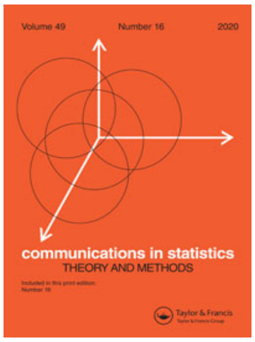

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

Testing equality of two normal covariance matrices with monotone missing data

Jianqi Yu, Kalimuthu Krishnamoorthy & Yafei He

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2020, VOL. 49, NO. 16, 3911–3918
https://doi.org/10.1080/03610926.2019.1591453

Taylor & Francis
Taylor & Francis Group

Testing equality of two normal covariance matrices with monotone missing data

Jianqi Yu[a], Kalimuthu Krishnamoorthy[b], and Yafei He[a]

[a]Guilin University of Technology, Guilin, Guangxi, China; [b]Department of Mathematics, University of Louisiana at Lafayette, Lafayette, Louisiana, USA

ABSTRACT
The problem of testing equality of two multivariate normal covariance matrices is considered. Assuming that the incomplete data are of monotone pattern, a quantity similar to the Likelihood Ratio Test Statistic is proposed. A satisfactory approximation to the distribution of the quantity is derived. Hypothesis testing based on the approximate distribution is outlined. The merits of the test are investigated using Monte Carlo simulation. Monte Carlo studies indicate that the test is very satisfactory even for moderately small samples. The proposed methods are illustrated using an example.

ARTICLE HISTORY
Received 22 September 2018
Accepted 2 March 2019

KEYWORDS
Likelihood ratio test; maximum likelihood estimators; missing data; monotone pattern; powers; sizes

MSC2000 SUBJECT CLASSIFICATIONS
62H12; 62H15

1. Introduction

The problem of Missing data arises very commonly in many practical situations. It arises, for example, during data gathering and recording, when the experiment involves a group of individuals over a period of time like in clinical trials or in a planned experiment which collects variables that are expensive to measure are collected only from a subset of a sample. The causes for missing data are not our concern, but to ignore the process that causes missing data, it is assumed that the data are missing at random (MAR). Lu and Copas (2004) pointed out that inference from the likelihood method ignoring the missing data mechanism is valid if and only if the missing data mechanism is MAR. For formal definition and exposition of MAR we refer to Little and Rubin (1987) or Little (1988).

There are a few missing patterns considered in the literature, but the incomplete data with monotone pattern (see display 1) not only occurs frequently in practice but also it is convenient for making inference. Moreover, if multivariate normality is assumed then the monotone pattern allows the exact calculation of the maximum likelihood estimators (MLEs), the likelihood ratio statistics and relevant distributions. Several authors have considered the monotone missing pattern under the normality assumption, and provided asymptotic as well as approximate test procedures about the normal mean vector. Anderson (1957), one of the earliest papers in this area, gives a simple approach to derive the MLEs and present them for a special case of monotone pattern and some

CONTACT Jianqi Yu 498616729@qq.com Guilin University of Technology, Guilin, Guangxi 540041, China.

other patterns. Kanda and Fujikoshi (1998) studied some basic properties of the MLEs based on monotone data. Many authors, like Bhargava (1962), Morrison and Bhoj (1973) and Naik (1975), developed asymptotic inferential procedures based on the likelihood ratio approach for the multivariate normal distribution. Krishnamoorthy and Pannala (1998, 1999) provided an accurate, simple approach to construct a confidence region for a normal mean vector. Hao and Krishnamoorthy (2001) developed an inferential procedure on a normal covariance matrix. Yu, Krishnamoorthy, and Pannala (2006) considered the problem of testing equality of two normal mean vectors with the assumption that the two covariance matrices are equal, while Krishnamoorthy and Yu (2012) considered the same problem of testing equality of two normal mean vectors without that assumption. Hence, an interesting problem arises: how to tell whether or not two normal covariance matrices are equal with monotone missing data. This is just the problem we considered in this article.

To formulate the problem, let $\boldsymbol{x}$ follow a p-variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. We write this as $\boldsymbol{x}\sim N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$. Let $\boldsymbol{y}\sim N_p(\boldsymbol{\beta},\boldsymbol{\Delta})$ independently of $\boldsymbol{x}$. Suppose that we have a sample of N_1 observations available on $\boldsymbol{x}$, and a sample of M_1 observations available on $\boldsymbol{y}$. Assume that the samples have the following monotone pattern:

$$\begin{array}{ll} x_{11},\ldots,x_{1N_k},\ldots,x_{1N_2},\ldots,x_{1N_1} & y_{11},\ldots,y_{1M_k},\ldots,y_{1M_2},\ldots,y_{1M_1} \\ x_{21},\ldots,x_{2N_k},\ldots,x_{2N_2} & y_{21},\ldots,y_{2M_k},\ldots,y_{2M_2} \\ \cdots\cdots & \cdots\cdots \\ x_{k1},\ldots,x_{kN_k} & y_{k1},\ldots,y_{kM_k} \end{array} \quad (1)$$

where x_{ij} is a $p_i\times 1$ vector, $j=1,\ldots N_i$, while y_{ij} is a $q_i\times 1$ vector, $j=1,\ldots M_i, i=1,\ldots,k$. In other words, in the x-sample, there are N_1 observations available on the first p_1 components, N_2 observations available on the first p_1+p_2 components, and so on. Notice that $N_1\geq N_2\geq\ldots\geq N_k, M_1\geq M_2\geq\ldots\geq M_k$, and $p_1+\ldots+p_k=q_1+\ldots+q_k=p$.

We want to test

$$H_0:\boldsymbol{\Sigma}=\boldsymbol{\Delta} \text{ vs. } H_a:\boldsymbol{\Sigma}\neq\boldsymbol{\Delta} \quad (2)$$

In the following section, we present first some preliminaries in the notations of Krishnamoorthy and Pannala (1998) for the data matrices in (1) with $k=2$ and $p_i=q_i$, $i=1$, 2. We present the MLEs of the relevant parameters in terms of these notations. As pointed in Yu, Krishnamoorthy and Pannala (2006), we do not need to consider the case of unequal monotone pattern, i.e, $p_i\neq q_i$, $i=1$, 2, since for any type of unequal monotone pattern the data can be rearranged to form an equal monotone pattern.

In Section 3, we developed inferential procedure for testing the equality of two normal covariance matrices with monotone missing pattern in (1) with $k=2$. We present a quantity similar to the Likelihood Ratio Test Statistic, and derive an approximation to its distribution. Then, the procedures for the hypothesis testing is outlined.

The accuracies of the approximation are verified by Monte Carlo simulation in Section 4. In Section 5, power comparisons between the test based on incomplete data and the test based on complete data obtained by discarding the extra data are made to demonstrate the advantage of keeping the extra data. Simulation studies show that the former is more powerful than the latter in all cases considered. The methods are illustrated using an example in Section 6.

2. Preliminaries

Consider the data matrices in (1) with $k=2$, $p_i = q_i$, $i=1$, 2, and partition the data matrices as follows:

$$\begin{aligned} \boldsymbol{x}_1 &= (x_{11}, \ldots, x_{1N_2}, \ldots, x_{1N_1})_{p_1 \times N_1} \\ \boldsymbol{x}_2 &= \begin{pmatrix} x_{11}, & \cdots & , x_{1N_2} \\ x_{21}, & \cdots & , x_{2N_2} \end{pmatrix}_{p \times N_2} \end{aligned} \tag{3}$$

That is, $\boldsymbol{x}_l$ is the submatrix of $\boldsymbol{x}$ in (1) formed by the first N_l columns and the first $p_1 + \ldots + p_l$ rows, $l = 1, 2$. Partition the matrix $\boldsymbol{y}$ similarly. That is,

$$\begin{aligned} \boldsymbol{y}_1 &= (y_{11}, \ldots, y_{1M_2}, \ldots, y_{1M_1})_{p_1 \times M_1} \\ \boldsymbol{y}_2 &= \begin{pmatrix} y_{11}, & \cdots & , y_{1M_2} \\ y_{21}, & \cdots & , y_{2M_2} \end{pmatrix}_{p \times M_2} \end{aligned} \tag{4}$$

Let $\bar{\boldsymbol{x}}_l$ and $\mathbf{S}_l$ denote respectively the sample mean vector and the sums of squares and products matrix based on $\boldsymbol{x}_l$, $l=1$, 2. Similarly, let $\boldsymbol{y}_l$ and $\mathbf{V}_l$ denote, respectively, the sample mean vector and the sums of squares and products matrix based on $\boldsymbol{y}_l$, $l=1$, 2. We partition these means and matrices accordingly as follows:

$$\bar{\boldsymbol{x}}_1 = \bar{x}_1^{(1)}, \ \bar{\boldsymbol{x}}_2 = \begin{pmatrix} \bar{x}_2^{(1)} \\ \bar{x}_2^{(2)} \end{pmatrix}, \ \mathbf{S}_1 = S_1^{(1,1)} \text{ and } \mathbf{S}_2 = \begin{pmatrix} S_2^{(1,1)} & S_2^{(1,2)} \\ S_2^{(2,1)} & S_2^{(2,2)} \end{pmatrix}$$

Notice that $\bar{x}_l^{(i)}$ is the mean of the ith block of the data matrix $\boldsymbol{x}_l, i = 1, \ldots, l$ and $l = 1, 2$. We also read $S_l^{(i,j)}$ as the (i, j)th component of $\mathbf{S}_l$ based on the data matrix $\boldsymbol{x}_l, l = 1, 2$.

The statistics $\overline{\boldsymbol{y}_l}$ and $\mathbf{V}_l$ based on the data matrix $\boldsymbol{y}_l$ in (4) are also partitioned like $\bar{\boldsymbol{x}}_l$ and $\mathbf{S}_l$. That is, $\bar{y}_l^{(i)}$ is the mean of the ith block of data matrix $\boldsymbol{y}_l, i = 1, \ldots, l$ and $l = 1, 2$, and $V_l^{(i,j)}$ is the (i, j)th component of $\mathbf{V}_l$, $i, j = 1, \ldots, l$ and $l=1$, 2.

Finally, we partition the parameters as follows:

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \ \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix}, \ \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \text{ and } \boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{\Delta}_{11} & \boldsymbol{\Delta}_{12} \\ \boldsymbol{\Delta}_{21} & \boldsymbol{\Delta}_{22} \end{pmatrix}$$

We have the following well known results:

$$\bar{\boldsymbol{x}}_1 = \bar{x}_1^{(1)} \sim N_{p_1}\left(\boldsymbol{\mu}_1, \frac{1}{N_1}\boldsymbol{\Sigma}_{11}\right), \ \overline{\boldsymbol{y}_1} = \bar{y}_1^{(1)} \sim N_{p_1}\left(\boldsymbol{\beta}_1, \frac{1}{M_1}\boldsymbol{\Delta}_{11}\right)$$

$$\mathbf{S}_1 = S_1^{(1,1)} \sim W_{p_1}(N_1 - 1, \boldsymbol{\Sigma}_{11}) \text{ and } \mathbf{V}_1 = V_1^{(1,1)} \sim W_{p_1}(M_1 - 1, \boldsymbol{\Delta}_{11}).$$

$$\bar{\boldsymbol{x}}_2 = \begin{pmatrix} \bar{x}_2^{(1)} \\ \bar{x}_2^{(2)} \end{pmatrix} \sim N_p\left(\boldsymbol{\mu}, \frac{1}{N_2}\boldsymbol{\Sigma}\right), \ \bar{\boldsymbol{y}}_2 = \begin{pmatrix} \bar{y}_2^{(1)} \\ \bar{y}_2^{(2)} \end{pmatrix} \sim N_p\left(\boldsymbol{\beta}, \frac{1}{M_2}\boldsymbol{\Delta}\right)$$

$$\mathbf{S}_2 = \begin{pmatrix} S_2^{(1,1)} & S_2^{(1,2)} \\ S_2^{(2,1)} & S_2^{(2,2)} \end{pmatrix} \sim W_p(N_2 - 1, \boldsymbol{\Sigma}), \ \mathbf{V}_2 = \begin{pmatrix} V_2^{(1,1)} & V_2^{(1,2)} \\ V_2^{(2,1)} & V_2^{(2,2)} \end{pmatrix} \sim W_p(M_2 - 1, \boldsymbol{\Delta})$$

Let

$$S_{2.1} = S_2^{(2,2)} - S_2^{(2,1)}\left(S_2^{(1,1)}\right)^{-1} S_2^{(1,2)}, \boldsymbol{\Sigma}_{2.1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$$
$$V_{2.1} = V_2^{(2,2)} - V_2^{(2,1)}\left(V_2^{(1,1)}\right)^{-1} V_2^{(1,2)}, \boldsymbol{\Delta}_{2.1} = \boldsymbol{\Delta}_{22} - \boldsymbol{\Delta}_{21}\boldsymbol{\Delta}_{11}^{-1}\boldsymbol{\Delta}_{12}$$

and

$$f_i = N_i - 1, g_i = M_i - 1, i = 1,2$$
$$\mathbf{P} = \mathbf{S}_2 + \mathbf{V}_2, P_{2.1} = P_{22} - P_{21}P_{11}^{-1}P_{12}$$

The statistics that we use to test $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Delta}$ vs. $H_a : \boldsymbol{\Sigma} \neq \boldsymbol{\Delta}$ is given by

$$\Lambda = -2(1-c_1)logT_1 - 2(1-c_2)logT_2 = \Lambda_1 + \Lambda_2$$

where

$$c_1 = \frac{\left(2p_1^2 + 3p_1 - 1\right)\left(\frac{1}{f_1} + \frac{1}{g_1} - \frac{1}{(f_1+g_1)}\right)}{6(p_1+1)}$$
$$c_2 = \frac{\left(2p^2 + 3p - 1\right)\left(\frac{1}{f_2} + \frac{1}{g_2} - \frac{1}{(f_2+g_2)}\right)}{6(p+1)}$$
$$T_1 = \frac{|S_1^{(1,1)}|^{f_1/2}|V_1^{(1,1)}|^{g_1/2}(f_1+g_1)^{p_1(f_1+g_1)/2}}{\left|\left(S_1^{(1,1)} + V_1^{(1,1)}\right)\right|^{(f_1+g_1)/2} f_1^{p_1 f_1/2} g_1^{p_1 g_1/2}}$$
$$T_2 = \frac{|S_{2.1}|^{f_2/2}|V_{2.1}|^{g_2/2}(f_2+g_2)^{p(f_2+g_2)/2}}{|P_{2.1}|^{(f_2+g_2)/2} f_2^{pf_2/2} g_2^{pg_2/2}}$$

The idea behind Λ is like this: if there are no additional observations on the first p_1 components, that is, there are only $N_2(M_2)$ observations on the first p_1 components of $X(Y)$, the appropriate test statistic should be (see Seber 1984, p. 449)

$$-2(1-c_2)logT$$

with

$$T = \frac{|\mathbf{S}_2|^{f_2/2}|\mathbf{V}_2|^{g_2/2}(f_2+g_2)^{p(f_2+g_2)/2}}{|(\mathbf{S}_2+\mathbf{V}_2)|^{(f_2+g_2)/2} f_2^{pf_2/2} g_2^{pg_2/2}}$$

Since

$$|\mathbf{S}_2| = |S_2^{(1,1)}| \times |S_{2.1}|$$
$$|\mathbf{V}_2| = |V_2^{(1,1)}| \times |V_{2.1}|$$
$$|(\mathbf{S}_2+\mathbf{V}_2)| = |\left(S_1^{(1,1)} + V_1^{(1,1)}\right)| \times |P_{2.1}|$$

T can be decomposed as product of two parts

$$\frac{|S_2^{(1,1)}|^{f_2/2}|V_2^{(1,1)}|^{g_2/2}(f_2+g_2)^{p(f_2+g_2)/2}}{\left|\left(S_2^{(1,1)} + V_2^{(1,1)}\right)\right|^{(f_2+g_2)/2} f_2^{pf_2/2} g_2^{pg_2/2}} \times T_2 = T_1^* \times T_2, say$$

It is easy to see that T_1^* is the appropriate statistic for testing $H_0 : \mathbf{\Sigma}^{(1,1)} = \mathbf{\Delta}^{(1,1)}$ vs. $H_a : \mathbf{\Sigma}^{(1,1)} \neq \mathbf{\Delta}^{(1,1)}$ if we have only $N_2(M_2)$ observations on the first p_1 components. Now we have more observations on the first p_1 components, T_1^* should be replaced by T_1.

3. Test for equal covariance matrices

To test $H_0 : \mathbf{\Sigma} = \mathbf{\Delta}$ vs. $H_a : \mathbf{\Sigma} \neq \mathbf{\Delta}$, we need to find the null distribution of Λ. Though the expression for Λ suggests that it is difficult to derive the exact null distribution of Λ, we could find out an approximate distribution for it.

From the well known properties of Wishart matrices we have the following results:

Let

$$\mathbf{W} = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} \sim W_p(n, \mathbf{\Sigma}), n > p > 1$$

where W_{11} is of order $p_1 \times p_1$ and W_{22} is of order $p_2 \times p_2$. Let $\mathbf{\Sigma}$ partitioned the same way and $\mathbf{\Sigma}_{2.1} = \mathbf{\Sigma}_{22} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}$, then

$$W_{2.1} = W_{22} - W_{21}W_{11}^{-1}W_{12} \sim W_{p_2}(n - p_1, \mathbf{\Sigma}_{2.1})$$

and $W_{2.1}$ is independnt of W_{11}, W_{12}

Proof See Seber (1984, Lemma 2.10, p. 50).

Firstly, it can be deduced from the above results that $S_2^{(1,1)}$ is independent of $S_{2.1}$, $V_2^{(1,1)}$ is independent of $V_{2.1}$. Hence, $(S_2^{(1,1)} + V_2^{(1,1)})$ is independent of $S_{2.1}$ and $V_{2.1}$. It is obvious that $(S_1^{(1,1)} + V_1^{(1,1)})$ is also independent of $S_{2.1}$ and $V_{2.1}$. Secondly, under $H_0 : \mathbf{\Sigma} = \mathbf{\Delta}, \mathbf{P} = \mathbf{S}_2 + \mathbf{V}_2 \sim W_p(f_2 + g_2, \mathbf{\Sigma})$. Thus $(S_2^{(1,1)} + V_2^{(1,1)}) = P^{(1,1)}$ is independent of $P_{2.1}$. Then again, it is obvious that $(S_1^{(1,1)} + V_1^{(1,1)})$ is also independent of $P_{2.1}$. So, we can conclude that $(S_1^{(1,1)} + V_1^{(1,1)})$ are independent of $S_{2.1}, V_{2.1}$, and $P_{2.1}$. Hence, under the null hypothesis, T_1^* and T_1 are independent of T_2.

Thirdly, we have approximately (See G.A.E Seber ([13], p. 449)

$$-2(1-c_2)logT \sim \chi^2(\nu)$$
$$-2(1-c_1)logT_1 \sim \chi^2(\nu_1)$$

with

$$c_1 = \frac{\left(2p_1^2 + 3p_1 - 1\right)\left(\frac{1}{f_1} + \frac{1}{g_1} - \frac{1}{(f_1 + g_1)}\right)}{6(p_1 + 1)}$$

$$c_2 = \frac{\left(2p^2 + 3p - 1\right)\left(\frac{1}{f_2} + \frac{1}{g_2} - \frac{1}{(f_2 + g_2)}\right)}{6(p + 1)}$$

$$\nu = \frac{1}{2}p(p+1), \nu_1 = \frac{1}{2}p_1(p_1 + 1)$$

Table 1. Critical values $\chi^2_{1-\alpha}(\nu)$ and Monte Carlo estimates of the size of the test with nominal level α (in parenthesis).

(N_1, N_2, M_1, M_2)	$(p_1 = 2, p_2 = 1)$ $\alpha = 0.05$	$\alpha = 0.01$	$(p_1 = 2, p_2 = 2)$ $\alpha = 0.05$	$\alpha = 0.01$
(18,13,12,8)	12.592 (0.0622)	16.811 (0.0134)	18.307 (0.0720)	23.209 (0.0164)
(18,13,20,15)	12.592 (0.0567)	16.811 (0.0120)	18.307 (0.0591)	23.209 (0.0128)
(25,17,25,18)	12.592 (0.0559)	16.811 (0.0111)	18.307 (0.0569)	23.209 (0.0122)
(30,20,35,20)	12.592 (0.0550)	16.811 (0.01205)	18.307 (0.0568)	23.209 (0.0164)
(40,25,30,10)	12.592 (0.0587)	16.811 (0.0124)	18.307 (0.0648)	23.209 (0.0149)
(40,15,30,10)	12.592 (0.0572)	16.811 (0.0126)	18.307 (0.0650)	23.209 (0.0144)
(40,30,40,32)	12.592 (0.0521)	16.811 (0.0102)	18.307 (0.0560)	23.209 (0.0113)

Since $-2(1-c_2)logT = (-2(1-c_2)logT_1') + (-2(1-c_2)logT_2)$, and $(-2(1-c_2)logT_1')$ is independent of $(-2(1-c_2)logT_2)$, we have

$$-2(1-c_2)logT_2 \sim \chi^2(\nu-\nu_1)$$

Hence,

$$\Lambda = -2(1-c_1)logT_1 - 2(1-c_2)logT_2 \sim \chi^2(\nu_1 + \nu - \nu_1) = \chi^2(\nu)$$

Thus, for a given level α and an observed value Λ_0 of Λ, the null hypothesis that $\mathbf{\Sigma} = \mathbf{\Delta}$ will be rejected whenever the p-value $P(\chi^2(\nu) > \Lambda_0) < \alpha$.

4. Accuracy of the approximations

We have used two approximations, one for approximating the distribution of T_1 and another for the distribution of T_2. So, to understand the accuracy of the approximation, we estimated the sizes of the test for hypotheses in (1) when nominal level is 0.05 or 0.01 using Monte Carlo simulation. Each simulation result is based on 100,000 runs.

To select the parameter configurations for Monte Carlo simulation, we note that the test is affine invariant, and so without loss of generality, we can assume that $\boldsymbol{\mu} = \boldsymbol{\beta} = 0$ and $\mathbf{\Sigma} = \mathbf{\Delta} = \mathbf{I}$ to estimate the sizes.

The estimated sizes are presented in Table 1 for two cases and a few selected sample sizes. One case is $p_1 = 2$ and $p_2 = 1$, and another case is $p_1 = p_2 = 2$. The sample sizes are chosen so that the number of data missing is relatively small in some cases, and large in other cases. It is clear from Table 1 that the estimated sizes are very close to nominal level for all the cases considered. But for very small samples, the sizes are not that accurate. In the worst situations, the estimated sizes are around 0.072 while the nominal level is 0.05.

5. Power study

To understand the nature of the powers of the test and the advantage of keeping additional data, we estimate the powers of the test based on incomplete data and powers based on" partially complete data" (the data obtained after deleting the last N_1-N_2 and M_1-M_2 observations on the first p_1 components) using simulation. The powers of the test based on partially complete data are given in parentheses. The powers are computed when $H_0 : \mathbf{\Sigma} = \mathbf{\Delta}$ and presented in Table 2.

Table 2. Powers of the test based on incomplete data and on partially complete data (in parentheses):$H_0 : \mathbf{\Sigma} = \mathbf{\Delta}$.

(N_1, N_2, M_1, M_2)	$(p_1 = 2, p_2 = 1)$ $\alpha = 0.05$ $\alpha = 0.01$	$(p_1 = 2, p_2 = 2)$ $\alpha = 0.05$ $\alpha = 0.01$
	$H_a : \mathbf{\Sigma} = I_p, \mathbf{\Delta} = diag(0.8, 0.6, 0.5)$	$H_a : \mathbf{\Sigma} = I_p, \mathbf{\Delta} = diag(0.5, 0.4, 0.2)$
(12,7,13,8)	0.110 (0.0765) 0.0284 (0.0169)	0.296 (0.221) 0.104 (0.0631)
(15,10,16,11)	0.129 (0.102) 0.0342 (0.0238)	0.468 (0.422) 0.213 (0.175)
(20,10,20,12)	0.146 (0.105) 0.0439 (0.0267)	0.549 (0.458) 0.284 (0.205)
(25,15,23,13)	0.179 (0.134) 0.0553 (0.0358)	0.685 (0.619) 0.413 (0.337)
(30,20,28,19)	0.235 (0.198) 0.0833 (0.0639)	0.866 (0.854) 0.668 (0.643)
(40,30,38,29)	0.355 (0.333) 0.153 (0.136)	0.981 (0.980) 0.921 (0.920)
	$H_a : \mathbf{\Sigma} = I_p, \mathbf{\Delta} = \begin{pmatrix} 2 & 0.4 & 0 \\ 0.4 & 1.5 & 0 \\ 0 & 0 & 1.3 \end{pmatrix}$	$H_a : \mathbf{\Sigma} = I_p, \mathbf{\Delta} = \begin{pmatrix} 2 & 0.4 & -0.3 \\ 0.4 & 1.5 & 0 \\ -0.3 & 0 & 1.3 \end{pmatrix}$
(12,7,13,8)	0.160 (0.0745) 0.0475 (0.0168)	0.166 0.0780) 0.0492 (0.0180)
(15,10,16,11)	0.193 (0.0981) 0.0595 (0.0228)	0.196 (0.101) 0.0631 (0.0247)
(20,10,20,12)	0.264 (0.100) 0.0961 (0.0248)	0.268 (0.106) 0.0995 (0.0258)
(25,15,23,13)	0.326 (0.134) 0.134 (0.0369)	0.335 (0.144) 0.141 (0.0400)
(30,20,28,19)	0.393 (0.196) 0.181 (0.0636)	0.409 (0.206) 0.189 (0.0685)
(40,30,38,29)	0.527 (0.322) 0.285 (0.131)	0.547 (0.345) 0.300 (0.143)

We observe from Table 2 that the powers of the tests are increasing as sample sizes increase; they are also increasing as $\mathbf{\Delta}$ moves away from $\mathbf{\Sigma}$. Thus, our proposed test possesses some natural power properties. Moreover, the powers of the test based on incomplete data are always larger than the corresponding powers based on partially complete data. This situation illustrates the advantage of keeping the extra data.

6. An illustrative example

We shall now illustrate the methods using the 'Fishers Iris Data' which represent measurements of the sepal length and width, and petal length and width in centimeters of fifty plants for each of three types of iris: Iris setosa, Iris versicolor and Iris virginica. The data sets are posted in many websites, and we downloaded them from http://javeeh.net/sasintro/intro151.html. For illustrative purposes, we use the data on virginica (x) and setosa (y).

Firstly, we created monotone patterns by discarding the last 18 measurements on x_3 (petal length of virginica) and x_4 (petal width of virginica), the last 20 measurements on y_3 (petal length of setosa) and y_4 (petal width of setosa). That is, we have $p_1 = 2, p_2 = 2, (N_1, N_2) = (50, 32), and (M_1, M_2) = (50, 30)$. Let $\mathbf{\Sigma}$ be covariance matrix of virginica, and $\mathbf{\Delta}$ be covariance matrix of setosa. We want to test

$$H_0 : \mathbf{\Sigma} = \mathbf{\Delta} \text{ vs. } H_a : \mathbf{\Sigma} \neq \mathbf{\Delta}$$

After careful calculation, we get $logT_1 = -71.602, logT_2 = -25.582$, and $\Lambda = 187.532$. The critical value $\chi^2_{0.95}(\nu) = 23.209$. Since Λ is much larger than the critical value, we have sufficient evidence to reject H_0 at 95% confidence level.

Secondly, we use only the partially complete data, that is, we delete the last 18 observations of virginica and 20 observations of setosa. Then, we get the $-2(1-c_2)logT =$

69.365. Since the critical value is also $\chi^2_{0.95}(\nu) = 23.209$, we have the same conclusion: to reject H_0 at 95% confidence level.

Also, as expected, the value of the test statistics based on incomplete data (187.532) is much larger than the one based on partially complete data (69.365). This illustrates the advantage of keeping the additional data too.

References

Anderson, T. W. 1957. Maximum likelihood estimates for a multivariate normal distribution when some observations are missing. *Journal of the American Statistical Association* 52 (278): 200–3. doi:10.1080/01621459.1957.10501379.

Bhargava, R. P. 1962. Multivariate tests of hypotheses with incomplete data. PhD diss., Stanford University.

Hao, J., and K. Krishnamoorthy. 2001. Inferences on normal covariance matrix and generalized variance with incomplete data. Journal of Multivariate Analysis 78 (1):62–82. doi:10.1006/jmva.2000.1939.

Kanda, T., and Y. Fujikoshi. 1998. Some basic properties of the MLEs for a multivariate normal distribution with monotone missing data. *American Journal of Mathematical and Management Sciences* 18:161–90. doi:10.1080/01966324.1998.10737458.

Krishnamoorthy, K., and M. Pannala. 1998. Some simple test procedures for normal mean vector with incomplete data. *Annals of the Institute of Statistical Mathematics* 50 (3):531–42. doi:10.1023/A:1003581513299.

Krishnamoorthy, K., and M. Pannala. 1999. Condence estimation of normal mean vector with incomplete data. *Canadian Journal of Statistics* 27 (2):395–407. doi:10.2307/3315648.

Krishnamoorthy, K., and J. Yu. 2012. Multivariate Behrens-Fisher problem with missing data. *Journal of Multivariate Analysis* 105 (1):141–50. doi:10.1016/j.jmva.2011.08.019.

Little, R. J. A. 1988. A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association* 83 (404):1198–202. doi:10.1080/01621459.1988.10478722.

Little, R. J. A., and D. B. Rubin. 1987. *Statistical analysis with missing data*. New York, NY: Wiley.

Lu, G. B., and J. B. Copas. 2004. Missing at random, likelihood ignorability and model completeness. The Annals of Statistics 32:754–65. doi:10.1214/009053604000000166.

Morrison, D. F., and D. Bhoj. 1973. Power of the likelihood ratio test on the mean vector of the multivariate normal distribution with missing observations. *Biometrika* 60 (2):365–8. doi:10.1093/biomet/60.2.365.

Naik, U. D. 1975. On testing equality of means of correlated variables with incomplete data. *Biometrika* 62 (3):615–22. doi:10.1093/biomet/62.3.615.

Seber, G. A. F. 1984. *Multivariate observations*. New York, NY: Wiley.

Yu, J., K. Krishnamoorthy, and M. Pannala. 2006. Two-sample inference for normal mean vectors based on monotone missing data. *Journal of Multivariate Analysis* 97 (10):2162–76. doi:10.1016/j.jmva.2006.07.002.