Multivariate normality test based on kurtosis with two-step monotone missing data

Eri Kurita* and Takashi Seo*

*Department of Applied Mathematics, Tokyo University of Science
1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan

Abstract

This paper deals with a sample measure of multivariate kurtosis, which is used as a test statistic in multivariate normality testing problems. We define a new multivariate sample kurtosis measure to provide a multivariate normality test for data with a two-step monotone missing structure. Furthermore, we derive its expectation and variance using a perturbation method. To evaluate the accuracy of a normal approximation, we conducted a Monte Carlo simulation for certain parameters. Finally, we present a numerical example to illustrate the proposed procedure.

Key Words and Phrases. Asymptotic expansion, Moment, Monte Carlo simulation, Multivariate kurtosis, Normal approximation.

1 Introduction

In a multivariate analysis, the assessment of whether multidimensional data follow a multivariate normal distribution is an important task, and many test procedures have been proposed and discussed for multivariate normality (MVN) tests (for related studies, see Farrel et al. [3], Hanusz et al. [5], Henze and Zirkler [8], Kollo [12], Rao et al. [15], Thode [18], and Zhou and Shao [22]). Among the test statistics used in MVN tests are those based on multivariate skewness or multivariate kurtosis, which are defined by Koziol [13], Mardia [14], and Srivastava [17], and their distributions are given for a large sample. In addition, Henze [7] discussed the asymptotic distribution of Mardia's kurtosis test statistic under nonnormality. An MVN test using a normalizing transformation for Mardia's multivariate kurtosis was recently given by Enomoto et al. [2]. Moreover, Yamada et al. [21] discussed a test for an MVN with two-step monotone missing data. By considering a case in which the data contain missing values, this study extends the sample measure of multivariate kurtosis defined by Mardia [14]. In this paper, we define a sample measure of multivariate kurtosis when the data have a two-step monotone pattern of missing observations. We also consider an MVN test under the assumption of a two-step monotone missing data. In particular, we focus on an MVN test statistic using multivariate kurtosis. For two-step monotone missing data, the maximum likelihood estimators (MLEs) of the mean vector and covariance matrix are given (see Kanda and Fujikoshi [9]). Tests of the mean vectors or covariance matrix using these MLEs were discussed by Hao and Krishnamoorthy [6], Tsukada [19], and Yagi et al. [20]. The sample measure of multivariate kurtosis discussed in this paper is also based on MLEs developed by Kanda and Fujikoshi [9]. By decomposing the multivariate kurtosis, the sample analogue of multivariate kurtosis with two-step monotone missing data can be defined, and asymptotic results of the expectation and variance are given using a perturbation method. For references partially related to the perturbation method described in this paper, see Kawasaki and Seo [10] and Kawasaki et al. [11]. The rest of this paper is organized as follows. Section 2 introduces two-step monotone missing data and provides a definition of the sample measure of multivariate kurtosis using such data. In Section 3, we derive the expectation and variance of the sample measure of multivariate kurtosis defined in Section 2, where the result is partly an approximation through an asymptotic expansion. In Section 4, some simulation results for two-step monotone missing data are presented to investigate the accuracy of the normal approximation of the test statistics proposed in this paper. We also provide a numerical example to illustrate the method in Section 4. Finally, we give some concluding remarks in Section 5.

2 Multivariate kurtosis with two-step monotone missing data

Let $\boldsymbol{x}$ be a p-dimensional random vector with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The population measure of multivariate kurtosis is defined as

$$\beta_{2,p} = \mathrm{E}[\{(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\}^2].$$

Under multivariate normality, we note that $\beta_{2,p} = p(p+2)$. Before defining the sample measure of multivariate kurtosis in the case of two-step monotone missing data, we provide a definition of the multivariate sample kurtosis under complete data for comparison and correspondence. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ be N random sample vectors from a p-variate population with mean vector $\boldsymbol{\mu}$ and covariance matrix

$\boldsymbol{\Sigma}$. Then, as a sample analogue of multivariate kurtosis, we define the following:

$$b_{2,p} = \frac{1}{N} \sum_{i=1}^{N} \{(\boldsymbol{x}_i - \overline{\boldsymbol{x}})^\top \boldsymbol{S}^{-1} (\boldsymbol{x}_i - \overline{\boldsymbol{x}})\}^2, \tag{1}$$

where $\overline{\boldsymbol{x}} = N^{-1} \sum_{i=1}^{N} \boldsymbol{x}_i$ and $\boldsymbol{S} = N^{-1} \sum_{i=1}^{N} (\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})^\top$. The definition in (1) is from Mardia [14], and has been applied as a test statistic for an MVN test by deriving the expectation and variance of $b_{2,p}$ under multivariate normality.

As a type of foreshadowing, in the case of the population multivariate kurtosis, let $Z = (\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})$, and thus $\beta_{2,p} = \mathrm{E}[Z^2]$; in addition, we can write $Z = Z_1 + Z_{2\cdot 1}$, where

$$Z_1 = (\boldsymbol{x}_1 - \boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}_{11}^{-1} (\boldsymbol{x}_1 - \boldsymbol{\mu}_1), \quad Z_{2\cdot 1} = (\boldsymbol{x}_{2\cdot 1} - \boldsymbol{\mu}_{2\cdot 1})^\top \boldsymbol{\Sigma}_{22\cdot 1}^{-1} (\boldsymbol{x}_{2\cdot 1} - \boldsymbol{\mu}_{2\cdot 1}),$$

and

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

$$\boldsymbol{x}_{2\cdot 1} = \boldsymbol{x}_2 - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{x}_1, \quad \boldsymbol{\mu}_{2\cdot 1} = \boldsymbol{\mu}_2 - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\mu}_1, \quad \boldsymbol{\Sigma}_{22\cdot 1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}.$$

Under multivariate normality, we note that $\mathrm{E}[Z_1^2] = p_1(p_1 + 2)$, and $\mathrm{E}[Z_{2\cdot 1}^2] = p_2(p_2 + 2)$. Because Z_1 and $Z_{2\cdot 1}$ are independent, it also indicates that $\beta_{2,p} = p(p + 2)$. Similar to this decomposition, consider the sample analogue of the multivariate kurtosis under two-step monotone missing data.

Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N_1}$ be N_1 p-variate random sample vectors and $\boldsymbol{x}_{1,N_1+1}, \ldots, \boldsymbol{x}_{1,N}$ be N_2 p_1-variate random sample vectors. Two-step monotone missing data can then be written as follows:

$$\begin{pmatrix} \boldsymbol{x}_{1,1}^\top & \boldsymbol{x}_{2,1}^\top \\ \vdots & \vdots \\ \boldsymbol{x}_{1,N_1}^\top & \boldsymbol{x}_{2,N_1}^\top \\ \boldsymbol{x}_{1,N_1+1}^\top & * \\ \vdots & \vdots \\ \boldsymbol{x}_{1,N}^\top & * \end{pmatrix},$$

where $\boldsymbol{x}_i = (\boldsymbol{x}_{1,i}^\top, \boldsymbol{x}_{2,i}^\top)^\top$, $i \in \{1, \ldots, N_1\}$, and "$*$" indicates a missing p_2-dimensional vector (Kanda and Fujikoshi [9]; Kawasaki and Seo [10]). Furthermore, we assume a multivariate normal distribution for two-step monotone missing data, i.e.,

$$\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N_1} \overset{i.i.d.}{\sim} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad \boldsymbol{x}_{1,N_1+1}, \ldots, \boldsymbol{x}_{1,N} \overset{i.i.d.}{\sim} N_{p_1}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}).$$

The sample measure of multivariate kurtosis can be defined as follows:

$$b_{2,p_1,p_2} = R_1 + R_2 + R_3, \tag{2}$$

where

$$R_1 = \frac{1}{N}\sum_{i=1}^{N} U_{1,i}^2,\ R_2 = \frac{1}{N_1}\sum_{i=1}^{N_1} U_{2\cdot1,i}^2,\ R_3 = \frac{2}{N_1}\sum_{i=1}^{N_1} U_{1,i}U_{2\cdot1,i},$$

and

$$U_{1,i} = (\boldsymbol{x}_{1,i} - \widehat{\boldsymbol{\mu}}_1)^\top \widehat{\boldsymbol{\Sigma}}_{11}^{-1}(\boldsymbol{x}_{1,i} - \widehat{\boldsymbol{\mu}}_1);\ U_{2\cdot1,i} = (\boldsymbol{x}_{2\cdot1,i} - \widehat{\boldsymbol{\mu}}_{2\cdot1})^\top \widehat{\boldsymbol{\Sigma}}_{22\cdot1}^{-1}(\boldsymbol{x}_{2\cdot1,i} - \widehat{\boldsymbol{\mu}}_{2\cdot1});$$

$\boldsymbol{x}_{2\cdot1,i} = \boldsymbol{x}_{2,i} - \widehat{\boldsymbol{\Sigma}}_{21}\widehat{\boldsymbol{\Sigma}}_{11}^{-1}\boldsymbol{x}_{1,i}$; $\widehat{\boldsymbol{\mu}}_1,\ \widehat{\boldsymbol{\Sigma}}_{11},\ \widehat{\boldsymbol{\mu}}_{2\cdot1}$, and $\widehat{\boldsymbol{\Sigma}}_{22\cdot1}$ are MLEs of $\boldsymbol{\mu}_1,\ \boldsymbol{\Sigma}_{11},\ \boldsymbol{\mu}_{2\cdot1} = \boldsymbol{\mu}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\mu}_1$, and $\boldsymbol{\Sigma}_{22\cdot1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$, respectively. The MLEs of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ were derived by Anderson and Olkin [1] for two-step monotone missing data. For k-step monotone missing data, including the case of two-step monotone missing data, see Jinadasa and Tracy [4] and Kanda and Fujikoshi [9]. Kanda and Fujikoshi [9] also provides the properties of their distributions. In addition, we use the following definition as a notation.

$$\overline{\boldsymbol{x}}_{(1)} = \frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{x}_i = \begin{pmatrix} \overline{\boldsymbol{x}}_{(1),1} \\ \overline{\boldsymbol{x}}_{(1),2} \end{pmatrix},\quad \boldsymbol{S}_{(1)} = \frac{1}{N_1}\sum_{i=1}^{N_1}(\boldsymbol{x}_i - \overline{\boldsymbol{x}}_{(1)})(\boldsymbol{x}_i - \overline{\boldsymbol{x}}_{(1)})^\top = \begin{pmatrix} \boldsymbol{S}_{(1),11} & \boldsymbol{S}_{(1),12} \\ \boldsymbol{S}_{(1),21} & \boldsymbol{S}_{(1),22} \end{pmatrix},$$

$$\overline{\boldsymbol{x}}_{(2)} = \frac{1}{N_2}\sum_{i=N_1+1}^{N}\boldsymbol{x}_{1,i},\ \boldsymbol{S}_{(2)} = \frac{1}{N_2}\sum_{i=N_1+1}^{N}(\boldsymbol{x}_{1,i} - \overline{\boldsymbol{x}}_{(2)})(\boldsymbol{x}_{1,i} - \overline{\boldsymbol{x}}_{(2)})^\top,$$

$$\overline{\boldsymbol{x}}_T = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_{1,i},\ \boldsymbol{S}_T = \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{x}_{1,i} - \overline{\boldsymbol{x}}_T)(\boldsymbol{x}_{1,i} - \overline{\boldsymbol{x}}_T)^\top.$$

We note that

$$\overline{\boldsymbol{x}}_T = r\overline{\boldsymbol{x}}_{(1),1} + (1-r)\overline{\boldsymbol{x}}_{(2)},\quad \boldsymbol{S}_T = r\boldsymbol{S}_{(1),11} + (1-r)\boldsymbol{S}_{(2)} + r(1-r)(\overline{\boldsymbol{x}}_{(1),1} - \overline{\boldsymbol{x}}_{(2)})(\overline{\boldsymbol{x}}_{(1),1} - \overline{\boldsymbol{x}}_{(2)})^\top,$$

where r is a constant such that $N_1 = rN,\ 0 < r < 1$. By substituting these MLEs into $U_{1,i}$ and $U_{2\cdot1,i}$, we can write

$$U_{1,i} = (\boldsymbol{x}_{1,i} - \overline{\boldsymbol{x}}_T)^\top \boldsymbol{S}_T^{-1}(\boldsymbol{x}_{1,i} - \overline{\boldsymbol{x}}_T), \tag{3}$$

$$U_{2\cdot1,i} = (\boldsymbol{x}_{2\cdot1,i} - \widehat{\boldsymbol{\mu}}_{2\cdot1})^\top \widehat{\boldsymbol{\Sigma}}_{22\cdot1}^{-1}(\boldsymbol{x}_{2\cdot1,i} - \widehat{\boldsymbol{\mu}}_{2\cdot1}), \tag{4}$$

respectively, where $\boldsymbol{x}_{2\cdot 1,i} = \boldsymbol{x}_{2,i} - \boldsymbol{S}_{(1),21}\boldsymbol{S}_{(1),11}^{-1}\boldsymbol{x}_{1,i}$, $\widehat{\boldsymbol{\mu}}_{2\cdot 1} = \overline{\boldsymbol{x}}_{(1),2} - \boldsymbol{S}_{(1),21}\boldsymbol{S}_{(1),11}^{-1}\overline{\boldsymbol{x}}_{(1),1}$, and $\widehat{\boldsymbol{\Sigma}}_{22\cdot 1} = \boldsymbol{S}_{(1),22} - \boldsymbol{S}_{(1),21}\boldsymbol{S}_{(1),11}^{-1}\boldsymbol{S}_{(1),12}$. In the next section, we consider the expectation and variance of b_{2,p_1,p_2} in (2).

3 Moments of multivariate sample kurtosis

In this section, we consider the first and second moments of b_{2,p_1,p_2} under multivariate normality. Before that, the expectation and variance of $b_{2,p}$ in (1), i.e., the multivariate sample kurtosis under complete data, is given by Mardia [14] as follows:

$$\mathrm{E}[b_{2,p}] = p(p+2)\frac{N-1}{N+1}, \tag{5}$$

$$\mathrm{Var}[b_{2,p}] = 8p(p+2)\frac{(N-3)(N-p-1)(N-p+1)}{(N+1)^2(N+3)(N+5)}. \tag{6}$$

For further details, see Siotani et al. [16]. Using this result in (5), we obtain

$$\mathrm{E}[R_1] = p_1(p_1+2)\frac{N-1}{N+1}. \tag{7}$$

For $\mathrm{E}[R_i]$, $i \in \{2,3\}$, although it is difficult to provide an exact expectation of R_i, we give an asymptotic expansion of $\mathrm{E}[R_i]$ using the perturbation expansion method, where N_1 and $N \to \infty$ with r $(= N_1/N)$ $\to \delta \in (0,1]$. As a result, we obtain

$$\mathrm{E}[R_2] = p_2(p_2+2) - \frac{2}{rN}p_2(p_2+2) + \mathrm{O}(N^{-\frac{3}{2}}), \quad \mathrm{E}[R_3] = 2p_1p_2 - \frac{4}{rN}p_1p_2 + \mathrm{O}(N^{-\frac{3}{2}}). \tag{8}$$

See Appendix A for details of the derivation. Therefore, the expectation of b_{2,p_1,p_2} is given by

$$\mathrm{E}[b_{2,p_1,p_2}] = p(p+2) + \frac{c}{N} + \mathrm{O}(N^{-\frac{3}{2}}), \quad c = -2\left\{\frac{1}{r}p(p+2) + \left(1-\frac{1}{r}\right)p_1(p_1+2)\right\}.$$

From this result, we propose an approximate asymptotic expansion given by

$$m_1 = p(p+2) + \frac{c}{N}. \tag{9}$$

Furthermore, because the exact expectations of R_1 is given by (7), as a nearly equal approximation of $\mathrm{E}[b_{2,p_1,p_2}]$, we can also propose the following:

$$m_2 = p(p+2) - \frac{2}{N+1}p_1(p_1+2) - \frac{2}{rN}p_2(2p_1+p_2+2). \tag{10}$$

Next, we consider the variance of b_{2,p_1,p_2}, which is given by

$$\mathrm{Var}[b_{2,p_1,p_2}] = \sum_{i=1}^{3} \mathrm{Var}[R_i] + 2\sum_{\substack{i=1 \\ i\neq j}}^{3}\sum_{j=1}^{3} \mathrm{Cov}[R_i, R_j].$$

From (6), which is the result of Mardia [14], we obtain

$$\mathrm{Var}[R_1] = 8p_1(p_1+2)\frac{(N-3)(N-p_1-1)(N-p_1+1)}{(N+1)^2(N+3)(N+5)}. \tag{11}$$

The variances of R_2 and R_3 also provide asymptotic results using the same method as the derivation of the expectation. To summarize these results, $\mathrm{Var}[R_2]$ and $\mathrm{Var}[R_3]$ are given by the following:

$$\mathrm{Var}[R_2] = \frac{1}{N}\sigma_2^2 + \mathrm{O}(N^{-\frac{3}{2}}), \quad \mathrm{Var}[R_3] = \frac{1}{N}\sigma_3^2 + \mathrm{O}(N^{-\frac{3}{2}}), \tag{12}$$

respectively, where

$$\sigma_2^2 = \frac{8}{r}p_2(p_2+2), \quad \sigma_3^2 = 8p_1p_2\left\{\left(\frac{1}{r}-1\right)p_2 + \frac{2}{r}\right\}.$$

See Appendix B for details of the derivation. Noting that the covariances between R_i and R_j, where $1 \leq i < j \leq 3$, are $\mathrm{O}(N^{-3/2})$, the variance of b_{2,p_1,p_2} is given by the following:

$$\mathrm{Var}(b_{2,p_1,p_2}) = \frac{1}{N}\sigma^2 + \mathrm{O}(N^{-\frac{3}{2}}), \quad \sigma^2 = 8\left[p_1(p_1+2) + \frac{1}{r}p_2(p_2+2) + p_1p_2\left\{\left(\frac{1}{r}-1\right)p_2 + \frac{2}{r}\right\}\right]. \tag{13}$$

Because the exact variance is given as in (11) for R_1, a nearly equal approximation of $\mathrm{Var}[b_{2,p_1,p_2}]$ can be proposed:

$$\nu^2 = \nu_1^2 + \frac{8}{rN}p_2\{(1-r)p_1p_2 + 2p_1 + p_2 + 2\}, \tag{14}$$

where

$$\nu_1^2 = 8p_1(p_1+2)\frac{(N-3)(N-p_1-1)(N-p_1+1)}{(N+1)^2(N+3)(N+5)}.$$

Thus, we propose the following three test statistics:

$$Z_{MM} = \frac{b_{2,p_1,p_2} - p(p+2)}{\sqrt{\dfrac{\sigma^2}{N}}}, \tag{15}$$

$$Z^*_{MM} = \frac{b_{2,p_1,p_2} - \left\{p(p+2) + \dfrac{c}{N}\right\}}{\sqrt{\dfrac{\sigma^2}{N}}}, \tag{16}$$

$$Z^{**}_{MM} = \frac{b_{2,p_1,p_2} - \left\{p(p+2) - \dfrac{2}{N+1}p_1(p_1+2) - \dfrac{2}{rN}p_2(2p_1+p_2+2)\right\}}{\sqrt{\nu_1^2 + \dfrac{8}{rN}p_2\{(1-r)p_1p_2 + 2p_1 + p_2 + 2\}}}. \tag{17}$$

Note that these test statistics are asymptotically distributed as $N(0,1)$, where N_1 and $N \to \infty$ with $r \to \delta \in (0,1]$.

4 Simulation studies and an example

In this section, the normal approximation for the three test statistics Z_{MM}, Z^*_{MM}, and Z^{**}_{MM} in (15), (16), and (17) is assessed based on a Monte Carlo simulation. Figure 1 presents histograms for Z_{MM}, Z^*_{MM}, and Z^{**}_{MM} from the results of 100,000 simulations. The random multidimensional data follow a multivariate standard normal distribution, and the histograms shown in Figure 1 are for the following parameter sets: $(p_1, p_2) = (2,2),\ (5,5)$, and $(N_1, N_2) = (40,10),\ (100,10),\ (400,10)$. To show whether the histograms have a standard normal distribution, a standard normal density curve is also included in Figure 1. It can be seen from the figure that the shape of the histogram approaches that of the density function of a standard normal distribution as sample size N_1 increases. In particular, the histogram of Z^{**}_{MM} is fit to a standard normal density curve even when the sample size is moderately small.

Table 1 lists the simulation values and theoretical results for the expectation and variance of b_{2,p_1,p_2}, respectively. For the simulations, 100,000 experiments were conducted for certain combinations of parameters. The parameter settings for the simulation are $(p_1, p_2) = (2,2), (3,3), (5,5)$ for the dimensions and $(N_1, N_2) = (m, n),\ m \in \{20, 30, 40, 100, 200, 400\},\ n \in \{10, 20\}$ for the sample sizes. For the theoretical results, the expectation is given based on the value of the expansion at up to N^{-1},

m_1, and the approximate value, m_2. It can be seen from Table 1 that both the empirical value and theoretical results m_1 and m_2 converge to $p(p+2)$, and in particular, the approximation m_2 is highly accurate for all cases. Regarding the variance, Table 1 shows that the empirical value of the variance multiplied by N converges to $8p(p+2)$ as N_1 increases. It can also be seen that the variance through the approximation ν^2 is an extremely good approximation for the majority of cases.

Next, as shown in Table 2, the empirical expectation, variance, skewness, and kurtosis of the three test statistics, Z_{MM}, Z^*_{MM}, and Z^{**}_{MM} are given. The settings for the simulation parameters are the same as those in Table 1. From Table 2, we can see that as the sample size increases, the expectation, variance, skewness, and kurtosis of any test statistics approach the corresponding values of the standard normal distribution of 0, 1, 0, and 3, respectively. In particular, it can be seen that the empirical expectation and variance of Z^{**}_{MM} converge to 0 and 1 more quickly than those of Z_{MM} and Z^*_{MM}. Note that, because these three test statistics differ only in terms of location and scale, their skewness and kurtosis are the same.

In Table 3, we give a type I error for the three test statistics. The parameter settings are the same as before, where $\alpha = 0.05$. From the results in Table 3, we can see that the type I errors of test statistics are closer to 0.05 as N_1 increases. In particular, it can be seen that the type I error of Z^{**}_{MM} is closer to 0.05 than the those of Z^*_{MM} and Z_{MM} in most cases. It can be seen that this result follows the same trend as Mardia's Z_M and Z^*_M for the complete data in Enomoto et al.'s study [2]. Note that, because Z_{MM} and Z^*_{MM} statistics differ only in terms of location, their type I errors are the same.

Finally, in this paper, we consider Fisher's Iris data to illustrate the proposed method. The Iris data handled here presents the measurements of the sepal length and width, and the pedal length and width, in centimeters of 50 Iris virginica plants. To make the dataset with two-step monotone missing data, we artificially selected 10 plants at random from the 50 plants, and from the data of these 10 plants, we excluded the data on the pedal length and width, leaving only the data on the sepal length and width. That is, we created a dataset in which 40 of the plants consisted of 4 variables, and 10 plants consisted of 2 variables. The dataset has $(p_1, p_2) = (2, 2)$ and $(N_1, N_2) = (40, 10)$. From these data, the values of b_{2,p_1,p_2} are calculated as $b_{2,p_1,p_2} = 24.09$, and the values of the test statistics are $Z_{MM} = 0.043$, $Z^*_{MM} = 0.554$, and $Z^{**}_{MM} = 0.574$. Thus, it can be seen that the multivariate normality is not rejected at the 5% level of significance. Incidentally, in the case of complete data $(p = 4, N = 50)$ for Iris virginica, Mardia's multivariate sample kurtosis, $b_{2,p} = 24.30$, and the value of the test statistic Z^*_M is 0.782, where the Z^*_M statistic is defined for the complete data. Therefore, the case of complete data is also not rejected.

5 Conclusion

In this paper, we defined a new sample measure of multivariate kurtosis when the type of data has a two-step monotone pattern of missing observations. This approach is based on Mardia's multivariate kurtosis, and we considered its sample version by decomposing the multivariate kurtosis. We then developed test statistics for an MVN test by asymptotically evaluating the expectation and variance using an asymptotic expansion procedure. In particular, in some parts of the decomposition of multivariate kurtosis, we also provide the exact expectations and variances in their sample version. Hence, it was possible to give a test statistic with a better approximation even when the sample size is moderately small. Future studies will involve extending the method to cases with more than a three-step monotone pattern and deriving a normalizing transformation statistic for the test statistic given in this paper.

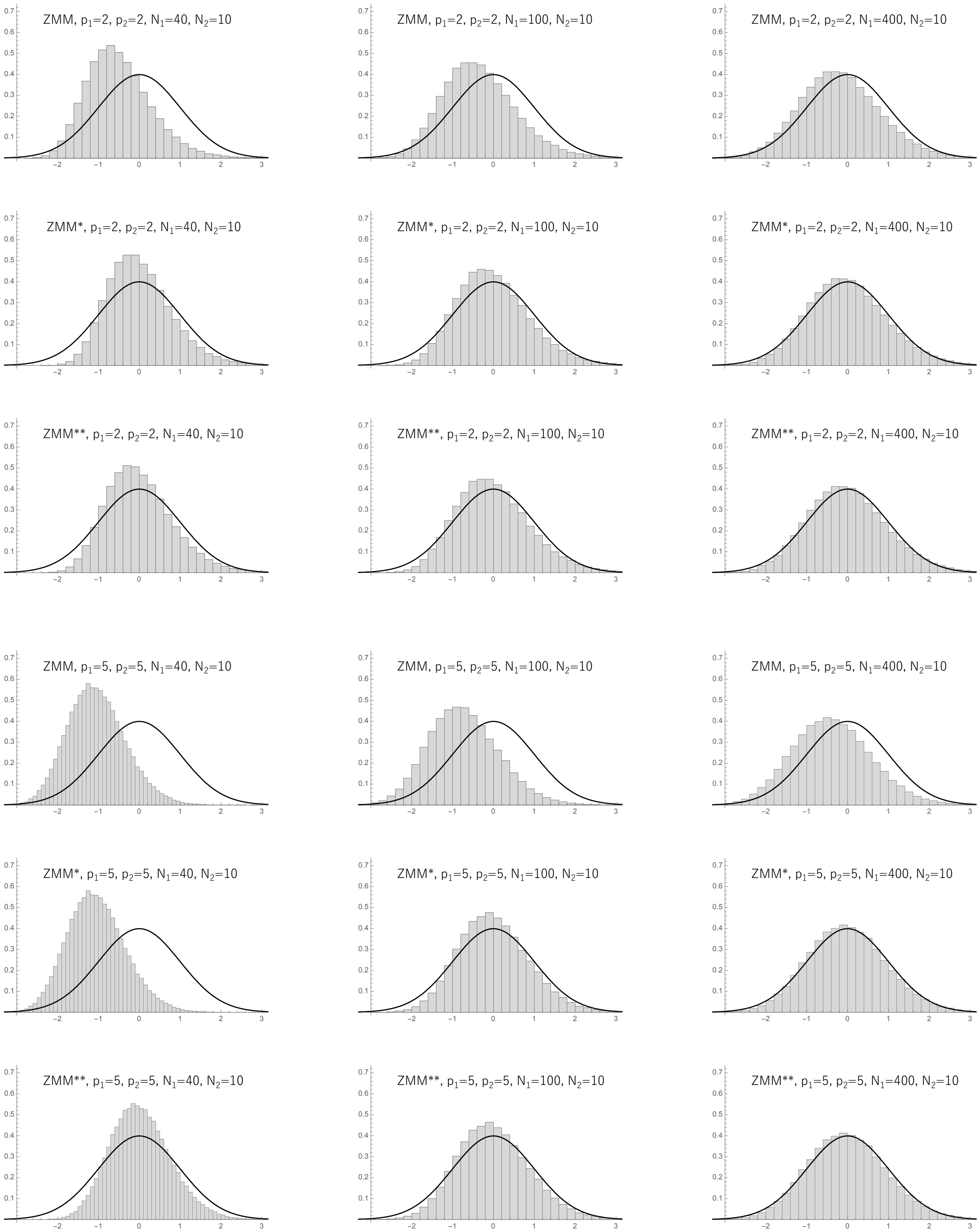

Figure 1: Histogram of the three test statistics Z_{MM}, Z^{*}_{MM}, and Z^{**}_{MM} in (15), (16), and (17) for dimensions $(p_1, p_2) = (2, 2)$, $(5, 5)$ and sample sizes $(N_1, N_2) = (40, 10)$, $(100, 10)$, and $(400, 10)$.

Table 1: Expectations and variances of b_{2,p_1,p_2} for $(p_1, p_2) = (2, 2), (3, 3)$, and $(5, 5)$.

				Simulation		Approximation			
p_1	p_2	N_1	N_2	$\mathrm{E}[b_{2,p_1,p_2}]$	$N\mathrm{Var}[b_{2,p_1,p_2}]$	m_1	m_2	σ^2	$N\nu^2$
2	2	20	10	21.97	124.99	21.87	21.88	288.00	260.57
		30		22.58	140.31	22.53	22.54	256.00	234.02
		40		22.91	149.50	22.88	22.89	240.00	221.68
		100		23.54	172.61	23.53	23.54	211.20	202.07
		200		23.76	180.12	23.76	23.76	201.60	196.63
		400		23.88	187.08	23.88	23.88	196.80	194.20
2	2	20	20	22.11	182.90	22.00	22.01	384.00	362.02
		30		22.67	188.44	22.61	22.62	320.00	301.68
		40		22.95	187.83	22.93	22.94	288.00	272.31
		100		23.55	190.68	23.55	23.55	230.40	221.98
		200		23.77	191.34	23.77	23.77	211.20	206.45
		400		23.88	192.64	23.88	23.88	201.60	199.06
3	3	20	10	43.92	242.12	43.70	43.73	624.00	567.75
		30		45.16	273.74	45.05	45.07	544.00	498.69
		40		45.81	294.54	45.75	45.76	504.00	466.12
		100		47.07	339.47	47.07	47.07	432.00	412.99
		200		47.53	360.61	47.53	47.53	408.00	397.63
		400		47.76	374.28	47.76	47.76	396.00	390.57
3	3	20	20	44.18	385.05	43.95	43.97	864.00	818.69
		30		45.30	386.79	45.20	45.21	704.00	666.12
		40		45.90	381.78	45.85	45.86	624.00	591.48
		100		47.10	382.23	47.09	47.09	480.00	462.45
		200		47.53	384.59	47.53	47.53	432.00	422.08
		400		47.77	383.67	47.76	47.76	408.00	402.70
5	5	20	10	109.72	576.25	109.17	109.24	1800.00	1647.50
		30		112.84	648.28	112.58	112.63	1520.00	1395.94
		40		114.50	692.16	114.35	114.38	1380.00	1275.65
		100		117.68	833.96	117.66	117.67	1128.00	1074.91
		200		118.82	890.34	118.82	118.82	1044.00	1014.88
		400		119.40	926.47	119.40	119.41	1002.00	986.71
5	5	20	20	110.32	1063.99	109.75	109.79	2640.00	2515.94
		30		113.19	1028.54	112.93	112.96	2080.00	1975.65
		40		114.72	1009.71	114.58	114.60	1800.00	1710.04
		100		117.75	984.64	117.72	117.72	1296.00	1246.94
		200		118.82	979.23	118.83	118.83	1128.00	1100.14
		400		119.40	961.76	119.41	119.41	1044.00	1029.06

Note. m_1 and m_2 are approximate values of $\mathrm{E}[b_{2,p_1,p_2}]$ given in (9) and (10), respectively. For the variance, σ^2 is an asymptotic variance of $N\mathrm{Var}[b_{2,p_1,p_2}]$ in (13) , and $N\nu^2$ is an approximate variance in (14) multiplied by N.

Table 2: Expectations, variances, skewness, and kurtosis for the Z_{MM}, Z^{*}_{MM}, and Z^{**}_{MM} test statistics given in (15), (16), and (17).

				Expectation			Variance			
p_1	p_2	N_1	N_2	Z_{MM}	Z^{*}_{MM}	Z^{**}_{MM}	Z_{MM}, Z^{*}_{MM}	Z^{**}_{MM}	Skewness	Kurtosis
2	2	20	10	-0.655	0.034	0.030	0.434	0.480	0.674	3.821
		30		-0.560	0.019	0.016	0.548	0.600	0.684	3.904
		40		-0.500	0.011	0.009	0.623	0.674	0.644	3.801
		100		-0.334	0.002	0.001	0.817	0.854	0.576	3.753
		200		-0.247	-0.006	-0.006	0.893	0.916	0.434	3.395
		400		-0.169	0.003	0.003	0.951	0.963	0.329	3.252
		20	20	-0.611	0.034	0.032	0.476	0.505	0.666	3.859
		30		-0.527	0.021	0.019	0.589	0.625	0.671	3.933
		40		-0.477	0.010	0.008	0.652	0.690	0.658	3.916
		100		-0.322	0.005	0.005	0.828	0.859	0.551	3.624
		200		-0.238	-0.001	-0.001	0.906	0.927	0.437	3.429
		400		-0.174	-0.003	-0.004	0.956	0.968	0.341	3.249
3	3	20	10	-0.894	0.049	0.044	0.388	0.426	0.476	3.377
		30		-0.771	0.029	0.025	0.503	0.549	0.519	3.478
		40		-0.689	0.019	0.016	0.584	0.632	0.496	3.437
		100		-0.469	0.001	0.000	0.786	0.822	0.445	3.434
		200		-0.339	0.000	0.000	0.884	0.907	0.360	3.273
		400		-0.242	0.000	0.000	0.945	0.958	0.268	3.169
3	3	20	20	-0.823	0.049	0.046	0.446	0.470	0.460	3.328
		30		-0.719	0.027	0.024	0.549	0.581	0.481	3.448
		40		-0.653	0.014	0.012	0.612	0.645	0.473	3.423
		100		-0.451	0.004	0.003	0.796	0.827	0.431	3.397
		200		-0.332	0.001	0.000	0.890	0.911	0.344	3.237
		400		-0.235	0.005	0.005	0.940	0.953	0.250	3.131
5	5	20	10	-1.327	0.072	0.065	0.320	0.350	0.284	3.099
		30		-1.162	0.041	0.036	0.427	0.464	0.328	3.167
		40		-1.047	0.029	0.024	0.502	0.543	0.335	3.167
		100		-0.725	0.004	0.002	0.739	0.776	0.345	3.242
		200		-0.528	0.002	0.002	0.853	0.877	0.278	3.174
		400		-0.385	-0.004	-0.004	0.927	0.942	0.219	3.124
		20	20	-1.191	0.070	0.067	0.403	0.423	0.286	3.148
		30		-1.055	0.040	0.037	0.494	0.521	0.297	3.161
		40		-0.965	0.024	0.021	0.561	0.590	0.321	3.187
		100		-0.685	0.010	0.009	0.760	0.790	0.322	3.182
		200		-0.520	-0.004	-0.005	0.868	0.890	0.274	3.169
		400		-0.378	-0.003	-0.003	0.921	0.935	0.197	3.084

Table 3: Empirical type I error of the Z_{MM}, Z^{*}_{MM}, and Z^{**}_{MM} test statistics given in (15), (16), and (17) for $(p_1, p_2) = (2,2), (3,3), (5,5)$, and $\alpha = 0.05$.

				Empirical type I error		
p_1	p_2	N_1	N_2	Z_{MM}	Z^{*}_{MM}	Z^{**}_{MM}
2	2	20	10	0.0075	0.0088	0.0110
		30		0.0132	0.0141	0.0169
		40		0.0176	0.0174	0.0206
		100		0.0328	0.0307	0.0337
		200		0.0394	0.0375	0.0397
		400		0.0459	0.0445	0.0459
2	2	20	20	0.0095	0.0103	0.0116
		30		0.0162	0.0160	0.0183
		40		0.0206	0.0197	0.0219
		100		0.0331	0.0322	0.0350
		200		0.0409	0.0386	0.0406
		400		0.0461	0.0448	0.0461
3	3	20	10	0.0274	0.0048	0.0060
		30		0.0297	0.0096	0.0121
		40		0.0332	0.0140	0.0170
		100		0.0408	0.0279	0.0309
		200		0.0448	0.0369	0.0392
		400		0.0469	0.0432	0.0445
3	3	20	20	0.0278	0.0068	0.0078
		30		0.0322	0.0122	0.0135
		40		0.0337	0.0152	0.0176
		100		0.0394	0.0288	0.0315
		200		0.0449	0.0371	0.0390
		400		0.0471	0.0432	0.0446
5	5	20	10	0.1274	0.0015	0.0019
		30		0.1039	0.0046	0.0062
		40		0.0909	0.0073	0.0093
		100		0.0667	0.0237	0.0269
		200		0.0587	0.0333	0.0357
		400		0.0555	0.0408	0.0423
5	5	20	20	0.1086	0.0038	0.0045
		30		0.0922	0.0072	0.0085
		40		0.0833	0.0104	0.0121
		100		0.0651	0.0247	0.0275
		200		0.0598	0.0352	0.0376
		400		0.0539	0.0406	0.0419

Appendix A. Derivation of $\mathrm{E}[R_2]$ and $\mathrm{E}[R_3]$

We derive asymptotic expansions of $\mathrm{E}[R_2]$ and $\mathrm{E}[R_3]$ using the perturbation method as follows. To avoid the dependency between $\boldsymbol{x}_{1,i}$ and $\overline{\boldsymbol{x}}_T$ and between $\boldsymbol{x}_{1,i}$ and $\boldsymbol{S}_T$ in (3), we use

$$\overline{\boldsymbol{x}}_T^{(i)} = \frac{1}{N-1}\sum_{\substack{\alpha=1\\ \alpha\neq i}}^{N}\boldsymbol{x}_{1,\alpha}, \quad \boldsymbol{S}_T^{(i)} = \frac{1}{N-1}\sum_{\substack{\alpha=1\\ \alpha\neq i}}^{N}(\boldsymbol{x}_{1,\alpha}-\overline{\boldsymbol{x}}_T^{(i)})(\boldsymbol{x}_{1,\alpha}-\overline{\boldsymbol{x}}_T^{(i)})^\top.$$

For $i \in \{1,\dots,N_1\}$, we can then write

$$\begin{aligned}\boldsymbol{x}_{1,i}-\overline{\boldsymbol{x}}_T &= \left(1-\frac{1}{N}\right)(\boldsymbol{x}_{1,i}-\overline{\boldsymbol{x}}_T^{(i)}),\\ \boldsymbol{S}_T^{-1} &= \left(1+\frac{1}{N}\right)\{\boldsymbol{S}_T^{(i)}\}^{-1} - \frac{1}{N}\{\boldsymbol{S}_T^{(i)}\}^{-1}(\boldsymbol{x}_{1,i}-\overline{\boldsymbol{x}}_T^{(i)})(\boldsymbol{x}_{1,i}-\overline{\boldsymbol{x}}_T^{(i)})^\top\{\boldsymbol{S}_T^{(i)}\}^{-1} + \mathrm{O}_p(N^{-2}).\end{aligned}$$

Then, $\overline{\boldsymbol{x}}_T^{(i)}$ and $\boldsymbol{S}_T^{(i)}$ can be written as

$$\begin{aligned}\overline{\boldsymbol{x}}_T^{(i)} &= \frac{N_1-1}{N-1}\overline{\boldsymbol{x}}_{(1),1}^{(i)} + \frac{N_2}{N-1}\overline{\boldsymbol{x}}_{(2)},\\ \boldsymbol{S}_T^{(i)} &= \frac{N_1-1}{N-1}\boldsymbol{S}_{(1),11}^{(i)} + \frac{N_1-1}{N-1}(\overline{\boldsymbol{x}}_{(1),1}^{(i)}-\overline{\boldsymbol{x}}_T^{(i)})(\overline{\boldsymbol{x}}_{(1),1}^{(i)}-\overline{\boldsymbol{x}}_T^{(i)})^\top\\ &\quad + \frac{N_2}{N-1}\boldsymbol{S}_{(2)} + \frac{N_2}{N-1}(\overline{\boldsymbol{x}}_{(2)}-\overline{\boldsymbol{x}}_T^{(i)})(\overline{\boldsymbol{x}}_{(2)}-\overline{\boldsymbol{x}}_T^{(i)})^\top.\end{aligned}$$

Herein, we use the perturbation method to expand the statistic $U_{1,i}$ in (3). Without a loss of generality, we can assume that $\boldsymbol{\mu}=\boldsymbol{0}$ and $\boldsymbol{\Sigma}=\boldsymbol{I}_p$. Let

$$\overline{\boldsymbol{x}}_{(1)}^{(i)} = \begin{pmatrix}\overline{\boldsymbol{x}}_{(1),1}^{(i)}\\ \overline{\boldsymbol{x}}_{(1),2}^{(i)}\end{pmatrix} = \frac{1}{\sqrt{N_1-1}}\begin{pmatrix}\boldsymbol{z}_1\\ \boldsymbol{z}_2\end{pmatrix},$$

$$\boldsymbol{S}_{(1)}^{(i)} = \begin{pmatrix}\boldsymbol{S}_{(1),11}^{(i)} & \boldsymbol{S}_{(1),12}^{(i)}\\ \boldsymbol{S}_{(1),21}^{(i)} & \boldsymbol{S}_{(1),22}^{(i)}\end{pmatrix} = \left(1-\frac{1}{N_1-1}\right)\left\{\begin{pmatrix}\boldsymbol{I}_{p_1} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}_{p_2}\end{pmatrix} + \frac{1}{\sqrt{N_1-1}}\begin{pmatrix}\boldsymbol{V}_{11} & \boldsymbol{V}_{12}\\ \boldsymbol{V}_{21} & \boldsymbol{V}_{22}\end{pmatrix}\right\},$$

$$\overline{\boldsymbol{x}}_{(2)} = \frac{1}{\sqrt{N_2}}\boldsymbol{z}_3,\ \boldsymbol{S}_{(2)} = \left(1-\frac{1}{N_2}\right)\left(\boldsymbol{I}_{p_1}+\frac{1}{\sqrt{N_2}}\boldsymbol{Z}\right),$$

where

$$\overline{\boldsymbol{x}}_{(1)}^{(i)} = \frac{1}{N_1-1}\sum_{\substack{\alpha=1\\ \alpha\neq i}}^{N_1}\boldsymbol{x}_\alpha, \quad \boldsymbol{S}_{(1)}^{(i)} = \frac{1}{N_1-1}\sum_{\substack{\alpha=1\\ \alpha\neq i}}^{N_1}(\boldsymbol{x}_\alpha-\overline{\boldsymbol{x}}_{(1)}^{(i)})(\boldsymbol{x}_\alpha-\overline{\boldsymbol{x}}_{(1)}^{(i)})^\top.$$

Then, $U_{1,i}$, $i \in \{1, \ldots, N_1\}$, in (3) can be expanded as

$$U_{1,i} = \boldsymbol{x}_{1,i}^\top \boldsymbol{x}_{1,i} - \frac{1}{\sqrt{N}} A_1 + \frac{1}{N} A_2 + \mathrm{O}_p(N^{-\frac{3}{2}}), \tag{18}$$

where

$$\begin{aligned}
A_1 &= \sqrt{r}(2\boldsymbol{x}_{1,i}^\top \boldsymbol{z}_1 + \boldsymbol{x}_{1,i}^\top \boldsymbol{V}_{11} \boldsymbol{x}_{1,i}) + \sqrt{1-r}(2\boldsymbol{x}_{1,i}^\top \boldsymbol{z}_3 + \boldsymbol{x}_{1,i}^\top \boldsymbol{Z} \boldsymbol{x}_{1,i}), \\
A_2 &= \boldsymbol{x}_{1,i}^\top \boldsymbol{x}_{1,i} - (\boldsymbol{x}_{1,i}^\top \boldsymbol{x}_{1,i})^2 \\
&\quad + \sqrt{r(1-r)}(2\boldsymbol{z}_1^\top \boldsymbol{z}_3 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{z}_1 \boldsymbol{z}_3^\top \boldsymbol{x}_{1,i} + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{V}_{11} \boldsymbol{z}_3 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{Z} \boldsymbol{z}_1 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{V}_{11} \boldsymbol{Z} \boldsymbol{x}_{1,i}) \\
&\quad + r\left\{ \boldsymbol{z}_1^\top \boldsymbol{z}_1 - (\boldsymbol{x}_{1,i}^\top \boldsymbol{z}_3)^2 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{V}_{11} \boldsymbol{z}_1 + \boldsymbol{x}_{1,i}^\top \boldsymbol{V}_{11}^2 \boldsymbol{x}_{1,i} \right\} \\
&\quad + (1-r)\left\{ \boldsymbol{z}_3^\top \boldsymbol{z}_3 - (\boldsymbol{x}_{1,i}^\top \boldsymbol{z}_1)^2 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{Z} \boldsymbol{z}_3 + \boldsymbol{x}_{1,i}^\top \boldsymbol{Z}^2 \boldsymbol{x}_{1,i} \right\}.
\end{aligned}$$

Next, we consider a stochastic expansion of $U_{2\cdot 1,i}$ in (4). Expanding in the same way as the perturbation expansion of $U_{1,i}$, we obtain the following:

$$U_{2\cdot 1,i} = \boldsymbol{x}_{2,i}^\top \boldsymbol{x}_{2,i} - \frac{1}{\sqrt{N}} B_1 + \frac{1}{N} B_2 + \mathrm{O}_p(N^{-\frac{3}{2}}), \tag{19}$$

where

$$\begin{aligned}
B_1 &= \frac{1}{\sqrt{r}}(2\boldsymbol{x}_{2,i}^\top \boldsymbol{z}_2 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{V}_{12} \boldsymbol{x}_{2,i} + \boldsymbol{x}_{2,i}^\top \boldsymbol{V}_{22} \boldsymbol{x}_{2,i}), \\
B_2 &= \frac{1}{r}\Big\{ \boldsymbol{z}_2^\top \boldsymbol{z}_2 - 2\boldsymbol{x}_{1,i}^\top \boldsymbol{x}_{1,i} \boldsymbol{x}_{2,i}^\top \boldsymbol{x}_{2,i} - (\boldsymbol{x}_{2,i}^\top \boldsymbol{x}_{2,i})^2 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{V}_{12} \boldsymbol{z}_2 + 2\boldsymbol{x}_{2,i}^\top \boldsymbol{V}_{21} \boldsymbol{z}_1 + 2\boldsymbol{x}_{2,i}^\top \boldsymbol{V}_{22} \boldsymbol{z}_2 \\
&\quad + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{V}_{11} \boldsymbol{V}_{12} \boldsymbol{x}_{2,i} + \boldsymbol{x}_{1,i}^\top \boldsymbol{V}_{12} \boldsymbol{V}_{21} \boldsymbol{x}_{1,i} + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{V}_{12} \boldsymbol{V}_{22} \boldsymbol{x}_{2,i} + \boldsymbol{x}_{2,i}^\top \boldsymbol{V}_{21} \boldsymbol{V}_{12} \boldsymbol{x}_{2,i} + \boldsymbol{x}_{2,i}^\top \boldsymbol{V}_{22}^2 \boldsymbol{x}_{2,i} \Big\}.
\end{aligned}$$

Calculating the expectations of the expansion of $U_{1,i}$ squared in (18) and the product of two expanded results in (18) and (19) with respect to $\boldsymbol{x}_{1,i}$, $\boldsymbol{x}_{2,i}$, $\boldsymbol{z}_j$, $j \in \{1, 2, 3\}$, $\boldsymbol{Z}$, and $\boldsymbol{V}$, we obtain (8).

Appendix B. Derivation of $\mathrm{Var}[R_2]$ and $\mathrm{Var}[R_3]$

First, for $i \neq j$, the second moments of R_2 and R_3 can be written as

$$\mathrm{E}[R_2^2] = \frac{1}{N_1}\mathrm{E}[U_{2\cdot 1,i}^4] + \left(1 - \frac{1}{N_1}\right)\mathrm{E}[U_{2\cdot 1,i}^2 U_{2\cdot 1,j}^2], \tag{20}$$

$$\mathrm{E}[R_3^2] = \frac{4}{N_1}\mathrm{E}[U_{1,i}^2 U_{2\cdot 1,i}^2] + 4\left(1 - \frac{1}{N_1}\right)\mathrm{E}[U_{1,i} U_{2\cdot 1,i} U_{1,j} U_{2\cdot 1,j}]. \tag{21}$$

For first terms on the right side of (20) and (21), we obtain

$$\mathrm{E}[U_{2\cdot1,i}^4] = p_2(p_2+2)(p_2+4)(p_2+6) + \mathrm{O}(N^{-1}),\ \mathrm{E}[U_{1,i}^2 U_{2\cdot1,i}^2] = p_1p_2(p_1+2)(p_2+2) + \mathrm{O}(N^{-1}).$$

Furthermore, in the second terms on the right side of (20) and (21), to avoid dependence among random variables, let

$$\overline{\boldsymbol{x}}_T^{(i,j)} = \frac{1}{\sqrt{N-2}} \sum_{\substack{\alpha=1 \\ \alpha\neq i \\ \alpha \neq j}}^{N} \boldsymbol{x}_{1,\alpha},\ \boldsymbol{S}_T^{(i,j)} = \frac{1}{N-2} \sum_{\substack{\alpha=1 \\ \alpha\neq i \\ \alpha \neq j}}^{N} (\boldsymbol{x}_{1,\alpha} - \overline{\boldsymbol{x}}_T^{(i,j)})(\boldsymbol{x}_{1,\alpha} - \overline{\boldsymbol{x}}_T^{(i,j)})^\top.$$

That is, $\boldsymbol{x}_T^{(i,j)}$ and $\boldsymbol{S}_T^{(i,j)}$ are the sample mean vector and sample covariance matrix with $\boldsymbol{x}_{1,i}$ and $\boldsymbol{x}_{1,j}$ removed from $\boldsymbol{x}_{1,1},\ldots,\boldsymbol{x}_{1,N}$. Therefore, we can write

$$\begin{aligned}
\overline{\boldsymbol{x}}_{1,i} - \overline{\boldsymbol{x}}_T &= \left(1 - \frac{1}{N}\right)(\boldsymbol{x}_{1,i} - \overline{\boldsymbol{x}}_T^{(i,j)}) - \frac{1}{N}(\boldsymbol{x}_{1,j} - \overline{\boldsymbol{x}}_T^{(i,j)}), \\
\boldsymbol{S}_T^{-1} &= \{\boldsymbol{S}_T^{(i,j)}\}^{-1} + \frac{1}{N}\Big[2\{\boldsymbol{S}_T^{(i,j)}\}^{-1} - \{\boldsymbol{S}_T^{(i,j)}\}^{-1}(\boldsymbol{x}_{1,j} - \overline{\boldsymbol{x}}_T^{(i,j)})(\boldsymbol{x}_{1,j} - \overline{\boldsymbol{x}}_T^{(i,j)})^\top\{\boldsymbol{S}_T^{(i,j)}\}^{-1} \\
&\quad - \{\boldsymbol{S}_T^{(i,j)}\}^{-1}(\boldsymbol{x}_{1,j} - \overline{\boldsymbol{x}}_T^{(i,j)})(\boldsymbol{x}_{1,j} - \overline{\boldsymbol{x}}_T^{(i,j)})^\top\{\boldsymbol{S}_T^{(i,j)}\}^{-1}\Big] + \mathrm{O}_p(N^{-2}).
\end{aligned}$$

Furthermore, $\overline{\boldsymbol{x}}_T^{(i,j)}$ and $\boldsymbol{S}_T^{(i,j)}$ can be written as

$$\begin{aligned}
\overline{\boldsymbol{x}}_T^{(i,j)} &= \frac{N_1-2}{N-2}\overline{\boldsymbol{x}}_{(1),1}^{(i,j)} + \frac{N_2}{N-1}\overline{\boldsymbol{x}}_{(2)}, \\
\boldsymbol{S}_T^{(i,j)} &= \frac{N_1-2}{N-2}\boldsymbol{S}_{(1),11}^{(i,j)} + \frac{N_1-2}{N-2}(\overline{\boldsymbol{x}}_{(1),1}^{(i,j)} - \overline{\boldsymbol{x}}_T^{(i,j)})(\overline{\boldsymbol{x}}_{(1),1}^{(i,j)} - \overline{\boldsymbol{x}}_T^{(i,j)})^\top \\
&\quad + \frac{N_2}{N-2}\boldsymbol{S}_{(2)} + \frac{N_2}{N-2}(\overline{\boldsymbol{x}}_{(2)} - \overline{\boldsymbol{x}}_T^{(i,j)})(\overline{\boldsymbol{x}}_{(2)} - \overline{\boldsymbol{x}}_T^{(i,j)})^\top,
\end{aligned}$$

and by using

$$\overline{\boldsymbol{x}}_{(1)}^{(i,j)} = \begin{pmatrix} \overline{\boldsymbol{x}}_{(1),1}^{(i,j)} \\ \overline{\boldsymbol{x}}_{(1),2}^{(i,j)} \end{pmatrix} = \frac{1}{\sqrt{N_1-2}} \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \end{pmatrix},$$

$$\boldsymbol{S}_{(1)}^{(i,j)} = \begin{pmatrix} \boldsymbol{S}_{(1),11}^{(i,j)} & \boldsymbol{S}_{(1),12}^{(i,j)} \\ \boldsymbol{S}_{(1),21}^{(i,j)} & \boldsymbol{S}_{(1),22}^{(i,j)} \end{pmatrix} = \left(1 - \frac{1}{N_1-2}\right)\left\{\begin{pmatrix} \boldsymbol{I}_{p_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{p_2} \end{pmatrix} + \frac{1}{\sqrt{N_1-2}}\begin{pmatrix} \boldsymbol{W}_{11} & \boldsymbol{W}_{12} \\ \boldsymbol{W}_{21} & \boldsymbol{W}_{22} \end{pmatrix}\right\},$$

where

$$\overline{\boldsymbol{x}}_{(1)}^{(i,j)} = \frac{1}{N_1-2} \sum_{\substack{\alpha=1 \\ \alpha\neq i \\ \alpha \neq j}}^{N_1} \boldsymbol{x}_{\alpha},\ \boldsymbol{S}_{(1)}^{(i,j)} = \frac{1}{N_1-2} \sum_{\substack{\alpha=1 \\ \alpha\neq i \\ \alpha \neq j}}^{N_1} (\boldsymbol{x}_{\alpha} - \overline{\boldsymbol{x}}_{(1)}^{(i,j)})(\boldsymbol{x}_{\alpha} - \overline{\boldsymbol{x}}_{(1)}^{(i,j)})^\top,$$

$U_{1,i}$ and $U_{2\cdot 1,i}$ are expanded as

$$U_{1,i} = \boldsymbol{x}_{1,i}^\top \boldsymbol{x}_{1,i} - \frac{1}{\sqrt{N}} C_1 + \frac{1}{N} C_2 + \mathrm{O}_p(N^{-\frac{3}{2}}),\ U_{2\cdot 1,i} = \boldsymbol{x}_{2,i}^\top \boldsymbol{x}_{2,i} - \frac{1}{\sqrt{N}} D_1 + \frac{1}{N} D_2 + \mathrm{O}_p(N^{-\frac{3}{2}}),$$

where

$$\begin{aligned}
C_1 &= \sqrt{r}(2\boldsymbol{x}_{1,i}^\top \boldsymbol{u}_1 + \boldsymbol{x}_{1,i}^\top \boldsymbol{W}_{11} \boldsymbol{x}_{1,i}) + \sqrt{1-r}(2\boldsymbol{x}_{1,i}^\top \boldsymbol{z}_3 + \boldsymbol{x}_{1,i}^\top \boldsymbol{Z} \boldsymbol{x}_{1,i}), \\
C_2 &= 2\boldsymbol{x}_{1,i}^\top \boldsymbol{x}_{1,i} - 2\boldsymbol{x}_{1,i}^\top \boldsymbol{x}_{1,j} - (\boldsymbol{x}_{1,i}^\top \boldsymbol{x}_{1,i})^2 - (\boldsymbol{x}_{1,i}^\top \boldsymbol{x}_{1,j})^2 \\
&\quad + \sqrt{r(1-r)}(2\boldsymbol{u}_1^\top \boldsymbol{z}_3 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{u}_1 \boldsymbol{z}_3^\top \boldsymbol{x}_{1,i} + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{W}_{11} \boldsymbol{z}_3 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{Z} \boldsymbol{u}_1 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{W}_{11} \boldsymbol{Z} \boldsymbol{x}_{1,i}) \\
&\quad + r\left\{\boldsymbol{u}_1^\top \boldsymbol{u}_1 - (\boldsymbol{x}_{1,i}^\top \boldsymbol{z}_3)^2 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{W}_{11} \boldsymbol{u}_1 + \boldsymbol{x}_{1,i}^\top \boldsymbol{W}_{11}^2 \boldsymbol{x}_{1,i}\right\} \\
&\quad + (1-r)\left\{\boldsymbol{z}_3^\top \boldsymbol{z}_3 - (\boldsymbol{x}_{1,i}^\top \boldsymbol{u}_1)^2 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{Z} \boldsymbol{z}_3 + \boldsymbol{x}_{1,i}^\top \boldsymbol{Z}^2 \boldsymbol{x}_{1,i}\right\}, \\
D_1 &= \frac{1}{\sqrt{r}}(2\boldsymbol{x}_{2,i}^\top \boldsymbol{u}_2 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{W}_{12} \boldsymbol{x}_{2,i} + \boldsymbol{x}_{2,i}^\top \boldsymbol{W}_{22} \boldsymbol{x}_{2,i}), \\
D_2 &= \frac{1}{r}\{\boldsymbol{x}_{2,i}^\top \boldsymbol{x}_{2,i} - 2\boldsymbol{x}_{2,i}^\top \boldsymbol{x}_{2,j} + \boldsymbol{u}_2^\top \boldsymbol{u}_2 - 2\boldsymbol{x}_{1,i}^\top \boldsymbol{x}_{1,i} \boldsymbol{x}_{2,i}^\top \boldsymbol{x}_{2,i} - 2\boldsymbol{x}_{1,i}^\top \boldsymbol{x}_{1,j} \boldsymbol{x}_{2,j}^\top \boldsymbol{x}_{2,i} - (\boldsymbol{x}_{2,i}^\top \boldsymbol{x}_{2,i})^2 - (\boldsymbol{x}_{2,i}^\top \boldsymbol{x}_{2,j})^2 \\
&\quad + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{W}_{12} \boldsymbol{u}_2 + 2\boldsymbol{x}_{2,i}^\top \boldsymbol{W}_{21} \boldsymbol{u}_1 + 2\boldsymbol{x}_{2,i}^\top \boldsymbol{W}_{22} \boldsymbol{u}_2 + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{W}_{11} \boldsymbol{W}_{12} \boldsymbol{x}_{2,i} + \boldsymbol{x}_{1,i}^\top \boldsymbol{W}_{12} \boldsymbol{W}_{21} \boldsymbol{x}_{1,i} \\
&\quad + 2\boldsymbol{x}_{1,i}^\top \boldsymbol{W}_{12} \boldsymbol{W}_{22} \boldsymbol{x}_{2,i} + \boldsymbol{x}_{2,i}^\top \boldsymbol{W}_{21} \boldsymbol{W}_{12} \boldsymbol{x}_{2,i} + \boldsymbol{x}_{2,i}^\top \boldsymbol{W}_{22}^2 \boldsymbol{x}_{2,i}\}.
\end{aligned}$$

Expansions of $U_{1,j}$ and $U_{2\cdot 1,j}$ are obtained by replacing the subscript i in $U_{1,i}$ and $U_{2\cdot 1,j}$ with the subscript j, respectively. Therefore, calculating the expectation of $U_{2\cdot 1,i}^2 U_{2\cdot 1,j}^2$ with respect to $\boldsymbol{x}_{1,i}$, $\boldsymbol{x}_{1,j}$, $\boldsymbol{x}_{2,i}$, $\boldsymbol{x}_{2,j}$, $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, and $\boldsymbol{W}$, and the expectation of $U_{1,i} U_{1,j} U_{2\cdot 1,i} U_{2\cdot 1,j}$ with respect to $\boldsymbol{x}_{1,i}$, $\boldsymbol{x}_{1,j}$, $\boldsymbol{x}_{2,i}$, $\boldsymbol{x}_{2,j}$, $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{z}_3$, $\boldsymbol{W}$, and $\boldsymbol{Z}$, we obtain

$$\begin{aligned}
\mathrm{E}[U_{2\cdot 1,i}^2 U_{2\cdot 1,j}^2] &= p_2^2(p_2+2)^2 - \frac{4}{rN} p_2(p_2+2)^3 + \mathrm{O}(N^{-\frac{3}{2}}), \\
\mathrm{E}[U_{1,i} U_{1,j} U_{2\cdot 1,i} U_{2\cdot 1,j}] &= p_1^2 p_2^2 - \frac{2}{N} p_1 p_2^2 - \frac{2}{rN} p_1^2 p_2 (2p_2+1) + \mathrm{O}(N^{-\frac{3}{2}}).
\end{aligned}$$

Thus, we obtain (12).

Acknowledgments

The second author was partly supported by a Grant-in-Aid for Scientific Research (C) (JSPS KAKENHI Grant Number JP21K11795).

References

[1] Anderson, T. W. and Olkin, I. (1985). Maximum-likelihood estimation of the parameters of a multivariate normal distribution. *Linear Algebra and Its Applications*, **70**, 147–171.

[2] Enomoto, R., Hanusz, Z., Hara, A. and Seo, T. (2020). Multivariate normality test using normalizing transformation for Mardia's multivariate kurtosis. *Communications in Statistics – Simulation and Computation*, **49**, 684–698.

[3] Farrel, P. J., Salibian-Barrera, M. and Naczk, K. (2007). On tests for multivariate normality and associated simulation studies. *Journal of Statistical Computation and Simulation*, **77**, 1065–1080.

[4] Jinadasa, K. G., and Tracy, D. S. (1992). Maximum likelihood estimation for multivariate normal distribution with monotone sample. *Communications in Statistics – Theory and Methods*, **21**(1), 41–50.

[5] Hanusz, Z., Enomoto, R., Seo, T. and Koizumi, K. (2018). A Monte Carlo comparison of jarque-Bera type test and Henze-Zirkler test of multivariate normality. *Communications in Statistics – Simulation and Computation*, **47**, 1439–1452.

[6] Hao, J., Krishnamoorthy, K. (2001). Inferences on a normal covariance matrix and generalized variance with monotone missing data. *Journal of Multivariate Analysis*, **78**, 62–82.

[7] Henze, N. (1994). On Mardia's kurtosis test for multivariate normality. *Communications in Statistics – Theory and Methods*, **23**, 1047–1061.

[8] Henze, N. and Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality. *Communications in Statistics – Theory and Methods*, **19**, 3595–3617.

[9] Kanda, T. and Fujikoshi, Y. (1998). Some basic properties of the MLE's for a multivariate normal distribution with monotone missing data. *American Journal of Mathematical and Management Sciences*, **18**, 161–190.

[10] Kawasaki, T. and Seo, T. (2016). Bias correction for T^2 type statistic with two-step monotone missing data. *Statistics*, **50**, 76–88.

[11] Kawasaki, T., Shutoh, N. and Seo, T. (2018). On the asymptotic distribution of T^2-type statistic with two-step monotone missing data. *Journal of Statistical Theory and Practice*, **12**, 657–668.

[12] Kollo, T. (2008). Multivariate skewness and kurtosis measures with an application in ICA. *Journal of Multivariate Analysis*, **99**, 2328–2338.

[13] Koziol, J. A. (1989). A note on measures of multivariate kurtosis. *Biometrical Journal*, **31**, 619–624.

[14] Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, **57**, 519–530.

[15] Rao, S. J., Taufer, E. and Terdik, G, H. (2021). On multivariate skewness and kurtosis. *Shnkhyā*, **83**, 607–644.

[16] Siotani, M., Hayakawa, T. and Fujikoshi, Y (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook.* Columbus, OH: American Science Press, Inc.

[17] Srivastava, M. S. (1984). A measure of skewness and kurtosis and a graphical method for assessing multivariate normality. *Statistics & Probability Letters*, **2**, 263–267.

[18] Thode, H. C. (2002). *Testing for Normality.* New York, NY: Marcel Dekker, Inc.

[19] Tsukada, S. (2014). Equivalence testing of mean vector and covariance matrix for multi-populations under a two-step monotone incomplete sample. *Journal of Multivariate Analysis*, **132**, 183–196.

[20] Yagi, A., Seo, T. and Hanusz, Z. (2019). Improved simplified T^2 test statistics for a mean vector with monotone missing data. *Communications in Statistics – Simulation and Computation*, **48**, 1606–1622.

[21] Yamada, T., Romer, M. M. and Richards, D. St. P. (2015). Kurtosis tests for multivariate normality with monotone incomplete data. *Test*, **24**, 532–557.

[22] Zhou, M., and Shao, Y. (2014). A powerful test for multivariate normality. *Journal of Applied Statistics*, **41**, 351–363.