. S. et al. (2013)** test - in comparing two population mean vectors with unequal covariance matrices, and to investigate the performance of these tests with increasing differences between population covariance matrices. The research will focus on comparing the mean vectors of two independent high-dimensional normal populations, where high-dimensionality refers to the situation where the number of variables (p) is greater than the sample size (n), i.e., p > n. Specifically, we will consider two independent normal samples with sizes n_1 and n_2, where $n = n_1 + n_2 - 2$. The population covariance matrices will be unknown and unequal but will have the same structure. There will be five covariance structures considered in this study including Sphericity, Compound Symmetry (CS), Heterogeneous Compound Symmetry (CSH), Toeplitz, and Block Diagonal matrix (BD).

2. Tests for Two High-Dimensional Mean Vectors with Unequal Covariance Matrices

2.1. Chen and Qin's Test

When the covariance matrices of two populations are unequal (i.e., $\mathbf{\Sigma}_1 \neq \mathbf{\Sigma}_2$) and the covariance matrices of both populations are unknown, one important method for testing the equality of two population mean vectors is proposed by **Chen, S. X. & Qin, Y. L. (2010)**. This method builds upon the work of **Bai, Z. & Saranadasa, H. (1996)**, who developed a method for the case where the covariance matrices of both populations are assumed to be equal. However, Chen and Qin extended this method to the case of unequal covariance matrices in a higher-dimensional setting.

The testing method proposed by **Chen, S. X. & Qin, Y. L. (2010)**, denoted by T_{CQ}, is defined as follows:

$$T_{CQ} = \frac{Q_n}{\hat{\sigma}_{CQ}} \tag{1}$$

where
$$Q_n = \left[(\overline{\mathrm{x}}_1 - \overline{\mathrm{x}}_2)'(\overline{\mathrm{x}}_1 - \overline{\mathrm{x}}_2) - \frac{tr(\mathbf{S}_1)}{n_1} - \frac{tr(\mathbf{S}_2)}{n_2} \right] / \sqrt{p}\ , \tag{2}$$

$$\hat{\sigma}^2_{CQ} = \frac{1}{p}\left[\frac{2}{n_1(n_1-1)} tr(\mathbf{\Sigma}_1^2) + \frac{2}{n_2(n_2-1)} tr(\mathbf{\Sigma}_2^2) + \frac{4}{n_1 n_2} tr(\mathbf{\Sigma}_1 \mathbf{\Sigma}_2) \right], \tag{3}$$

$$tr(\mathbf{\Sigma}_i^2) = \frac{1}{n_i(n_i-1)} \left\{ \sum_{j \neq k}^{n_i} (\mathbf{x}_{ij} - \tilde{\mathbf{x}}_{i(j,k)}) \mathbf{x}_{ij}^{'} (\mathbf{x}_{ij} - \tilde{\mathbf{x}}_{i(j,k)}) \mathbf{x}_{ik}^{'} \right\},\ i = 1, 2,$$

$$tr(\mathbf{\Sigma}_1 \mathbf{\Sigma}_2) = \frac{1}{n_1 n_2} tr\left[\sum_{j=1}^{n_1} \sum_{k=1}^{n_2} (\mathbf{x}_{1j} - \tilde{\mathbf{x}}_{1(j)}) \mathbf{x}'_{1j} (\mathbf{x}_{2k} - \tilde{\mathbf{x}}_{2(\mathrm{k})}) \mathbf{x}'_{2k} \right],$$

where
$$\tilde{\mathbf{x}}_{i(j,k)} = \frac{1}{n_i - 2}(n_i \tilde{\mathbf{x}}_i - \mathbf{x}_{ij} - \mathbf{x}_{ik}),\ i = 1, 2;\ j, k = 1, 2, ..., n_i\ ,$$

$$\tilde{\mathbf{x}}_{2(\mathrm{k})} = \frac{1}{n_i}(n_i \tilde{\mathbf{x}}_i - \mathbf{x}_{ik}),\ i = 1, 2;\ k = 1, 2, ..., n_i$$

The test using the T_{CQ} statistic will reject the null hypothesis at a significant level of α, when the value of $T_{CQ} > Z_\alpha$ where Z_α is the upper α quantile of $N(0, 1)$.

The assumptions of T_{CQ} are

(1) $n_1/(n_1+n_2) \to k \in (0,1)$ where $n \to \infty$

(2) $(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)'\Sigma_i(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2) = o(n^{-1}tr\left\{(\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2)^2\right\}),\ i=1,2$

In addition, the test proposed by **Chen, S. X. & Qin, Y. L. (2010)** has the property of being invariant under orthogonal transformation, but it is not invariant under scalar transformation. Furthermore, the calculation to obtain the test statistic can be quite complex.

2.2. Modified Chen and Qin's Test

Although the test T_{CQ} in (1) is more widely applicable than the test of **Bai, Z. & Saranadasa, H. (1996)** because it does not require the same covariance matrices for both populations, the estimator $\hat{\sigma}^2_{CQ}$ in (3) is still not very efficient. **Srivastava, M. S. et al. (2013)** therefore suggested using the UMVUE (Uniformly minimum variance unbiased estimator) under the normal distribution of $tr(\boldsymbol{\Sigma}_i^2)/p$ instead of $\hat{\sigma}^2_{CQ}$ using in (3), which would improve the efficiency of the test. The estimator proposed by **Srivastava, M. S. et al. (2013)** is as follows:

$$\hat{\sigma}^2_U = \frac{2}{n_1^2}\hat{a}_{21} + \frac{2}{n_2^2}\hat{a}_{22} + \frac{4}{pn_1n_2}tr(\mathbf{S}_1\mathbf{S}_2),$$

where $$\hat{a}_{2i} = \frac{(n_1-1)^2}{pn_1(n_1-2)}\left[tr(\mathbf{S}_i^2) - \frac{1}{n_i-1}\left(tr(\mathbf{S}_i)\right)^2\right],\ i=1,2$$

The method proposed by **Srivastava, M. S. et al. (2013)** modifies the estimator of the covariance matrix in the test statistic of **Chen, S. X. & Qin, Y. L. (2010)** to improve its performance. The modified test statistic, denoted by T_{MCQ}, is defined as follows:

$$T_{MCQ} = \frac{Q_n}{\hat{\sigma}_{Q_n}} \tag{4}$$

where $$Q_n = (\bar{\mathrm{x}}_1-\bar{\mathrm{x}}_2)'(\bar{\mathrm{x}}_1-\bar{\mathrm{x}}_2) - \frac{tr(\mathbf{S}_1)}{n_1} - \frac{tr(\mathbf{S}_2)}{n_2},$$

$$\hat{\sigma}^2_{Q_n} = \frac{2}{n_1(n_1-1)}tr(\boldsymbol{\Sigma}_1^2) + \frac{2}{n_2(n_2-1)}tr(\boldsymbol{\Sigma}_2^2) + \frac{4}{n_1n_2}tr(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2)$$

$$tr(\boldsymbol{\Sigma}_i^2) = \frac{(n_i-1)^2}{(n_i+1)(n_i-2)}\left\{tr(\mathbf{S}_i^2) - \frac{1}{n_i-1}\left(tr(\mathbf{S}_i)\right)^2\right\}$$

$$\mathbf{S}_i = \frac{1}{n_i-1}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)'(x_{ij}-\bar{x}_i)\ ,\ i=1,2$$

$$tr(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2) = tr(\mathbf{S}_1\mathbf{S}_2)$$

The test using the T_{MCQ} statistic will reject the null hypothesis at a significant level of α, when the value of $T_{MCQ} > Z_\alpha$ where Z_α is the upper α quantile of $N(0, 1)$.

Although the T_{MCQ} test has fewer initial assumptions than the **Bai, Z. & Saranadasa, H. (1996)** test, its weakness is that it is only invariant under orthogonal transformations and not invariant under scalar transformations. In addition, the calculation to find the test statistic is quite complex.

2.3. Srivastava, Katayama and Kano's Test

One important testing method proposed by **Srivastava, M. S. et al. (2013)** is the T_{MCQ} test, which has the property of being invariant under scalar transformation. The T_{SKK} test statistic is as follows:

$$T_{SKK} = \frac{(\bar{x}_1 - \bar{x}_2)'\hat{\mathbf{D}}^{-1}(\bar{x}_1 - \bar{x}_2) - p}{\left[pVar(\hat{q}_n)(1 + \frac{tr(\mathbf{R}^2)}{p^{3/2}}) \right]}, \tag{5}$$

where $\hat{\mathbf{D}} = \frac{\hat{\mathbf{D}}_1}{n_1} + \frac{\hat{\mathbf{D}}_2}{n_2}$, $\hat{\mathbf{D}}_i = diag(s_{i11}, s_{i22}, ..., s_{ipp})$, $i = 1, 2$,

s_{ikk} are elements on the main diagonal of the pooled sample covariance matrix $\mathbf{S}_i$,

$$Var(\hat{q}_n) = \frac{2tr(\mathbf{R}^2)}{p} - \frac{2\left[tr(\hat{\mathbf{D}}^{-1}\mathbf{S}_1)\right]^2}{pn_1^2(n_1 - 1)} - \frac{2\left[tr(\hat{\mathbf{D}}^{-1}\mathbf{S}_2)\right]^2}{pn_2^2(n_2 - 1)},$$

and $\mathbf{R} = \hat{\mathbf{D}}^{-1/2}\left(\frac{\mathbf{S}_1}{n_1} + \frac{\mathbf{S}_2}{n_2}\right)\hat{\mathbf{D}}^{-1/2}$

The test using the T_{SKK} statistic will reject the null hypothesis at a significant level of α, when the value of $T_{SKK} > Z_{1-\alpha}$ where Z is distributed as a standard normal distribution.

The assumptions of T_{SKK} are

(1) $0 < c_1 < \min_{1 \le k \le p} \sigma_{ikk} \le \max_{1 \le k \le p} \sigma_{ikk} < c_2 < \infty$

(2) $\lim_{p \to \infty} tr(\mathrm{R}^4)/[\mathrm{tr}(\mathrm{R}^2)]^2 = 0 \;\; n \to \infty$

(3) $n_1/(n_1 + n_2) \to k \in (0,1)$, where $n = n_1 + n_2 \to \infty$

(4) $n_{\min} = O(p^{\delta}),\ \delta > 1/2,\ n_{\min} = \min(n_1, n_2)$

The important advantage that both T_{SKK} and T_{MCQ} tests have in common is that they do not rely on any assumptions about the distribution of the population, meaning that they can be applied to non-normally distributed data.

In addition to the above testing methods, there are several other methods proposed for testing the equality of mean vectors for two high-dimensional populations, such as **Sukcharoen, P. & Chongcharoen, S. (2019)** who proposed a method for testing the equality of mean vectors for two populations in cases where the joint covariance matrix of the two populations is not the same, by utilizing the concept of retaining information from the joint covariance matrix of the sample as much as possible. The testing method developed by **Jiamwattanapong, K. & Chongcharoen, S. (2017)** is based on the case where the joint covariance matrix of the two populations is equal. The testing statistic proposed by **Sukcharoen, P. & Chongcharoen, S. (2019)** has the property of being invariant under scalar transformation and has an approximate distribution when the number of variables is large. The proposed method is more efficient than other methods, especially when the joint covariance matrix of the sample can be arranged into a block diagonal matrix. However, the weakness of this method is that the data must have a normal distribution and the process of arranging the variables that are related into the same block takes a long time.

Based on the literature review above, the present study aims to investigate the performance of two tests: the modified method of **Chen, S. X. & Qin, Y. L. (2010)** (T_{MCQ}) and the method of **Srivastava, M. S. et al. (2013)** (T_{SKK}).

3. Simulation Procedure

3.1. Phase 1 Simulation

Let the first sample $\mathbf{x}_{11}, \mathbf{x}_{12}, ..., \mathbf{x}_{1n_1}$ be drawn from a p-variable normal population $N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1^{(s)})$, where s = 1, 2, …, 5 representing different covariance structures: 1) Sphericity, 2) Compound Symmetric (CS), Heterogeneous Compound Symmetry (CSH), 4) Toeplitz, and 5) Block Diagonal (BD). The second sample $\mathbf{x}_{21}, \mathbf{x}_{22}, ..., \mathbf{x}_{2n_2}$, which is independent of the first one, is drawn from $N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2^{(s)})$, where $p > n$, n: $n_1 + n_2$

- 2 and $n_1 = n_2$. Five covariances structured were formed as follows:

1) Sphericity

$$\mathbf{\Sigma}_k^{(1)} = \left(\sigma_{ij}\right), \sigma_{ij} = \begin{cases} c & ,i = j \\ 0 & ,i < j \end{cases} \quad ,k = 1,2, \tag{6}$$

where $\mathbf{\Sigma}_1^{(1)}$ and $\mathbf{\Sigma}_2^{(1)}$ are formed with $c = 1$ and $c = 2$ respectively

2) Compound Symmetry (CS)

$$\mathbf{\Sigma}_k^{(2)} = cI_p + (1-c)1_p 1_p' \quad ,k = 1,2, \tag{7}$$

where 1_p is a px1 vector in which all entries are 1's,

$\mathbf{\Sigma}_1^{(2)}$ and $\mathbf{\Sigma}_2^{(2)}$ are formed with $c = 1$ and $c = 2$ respectively

3) Heterogeneous Compound symmetry (CSH)

$$\mathbf{\Sigma}_k^{(3)} = \left(\sigma_{ij}\right),\ \sigma_{ij} = \begin{cases} \sigma_{ij} \sim U(c, c+1) & ,i = j \\ 0.5\sqrt{\sigma_{ii}\sigma_{jj}} & ,i < j \end{cases} \quad ,k = 1,2, \tag{8}$$

where $\mathbf{\Sigma}_1^{(3)}$ and $\mathbf{\Sigma}_2^{(3)}$ are formed with $c = 1$ and $c = 3$ respectively,

$U(c, c+1)$ is a continuous uniform distribution

4) Toeplitz

$$\mathbf{\Sigma}_k^{(4)} = \left(\sigma_{ij}\right),\ \sigma_{ij} = \begin{cases} 1 & ,i = j \\ \sigma_{i(i+1)} = c & ,i = 1,2,\ldots,p-1 \\ 0 & ,\text{others} \end{cases} \quad ,k = 1,2 \tag{9}$$

where $\mathbf{\Sigma}_1^{(4)}$ and $\mathbf{\Sigma}_2^{(4)}$ are formed with $c = 0.1$ and $c = 0.2$ respectively

5) Block Diagonal Matrix (BD)

$$\mathbf{\Sigma}_k^{(5)} = diag(\mathbf{\Sigma}_{11}, \mathbf{\Sigma}_{22}, \ldots, \mathbf{\Sigma}_{(m-1)(m-1)}, \mathbf{\Sigma}_{mm}), \tag{10}$$

where $\mathbf{\Sigma}_{ii} = cI_p + (1-c)1_p 1_p'$ with the sizes of $\mathbf{\Sigma}_{11}, \Sigma_{22}, \ldots, \mathbf{\Sigma}_{(m-1)(m-1)}$ are equal to $n-6$ and the rest is the size of $\mathbf{\Sigma}_{mm}$,

$\mathbf{\Sigma}_1^{(5)}$ and $\mathbf{\Sigma}_2^{(5)}$ are formed with $c = 0.3$ and $c = 0.6$ respectively.

Under the alternative hypothesis, set the population mean vectors as $\boldsymbol{\mu}_1 = \mathbf{0}$ and $\boldsymbol{\mu}_2 = (\delta_1, \ldots, \delta_p)'$, where $\delta_{2k-1} = 0$, and $\delta_{2k} \sim U(-0.4, 0.6),\ k = 1, \ldots, p/2$.

Set the sample size (n_i) and the number of variables (p) as follows: at $n_1 = n_2$ = 20, set p = {50, 100, 200, 300, 400}; at $n_1 = n_2$ = 40, set p = {100, 200, 300, 400}; and at $n_1 = n_2$ = 60, set p = {150, 200, 300, 400}. Each condition was iterated 1,000 times with a nominal significance level (α) of 0.05.

The performance of a test refers to the criteria used to determine which test provides an attained significance level (ASL) close to the specified nominal significance level and if the ASL of a test falls within an acceptable range, the test is considered to have acceptable power. Tests that have higher empirical power are considered to be more efficient. In studying the performance of a testing method, there are criteria used for evaluation, including the ASL value which follows Cochran's criterion **(Cochran, W. G., 1954)**, or Bradley's liberal criterion **(Bradley, J. V., 1978)** for the predetermined significance level (α). If the ASL value falls within the range of [0.040, 0.060], then the testing method is considered to have good performance or acceptable in terms of Type I error probability. Additionally, if the ASL value falls within an acceptable range, the test with higher empirical power is generally considered to be better. Finally, if the empirical power value approaches 1, the test is considered to have high testing power.

3.2. Phase 2 Simulation

Phase 2 aims to assess the performance of the tests under conditions where the covariance matrices of the second population differ more from those of the first population. In Phase 1, we discovered that T_{MCQ} displayed an acceptable level of performance for the covariance structures of Sphericity, Toeplitz, and BD. Similarly, T_{SKK} was found to perform acceptably for the covariance structures of CS, Toeplitz, and BD.

To simulate this scenario, we set the population mean vectors to be identical to those in Phase 1. However, we altered the covariance matrix of the second population to be more distinct by modifying the constant c forming the covariance matrix. Specifically, for Sphericity and CS, the constant c was adjusted from 2 to 3. For Toeplitz, it was adjusted from 0.2 to 0.3, and for BD, it was adjusted from 0.6 to 0.9. Each condition was iterated 1,000 times with a nominal significance level (α) of 0.05.

4. Results

The results are presented in two parts. The first part includes the results from Phase 1, where the performance of two tests, T_{MCQ} and T_{SKK}, was compared under a *p*-variate normal distribution with unequal population covariance matrices and equal sample sizes. Tables 1 and 2 show the results of this comparison. The second part includes the results from Phase 2, which investigated the performance of the two tests under different population covariance matrices. Phase 2 focused only on the covariance structures where the tests performed well, as determined by the results from Phase 1. The results from Phase 2 are presented in Tables 3 and 4.

Table.1. ASL of T_{MCQ} and T_{SKK} under five covariance structures with nominal level 0.05

n_i	p	Sphericity		CS		CSH		Toeplitz		BD	
		T_{MCQ}	T_{SKK}	T_{MCQ}	T_{SKK}	T_{MCQ}	T_{SKK}	T_{MCQ}	T_{SKK}	T_{MCQ}	T_{SKK}
20	50	0.0638	0.0918	0.0726	0.0632	0.0726	0.0370	0.0640	0.0850	0.0586	0.0742
	100	0.0604	0.1046	0.0804	0.0644	0.0830	0.0354	0.0634	0.0956	0.0662	0.0882
	200	0.0586	0.1172	0.0720	0.0518	0.0786	0.0276	0.0580	0.1050	0.0566	0.0992
	300	0.0546	0.1316	0.0770	0.0488	0.0696	0.0182	0.0610	0.1204	0.0606	0.1070
	400	0.0530	0.1498	0.0868	0.0512	0.0730	0.0194	0.0538	0.1292	0.0560	0.1128
40	100	0.0568	0.0704	0.0744	0.0532	0.0752	0.0276	0.0626	0.0704	0.0570	0.0622
	200	0.0574	0.0788	0.0756	0.0502	0.0652	0.0190	0.0528	0.0654	0.0604	0.0694
	300	0.0546	0.0804	0.0756	0.0450	0.0656	0.0162	0.0520	0.0714	0.0528	0.0728
	400	0.0506	0.0792	0.0752	0.0410	0.0640	0.0118	0.0582	0.0804	0.0554	0.0746
60	150	0.0584	0.0652	0.0718	0.0464	0.0698	0.0186	0.0516	0.0576	0.0558	0.0540
	200	0.0566	0.0646	0.0752	0.0476	0.0760	0.0172	0.0552	0.0648	0.0572	0.0638
	300	0.0684	0.0812	0.0728	0.0424	0.0690	0.0114	0.0556	0.0678	0.0550	0.0624
	400	0.0540	0.0690	0.0684	0.0380	0.0724	0.0104	0.0500	0.0636	0.0592	0.0672

The findings presented in Table 1 demonstrate that the ASL values for both T_{MCQ} and T_{SKK} were influenced by the choice of covariance structure. The T_{MCQ} exhibited acceptable performance and outperformed the T_{SKK} for the Sphericity, Toeplitz, and BD covariance structures, while the T_{SKK} was found to be superior to T_{MCQ} for the CS covariance structure. Notably, the T_{SKK} showed satisfactory performance for Toeplitz and BD covariance structures only when the sample size was at least 60. It is also evident that the performance of T_{SKK} improves with increasing sample size. Overall, the T_{MCQ} was observed to outperform T_{SKK} for the Toeplitz and BD covariance structures for small sample sizes. However, both tests showed poor performance for the CSH covariance structure.

Table.2. Empirical power of T_{MCQ} and T_{SKK} under five covariance structures with nominal level 0.05

n_i	p	Sphericity		CS		CSH		Toeplitz		BD	
		T_{MCQ}	T_{SKK}	T_{MCQ}	T_{SKK}	T_{MCQ}	T_{SKK}	T_{MCQ}	T_{SKK}	T_{MCQ}	T_{SKK}
20	50	0.4428	0.5034	0.2780	0.2402	0.1070	0.0524	0.6354	0.6716	0.5884	0.6084
	100	0.6384	0.7122	0.2900	0.2320	0.1150	0.0500	0.8576	0.8870	0.8144	0.8410
	200	0.8544	0.9122	0.2888	0.2044	0.1048	0.0346	0.9760	0.9844	0.9668	0.9782
	300	0.9522	0.9772	0.2796	0.1782	0.1048	0.0312	0.9990	0.9992	0.9946	0.9970
	400	0.9826	0.9928	0.2774	0.1668	0.1004	0.0268	1.0000	1.0000	0.9992	0.9994
40	100	0.9622	0.9662	0.7280	0.5606	0.1394	0.0516	0.9974	0.9974	0.9962	0.9952
	200	0.9990	0.9994	0.7884	0.5284	0.1362	0.0364	1.0000	1.0000	1.0000	1.0000
	300	0.9998	0.9998	0.8084	0.5012	0.1296	0.0300	1.0000	1.0000	1.0000	1.0000
	400	1.0000	1.0000	0.8262	0.4798	0.1420	0.0292	1.0000	1.0000	1.0000	1.0000
60	150	0.9998	0.9998	0.9866	0.9132	0.1804	0.0466	1.0000	1.0000	1.0000	1.0000
	200	1.0000	1.0000	0.9946	0.9142	0.1750	0.0448	1.0000	1.0000	1.0000	1.0000
	300	1.0000	1.0000	0.9974	0.9122	0.1820	0.0402	1.0000	1.0000	1.0000	1.0000
	400	1.0000	1.0000	0.9990	0.9066	0.1874	0.0396	1.0000	1.0000	1.0000	1.0000

Based on Table 1, the ASL values of T_{MCQ} were found to be within an acceptable range for the covariance structures of Sphericity, Toeplitz, and BD. Table 2 presents the results that show for these covariance structures, the empirical power of T_{MCQ} was satisfactory. For the covariance structures of CS, Toeplitz, and BD where the ASL values of T_{SKK} performed acceptably, the empirical power of T_{SKK} was high. Both tests showed an increase in empirical power as the sample size increased. In addition, T_{SKK} exhibited slightly higher empirical power than T_{MCQ} for the covariance structures of Toeplitz, and BD.

Table.3. ASL and empirical power of T_{MCQ} in Phase 2 with nominal level 0.05

n_i	p	Sphericity		Toeplitz		BD	
		ASL	Empirical Power	ASL	Empirical Power	ASL	Empirical Power
20	50	0.0660	0.3224	0.0670	0.6378	0.0618	0.6072
	100	0.0540	0.4804	0.0562	0.8548	0.0602	0.8372
	200	0.0578	0.6978	0.0514	0.9820	0.0558	0.9794
	300	0.0552	0.8428	0.0546	0.9974	0.0596	0.9962
	400	0.0564	0.9102	0.0522	0.9996	0.0562	0.9996
40	100	0.0558	0.8682	0.0584	0.9980	0.0608	0.9982
	200	0.0576	0.9844	0.0536	1.0000	0.0608	1.0000
	300	0.0556	0.9982	0.0550	1.0000	0.0586	1.0000
	400	0.0590	1.0000	0.0496	1.0000	0.0512	1.0000
60	150	0.0526	0.9978	0.0574	1.0000	0.0616	1.0000
	200	0.0564	0.9992	0.0598	1.0000	0.0622	1.0000
	300	0.0562	1.0000	0.0540	1.0000	0.0624	1.0000
	400	0.0536	1.0000	0.0560	1.0000	0.0540	1.0000

Table 3 presents the ASL and empirical power of the T_{MCQ} test in Phase 2, where the two population covariance matrices differed more, with a nominal level of 0.05. In this phase, the covariance structures (Sphericity, Toeplitz, and BD) were examined because the results from Phase 1 indicated that the T_{MCQ} test performed acceptably only for these situations. The ASL values range from 0.0496 to 0.0670, indicating that the test has good performance. The empirical power increases with an increase in the sample size, indicating that the test has high power to detect significant differences. Overall, the T_{MCQ} test still performs well even when the covariance matrices between two populations differ more.

Table.4. ASL and empirical power of T_{SKK} in Phase 2 with nominal level 0.05

n_i	p	CS		Toeplitz		BD	
		ASL	Empirical Power	ASL	Empirical Power	ASL	Empirical Power
20	50	0.0642	0.3124	0.0858	0.6620	0.0786	0.6388
	100	0.0668	0.3388	0.0872	0.8844	0.0864	0.8662
	200	0.0642	0.3210	0.1036	0.9886	0.1006	0.9876
	300	0.0612	0.3060	0.1110	0.9988	0.1168	0.9974
	400	0.0524	0.2948	0.1182	0.9998	0.1252	1.0000
40	100	0.0566	0.7686	0.0652	0.9980	0.0658	0.9978
	200	0.0532	0.8048	0.0684	1.0000	0.0746	1.0000
	300	0.0504	0.7974	0.0754	1.0000	0.0780	1.0000
	400	0.0494	0.7914	0.0714	1.0000	0.0764	1.0000
60	150	0.0490	0.9880	0.0638	1.0000	0.0668	1.0000
	200	0.0548	0.9936	0.0638	1.0000	0.0674	1.0000
	300	0.0470	0.9934	0.0648	1.0000	0.0726	1.0000
	400	0.0436	0.9932	0.0698	1.0000	0.0668	1.0000

Table 4 presents the ASL and empirical power of the T_{SKK} test in Phase 2 with a nominal level of 0.05. The study investigated three covariance structures (CS, Toeplitz, and BD) since the results from Phase 1 indicated that the T_{SKK} test performed acceptably only for these situations. The ASL values ranged from 0.0436 to 0.0548 for the CS covariance structure, indicating that the T_{SKK} test performed well in this type of covariance structure. However, the ASL values were slightly higher than 0.05 for the Toeplitz and BD covariance structures, indicating that the T_{SKK} test performed worse when the two population covariance matrices differed more for these covariance structures. The empirical power of the test also increased with an increase in the sample size.

5. Conclusion and Discussion

5.1. Conclusion

For two independent high-dimensional datasets with multivariate normal distributions, where the population covariance matrices are unknown and unequal but have the same covariance structure, we studied the case where the sample sizes for both datasets are equal ($n_1 = n_2$) and fall in the range of 20-60, with the number of variables (p) not exceeding 400 and p > n, where $n = n_1 + n_2 - 2$. We divided the results into two phases.

The results from Phase 1 showed that the choice of covariance structure had an impact on the performance of both tests. The T_{MCQ} performed acceptably for the Sphericity, Toeplitz, and BD covariance structures. The T_{SKK} performed acceptably for the CS covariance structure, and for the covariance structures of Toeplitz and BD, the test performed well only when the sample size was at least 60. the T_{MCQ} is more efficient than the T_{SKK} under the covariance structures of Sphericity, Toeplitz, and BD, while the T_{SKK} is more efficient than T_{MCQ} under the CS covariance structure. However, the performance of the T_{SKK} improves with increasing sample size, particularly when the sample size is at least 60. Both tests showed increased empirical power as the sample size increased. Under the covariance structure of CSH, it was found that both the T_{MCQ} and T_{SKK} are still underperforming.

In Phase 2, we examined the performance of each test in detecting the difference between two mean vectors when the difference between the two covariance matrices was increased. It can be concluded that T_{MCQ} continued to perform well for the covariance structures of Sphericity, Toeplitz, and BD. The empirical power of T_{MCQ} increased with larger sample sizes. The T_{SKK} continued to perform well for the CS structure of all cases studied, and the test performed acceptably under the Toeplitz, and BD, particularly for sample size was at least 60. Nevertheless, the T_{SKK} still demonstrated high power with larger sample sizes.

Based on the results, we recommend the following:

- For Sphericity, Toeplitz, or BD covariance structures, either T_{MCQ} or T_{SKK} can be used when the sample size is at least 60. For smaller sample sizes, use T_{MCQ}.
- For Compound Symmetry covariance structure, use T_{SKK}.

- For Heterogeneous Compound Symmetry covariance structures, use caution or consider using other testing methods, as neither T_{MCQ} nor T_{SKK} performed well in this case.
- For other types of covariance structures, experiment with various tests and evaluate their conclusions before making a decision.

Overall, our findings confirm that the choice of covariance structure affects the performance of the tests. Both T_{MCQ} and T_{SKK} showed increased power of the test with larger sample sizes, and T_{MCQ} generally performed well for Sphericity, Toeplitz, and BD covariance structures, while T_{SKK} performed well for the CS structure. However, caution is advised when dealing with Heterogeneous Compound Symmetry structures, as neither test performed well in this case.

5.2. Discussion

The T_{MCQ} test statistic proposed by **Chen, S. X. & Qin, Y. L. (2010)** was found to be ineffective when applied to test the mean vector of two populations under the compound symmetry (CS) and heterogeneous compound symmetry (CSH) structures of the covariance matrix. This may be because the data used in this study did not conform to the underlying assumptions of the test, which are difficult to verify. The crucial assumption of this test is that $(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_i(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)=o(n^{-1}tr\left\{(\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2)^2\right\})$. If the data used for testing do not meet this assumption, the test will be less effective than it should be. This finding is consistent with the results of **Srivastava et al. (2013)**.

The study found that the T_{SKK} test, proposed by **Srivastava, M. S. et al. (2013)**, demonstrated higher overall performance when each sample size was 60 or greater. These results are consistent with a previous simulation study by **Srivastava, M. S. et al. (2013)**, which assumed a Sphericity structure for the covariance matrix of both populations and studied sample sizes of 30 or more (with a number of variables over 60). The T_{SKK} test was also found to be equally effective in that study.

Acknowledgments

We are grateful for the support provided by Rajamangala University of Technology Rattanakosin (RMUTR) through the research funds that enabled this study. In addition, we would like to express my appreciation to Professor Samruam Chongcharoen for his invaluable advice throughout this research project. His contribution has been critical to the success of this study.

References

Alon, U., Barkai, N., Notterman, D. A., Gish, K., Ybarra, S., Mack, D., & Levine, A. J. (1999). Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proceedings of the National Academy of Sciences, 96(12), 6745-6750.

Bai, Z., & Saranadasa, H. (1996). Effect of high dimension: by an example of a two sample problem. Statistica Sinica, 6(2), 311-329.

Bradley, J. V. (1978). Robustness? In the smoothing of data. In Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Vol. 1, pp. 215-236). University of California Press.

Bühlmann, P., & Van De Geer, S. (2011). Statistics for high-dimensional data: Methods, theory and applications. Springer Science & Business Media.

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

Cochran, W. G. (1954). Some methods for strengthening the common chi-square tests. Biometrics, 10(4), 417-451.

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of the American Statistical Association, 96(456), 1348-1360.

Gelper, S., Friedl, H., & Croux, C. (2017). A primer on partial least squares structural equation modeling (PLS-SEM). Sage Publications.

Hotelling, H. (1931). The generalization of Student's ratio. Annals of Mathematical Statistics, 2(3), 360-378.

Hu, J., Bai, Z., Wang, C., & Wang, W. (2017). A new high-dimensional test for the mean vector with unequal covariance matrices. Journal of Multivariate Analysis, 153, 113-127.

Jiamwattanapong, K., & Chongcharoen, S. (2017). Two-sample test for high-dimensional mean vectors with unequal covariance matrices. Journal of Multivariate Analysis, 155, 128-146.

MIT Technology Review. (2014, February 26). The world's 50 smartest companies. MIT Technology Review.

Schott, J. R. (2005). Testing for equality of multiple covariance matrices. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67(1), 241-253.

Schreiber, J. B. (2017). Understanding the interaction between measurement and analysis: An illustration using the Partial Least Squares method. International Journal of Business Analytics, 4(2), 240-258.

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

Srivastava, M. S., Katayama, S., & Kano, Y. (2013). A new test for equality of covariance matrices in high-dimensional settings. Journal of Multivariate Analysis, 116, 456-470.

Sukcharoen, P., & Chongcharoen, S. (2019). A test for equality of high-dimensional mean vectors with unequal covariance matrices. Communications in Statistics-Simulation and Computation, 48(3), 1046-1060.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 58(1), 267-288.