Unbiased estimators for the covariance matrix under a monotone incomplete sample

TSUKADA Shin-ichi

Abstract

In this article, we consider an inference for a covariance matrix under monotone incomplete sample. The maximum likelihood estimator for a mean vector is unbiased but that of the covariance matrix is not unbiased. We derive an unbiased estimator for the covariance matrix using some fundamental properties of Wishart matrix. The accuracy of the estimators is investigated by numerical simulation.

1 Introduction

There are situations where some variates are missing in multivariate statistical analysis, for example, some of the variables to be measured are too expensive. The problem of missing data is an important applied problem. For analyzing these data, various statistical methods have been developed by Anderson [1], Anderson and Olkin [2], Dempster, Laird and Rubin [5], Srivastava [13] and Little and Rubin [9].

In this paper, we consider a k-step monotone incomplete sample. Let $\boldsymbol{x}$ be distributed as $N_p(\boldsymbol{\mu}, \Sigma)$, and $\boldsymbol{x}^{(i)}$ the vector of the first q_i elements of $\boldsymbol{x}$, where $p = q_1 > q_2 > \cdots > q_k > 0$. We partition $\boldsymbol{x}$ as

$$\boldsymbol{x} = (\boldsymbol{x}_1', \ldots, \boldsymbol{x}_k')', \quad \boldsymbol{x}_i : p_i \times 1,$$

and $p_1 + \cdots + p_i = q_{k-i+1}$ for $i = 1, \ldots, k$. Then

$$\boldsymbol{x}^{(1)} = (\boldsymbol{x}_1', \ldots, \boldsymbol{x}_k')', \boldsymbol{x}^{(2)} = (\boldsymbol{x}_1', \ldots, \boldsymbol{x}_{k-1}')', \ldots, \boldsymbol{x}^{(k)} = \boldsymbol{x}_1.$$

Suppose that we have N_1 observations on $\boldsymbol{x}^{(1)}$, N_2 observations on $\boldsymbol{x}^{(2)}$, and so on. Let $\boldsymbol{x}_j^{(i)}$ be the j-th observation on $\boldsymbol{x}^{(i)}$. Here it is assumed that the marginal density function of the observed data set $\{\boldsymbol{x}_1^{(1)}, \ldots, \boldsymbol{x}_{N_1}^{(1)}, \ldots, \boldsymbol{x}_1^{(k)}, \ldots, \boldsymbol{x}_{N_k}^{(k)},\}$ is

$$\prod_{i=1}^{k} \prod_{j=1}^{N_i} f(\boldsymbol{x}_j^{(i)} | \boldsymbol{\mu}_{[i]}, \Sigma_{(1,\ldots,i)(1,\ldots,i)}), \tag{1}$$

where $f(\boldsymbol{x}_j^{(i)} | \boldsymbol{\mu}_{[i]}, \Sigma_{(1,\ldots,i)(1,\ldots,i)})$ is the density function of $N_{q_{k-i+1}}\left(\boldsymbol{\mu}_{[i]}, \Sigma_{(1,\ldots,i)(1,\ldots,i)}\right)$ and

$$\boldsymbol{\mu}_{[i]} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \vdots \\ \boldsymbol{\mu}_i \end{pmatrix}, \qquad \Sigma_{(1,\ldots,i)(1,\ldots,i)} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1i} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{i1} & \Sigma_{i2} & \cdots & \Sigma_{ii} \end{pmatrix},$$

where $\boldsymbol{\mu}_j$ is a p_j-dimensional vector and Σ_{jl} is a (p_j, p_l) matrix.

Anderson and Olkin [2] consider the 2-step monotone sample and derive the maximum likelihood estimator(MLE) $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$ for the mean vector $\boldsymbol{\mu}$ and the covariance matrix Σ based on the density function (1). Fujisawa [6] has obtained the estimators by the conditional approach. Kanda and Fujikoshi [8] investigate some fundamental properties of the MLE, and indicate that the MLE $\hat{\boldsymbol{\mu}}$ for the mean vector is unbiased but the MLE $\hat{\Sigma}$ for the covariance matrix is not unbiased. In general, it becomes difficult to derive the exact properties of these estimators except for some special cases. They study the asymptotic properties. Chang and Richards [3], [4] derive a stochastic representation for the exact distribution of the MLE $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$ under the 2-step monotone sample. They obtain ellipsoidal confidence region for $\boldsymbol{\mu}$ and deal with the hypothesis testing for the covariance matrix. Provost [11] considers the mutual independence of covariance matrix under the 2-step monotone incomplete sample and derives the likelihood ratio criterion. Hao and Krishnamoorthy [7] deal with the hypothesis testing that the covariance matrix is equal to a specified matrix and that the mean vector and the covariance matrix equal to a given vector and matrix under the k-step monotone incomplete sample. They derive the likelihood ratio criteria and asymptotic null distribution.

For $k = 2$ or $k = 3$, we derive an unbiased estimator for the covariance matrix using some fundamental properties by Kanda and Fujikoshi [8]. We deal with the 2-step and the 3-step monotone incomplete sample in Section 2 and 3, respectively. In Section 4, the accuracy of the unbiased estimators is investigated by numerical simulation.

2 2-step monotone incomplete data

Let the p-dimensional variate $\boldsymbol{x}$ be decomposed as $(\boldsymbol{x}_1', \boldsymbol{x}_2')$, where $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are p_1 and p_2-dimensional vectors, respectively. Suppose that we have N_1 observations on the full set of variables, i.e., $\boldsymbol{x}$, and N_2 observations on $\boldsymbol{x}_1$, and these observations are independently distributed. That is, we have the following observations:

$$\begin{pmatrix} \boldsymbol{x}_{11} \\ \boldsymbol{x}_{21} \end{pmatrix}, \begin{pmatrix} \boldsymbol{x}_{12} \\ \boldsymbol{x}_{22} \end{pmatrix}, \cdots \begin{pmatrix} \boldsymbol{x}_{1N_1} \\ \boldsymbol{x}_{2N_1} \end{pmatrix}, \begin{pmatrix} \boldsymbol{x}_{1N_1+1} \\ * \end{pmatrix}, \cdots \begin{pmatrix} \boldsymbol{x}_{1N_1+N_2} \\ * \end{pmatrix}.$$

Let $\bar{\boldsymbol{x}}^{(1)}$ denote the sample mean of $\boldsymbol{x}$ based on the N_1 observations, and $\bar{\boldsymbol{x}}^{(1)} = (\bar{\boldsymbol{x}}_1^{(1)\prime}, \bar{\boldsymbol{x}}_2^{(1)\prime})'$, $\bar{\boldsymbol{x}}_i^{(1)}$: $p_i \times 1$. Let $\bar{\boldsymbol{x}}^{(2)}$ denote the sample mean of the p_1-dimensional elements of $\boldsymbol{x}_1$ based on the N_2 observations. Throughout this article, we use the letter α only as running suffix for sample observations. Then the sample covariance matrix based on the N_1 and N_2 observations are expressed as

$$S^{(1)} = \frac{1}{n_1}\sum_{\alpha=1}^{N_1}(\boldsymbol{x}_\alpha^{(1)} - \bar{\boldsymbol{x}}^{(1)})(\boldsymbol{x}_\alpha^{(1)} - \bar{\boldsymbol{x}}^{(1)})', \qquad S^{(2)} = \frac{1}{n_2}\sum_{\alpha=1}^{N_2}(\boldsymbol{x}_\alpha^{(2)} - \bar{\boldsymbol{x}}^{(2)})(\boldsymbol{x}_\alpha^{(2)} - \bar{\boldsymbol{x}}^{(2)})',$$

respectively, where $n_i = N_i - 1$, $i = 1, 2$. Let the partitions of $\boldsymbol{\mu}$, Σ and $S^{(1)}$ corresponding to the ones of $\boldsymbol{x}$ be

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \quad S^{(1)} = \begin{pmatrix} S_{11}^{(1)} & S_{12}^{(1)} \\ S_{21}^{(1)} & S_{22}^{(1)} \end{pmatrix}.$$

2.1 Maximum likelihood estimator

Let the maximum likelihood estimator of $\boldsymbol{\mu}$ and Σ denote by $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$, respectively, which are partitioned in the same way as $\boldsymbol{\mu}$ and Σ. We can represent $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$ as follows:

$$\hat{\boldsymbol{\mu}}_1 = \frac{1}{N}\left(N_1\bar{\boldsymbol{x}}_1^{(1)} + N_2\bar{\boldsymbol{x}}^{(2)}\right), \quad \hat{\boldsymbol{\mu}}_2 = \bar{\boldsymbol{x}}_2^{(1)} - \hat{\Sigma}_{21}\hat{\Sigma}_{11}^{-1}\left(\bar{\boldsymbol{x}}_1^{(1)} - \hat{\boldsymbol{\mu}}_1\right),$$

$$\hat{\Sigma}_{11} = \frac{1}{N}\left(W_{11}^{(1)} + W^{(2)}\right), \quad \hat{\Sigma}_{12} = \hat{\Sigma}_{11}\left(W_{11}^{(1)}\right)^{-1}W_{12}^{(1)}, \quad \hat{\Sigma}_{22} = \frac{1}{N_1}W_{22\cdot1}^{(1)} + \hat{\Sigma}_{21}\hat{\Sigma}_{11}^{-1}\hat{\Sigma}_{12}, \quad (2)$$

where $N = N_1 + N_2$,

$$W^{(1)} = n_1S^{(1)}, \quad W^{(2)} = n_2S^{(2)} + \frac{N_1N_2}{N}\left(\bar{\boldsymbol{x}}_1^{(1)} - \bar{\boldsymbol{x}}^{(2)}\right)\left(\bar{\boldsymbol{x}}_1^{(1)} - \bar{\boldsymbol{x}}^{(2)}\right)',$$

$$W^{(1)} = \begin{pmatrix} W_{11}^{(1)} & W_{12}^{(1)} \\ W_{21}^{(1)} & W_{22}^{(1)} \end{pmatrix}, \quad W_{22\cdot1}^{(1)} = W_{22}^{(1)} - W_{21}^{(1)}\left(W_{11}^{(1)}\right)^{-1}W_{12}^{(1)}.$$

Kanda and Fujikoshi [8] prove the following lemma which is useful in deriving the expectation and variance of $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$.

Lemma 2.1. *Suppose that A is distributed as a Wishart distribution $W_p(\Sigma, n)$ and $n \geq p$, where A is partitioned as in the partition of Σ, and let $A_{22\cdot1} = A_{22} - A_{21}A_{11}^{-1}A_{12}$. Then we have*

(i) $A_{22\cdot1} \sim W_{p_2}(\Sigma_{22\cdot1}, n - p_1)$, and $A_{22\cdot1}$ is independent of A_{11} and A_{12},

(ii) The conditional distribution of $\mathrm{vec}(A_{12})$ given A_{11} is distributed as normal distribution with the mean vector $\mathrm{vec}(A_{11}\Sigma_{11}^{-1}\Sigma_{12})$ and the covariance matrix $\Sigma_{22\cdot1}\otimes A_{11}$. In particular $E[A_{11}^{-1}A_{12}] = \Sigma_{11}^{-1}\Sigma_{12}$, where $\mathrm{vec}(C)$ denotes the column vector formed by stacking the columns of C under each other,

(iii) $A_{11} \sim W_{p_1}(\Sigma_{11}, n)$,

(iv) if $n - p - 1 > 0$, then $E[A^{-1}] = \dfrac{1}{n-p-1}\Sigma^{-1}$

(v) if $n - p - 1 > 0$, then $E[A_{21}A_{11}^{-1}CA_{11}^{-1}A_{12}] = E[trA_{11}^{-1}C]\Sigma_{22\cdot1} + \Sigma_{21}\Sigma_{11}^{-1}E[C]\Sigma_{11}^{-1}\Sigma_{12}$, where C is a random matrix depending on A_{11}.

Proof. The results from (i) to (iv) are well known. For a proof, see Muirhead[10] and Siotani, Hayakawa and Fujikoshi[12]. The result (v) follows from (ii) and (iv). □

Using the above lemma, the expectation and variance of $\hat{\boldsymbol{\mu}}$ and the expectation of $\hat{\Sigma}$ are obtained as follows:

Theorem 2.1 (Kanda and Fujikoshi [8])**.** *Suppose that $N_1 > p$. Then the mean and the covariance matrix of $\hat{\boldsymbol{\mu}}$ and the mean of $\hat{\Sigma}$ are given by*

(i) $E[\hat{\boldsymbol{\mu}}] = \boldsymbol{\mu}$,

(ii) $\mathrm{Var}[\hat{\boldsymbol{\mu}}] = \dfrac{1}{N}\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & N\mathrm{Var}[\hat{\boldsymbol{\mu}}_2] \end{pmatrix}, \quad (N_1 > p_1 + 2),$

where

$$\mathrm{Var}[\hat{\boldsymbol{\mu}}_2] = \frac{1}{N_1}\left(\Sigma_{22} - \frac{N_2}{N}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right) + \frac{N_2p_1}{NN_1(N_1 - p_1 - 2)}\Sigma_{22\cdot1},$$

(iii) $E[\hat{\Sigma}] = \frac{N-1}{N}\Sigma + \frac{1}{N}\begin{pmatrix} O & O \\ O & b_0\Sigma_{22\cdot 1} \end{pmatrix}$,

where

$$b_0 = -\frac{N_2\left\{N_1 - (p_1+1)(p_1+2)\right\}}{N_1(N_1 - p_1 - 2)}.$$

Proof. The results for the mean vector are derived by using $\bar{\boldsymbol{x}}^{(1)} \sim N(\boldsymbol{\mu}, N_1^{-1}\Sigma)$, $\bar{\boldsymbol{x}}^{(2)} \sim N(\boldsymbol{\mu}_1, N_2^{-1}\Sigma_{11})$ and Lemma 2.1. The result for the covariance matrix is derived by using $W^{(1)} \sim W_p(\Sigma, N_1 - 1)$, $W^{(2)} \sim W_{p_1}(\Sigma_{11}, N_2)$, and Lemma 2.1 and that $W^{(1)}$ and $W^{(2)}$ are independently distributed. □

We may see that $\hat{\boldsymbol{\mu}}$ is unbiased and $\hat{\Sigma}$ is biased. Kanda and Fujikoshi [8] recommend a usual correction

$$\tilde{\Sigma} = \frac{N}{N-1}\hat{\Sigma} \tag{3}$$

for an estimator of Σ. In the next subsection, we will obtain the unbiased estimator of Σ.

2.2 Unbiased estimator of covariance matrix

From Theorem 2.1, the MLE $\hat{\Sigma}$ of the covariance matrix is biased. In this section, we obtain the unbiased estimator for Σ as follows.

Theorem 2.2. *Let*

$$\tilde{\tilde{\Sigma}} = \begin{pmatrix} \tilde{\Sigma}_{11} & \tilde{\Sigma}_{12} \\ \tilde{\Sigma}_{21} & \tilde{\tilde{\Sigma}}_{22} \end{pmatrix}, \tag{4}$$

where

$$\tilde{\tilde{\Sigma}}_{22} = \tilde{\Sigma}_{22} - c_0\hat{\Sigma}_{22\cdot 1}, \quad c_0 = \frac{N_2(p_1+1)(p_1+2) - N_1N_2}{(N-1)(N_1-p_1-2)(N_1-p_1-1)}.$$

Suppose that $N_1 > \max(p, p_1+2)$. *Then we have*

$$E\left[\tilde{\tilde{\Sigma}}\right] = \Sigma.$$

Proof. Since it is trivial that $E\left[\tilde{\Sigma}_{11}\right] = \Sigma_{11}$, $E\left[\tilde{\Sigma}_{12}\right] = \Sigma_{12}$ and $E\left[\tilde{\Sigma}_{21}\right] = \Sigma_{21}$, we prove that $E\left[\tilde{\tilde{\Sigma}}_{22}\right] = \Sigma_{22}$. It follows from Theorem2.1 (iii) that

$$E\left[\tilde{\Sigma}_{22}\right] = \left(1 + \frac{b_0}{N-1}\right)\Sigma_{22} - \frac{b_0}{N-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}. \tag{5}$$

Using the equation (2), the estimator $\hat{\Sigma}_{21}\hat{\Sigma}_{11}^{-1}\hat{\Sigma}_{12}$ is written as follows:

$$\hat{\Sigma}_{21}\hat{\Sigma}_{11}^{-1}\hat{\Sigma}_{12} = \frac{1}{N}W_{21}^{(1)}\left(W_{11}^{(1)}\right)^{-1}W_{12}^{(1)} + \frac{1}{N}W_{21}^{(1)}\left(W_{11}^{(1)}\right)^{-1}W^{(2)}\left(W_{11}^{(1)}\right)^{-1}W_{12}^{(1)}.$$

Since

$$E\left[W_{21}^{(1)}\left(W_{11}^{(1)}\right)^{-1}W_{12}^{(1)}\right] = p_1\Sigma_{22} + (N_1 - p_1 - 1)\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12},$$

$$E\left[W_{21}^{(1)}\left(W_{11}^{(1)}\right)^{-1}W^{(2)}\left(W_{11}^{(1)}\right)^{-1}W_{12}^{(1)}\right] = \frac{N_2p_1}{N_1-p_1-2}\Sigma_{22} + \frac{N_2(N_1-2p_1-2)}{N_1-p_1-2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$$

from Lemma2.1, one sees that

$$E\left[\hat{\Sigma}_{21}\hat{\Sigma}_{11}^{-1}\hat{\Sigma}_{12}\right] = \frac{p_1}{N}\frac{N-p_1-2}{N_1-p_1-2}\Sigma_{22} + \frac{B_0}{N(N_1-p_1-2)}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}, \tag{6}$$

where

$$B_0 = N_1^2 + N_1(N_2-2p_1-3) - (p_1+1)(2N_2-p_1-2).$$

From Theorem2.1 (iii), (5) and (6), we obtain that

$$E\left[\tilde{\tilde{\Sigma}}_{22}\right] = E\left[\tilde{\Sigma}_{22}\right] - c_0 E\left[\hat{\Sigma}_{22}\right] + c_0 E\left[\hat{\Sigma}_{21}\hat{\Sigma}_{11}^{-1}\hat{\Sigma}_{12}\right] = \Sigma_{22}.$$

□

The effect of the unbiasedness is investigated by the numerical simulation in Section 4.

3 3-step monotone incomplete data

As in Section 2, we consider the case of the 3-step monotone incomplete sample. Let the p-dimensional variate $\boldsymbol{x}$ be decomposed as $(\boldsymbol{x}_1', \boldsymbol{x}_2', \boldsymbol{x}_3')$, where $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$ are p_1, p_2 and p_3-dimensional vectors, respectively. Suppose that we have N_1 observations on $\boldsymbol{x}$, N_2 observations on $(\boldsymbol{x}_1', \boldsymbol{x}_2')'$, and N_3 observations on $\boldsymbol{x}_1$, and that these observations are independently distributed. That is, we have the following observations:

$$\begin{pmatrix} \boldsymbol{x}_{11} \\ \boldsymbol{x}_{21} \\ \boldsymbol{x}_{31} \end{pmatrix}, \cdots, \begin{pmatrix} \boldsymbol{x}_{1N_1} \\ \boldsymbol{x}_{2N_1} \\ \boldsymbol{x}_{3N_1} \end{pmatrix}, \begin{pmatrix} \boldsymbol{x}_{1N_1+1} \\ \boldsymbol{x}_{2N_1+1} \\ * \end{pmatrix}, \cdots, \begin{pmatrix} \boldsymbol{x}_{1N_1+N_2} \\ \boldsymbol{x}_{2N_1+N_2} \\ * \end{pmatrix},$$

$$\begin{pmatrix} \boldsymbol{x}_{1N_1+N_2+1} \\ * \\ * \end{pmatrix}, \cdots, \begin{pmatrix} \boldsymbol{x}_{1N_1+N_2+N_3} \\ * \\ * \end{pmatrix}.$$

Let $\bar{\boldsymbol{x}}^{(1)} = (\bar{\boldsymbol{x}}_1^{(1)\prime}, \bar{\boldsymbol{x}}_2^{(1)\prime}, \bar{\boldsymbol{x}}_3^{(1)\prime})'$ denote the sample mean of $\boldsymbol{x}$ based on the N_1 observations, and $\bar{\boldsymbol{x}}^{(2)} = (\bar{\boldsymbol{x}}_1^{(2)\prime}, \bar{\boldsymbol{x}}_2^{(2)\prime})'$ denote the sample mean for the first (p_1+p_2)-dimensional elements of $\boldsymbol{x}$ based on the N_2 observations, respectively, and $\bar{\boldsymbol{x}}^{(3)}$ denote the sample mean for p_1-dimensional elements of $\boldsymbol{x}_1$ based on the N_3 observations. The corresponding sample covariance matrices are denoted by $S^{(i)}$ $(i=1,2,3)$ as follows:

$$\begin{aligned} S^{(1)} &= \frac{1}{n_1}\sum_{\alpha=1}^{N_1}(\boldsymbol{x}_\alpha^{(1)} - \bar{\boldsymbol{x}}^{(1)})(\boldsymbol{x}_\alpha^{(1)} - \bar{\boldsymbol{x}}^{(1)})', \\ S^{(2)} &= \frac{1}{n_2}\sum_{\alpha=1}^{N_2}\begin{pmatrix} \boldsymbol{x}_{1\alpha}^{(2)} - \bar{\boldsymbol{x}}_1^{(2)} \\ \boldsymbol{x}_{2\alpha}^{(2)} - \bar{\boldsymbol{x}}_2^{(2)} \end{pmatrix}\begin{pmatrix} \boldsymbol{x}_{1\alpha}^{(2)} - \bar{\boldsymbol{x}}_1^{(2)} \\ \boldsymbol{x}_{2\alpha}^{(2)} - \bar{\boldsymbol{x}}_2^{(2)} \end{pmatrix}', \\ S^{(3)} &= \frac{1}{n_3}\sum_{\alpha=1}^{N_3}(\boldsymbol{x}_\alpha^{(3)} - \bar{\boldsymbol{x}}^{(3)})(\boldsymbol{x}_\alpha^{(3)} - \bar{\boldsymbol{x}}^{(3)})', \end{aligned}$$

where $n_i = N_i - 1,\ (i=1,2,3)$ and $N = N_1 + N_2 + N_3$.

Let the partitions of $\boldsymbol{\mu}$ and Σ corresponding to the ones of $\boldsymbol{x}$ be

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \boldsymbol{\mu}_3 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix} = \begin{pmatrix} \Sigma_{(12)(12)} & \Sigma_{(12)3} \\ \Sigma_{3(12)} & \Sigma_{33} \end{pmatrix}.$$

Similar partitions and notations are used for $S^{(i)}$ and for other matrices.

3.1 Maximum likelihood estimator

We can represent $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$ as follows:

$$\begin{aligned}
\hat{\boldsymbol{\mu}}_1 &= \frac{1}{N}\left(N_1\bar{\boldsymbol{x}}_1^{(1)} + N_2\bar{\boldsymbol{x}}_1^{(2)} + N_3\bar{\boldsymbol{x}}^{(3)}\right), \\
\hat{\boldsymbol{\mu}}_2 &= \frac{1}{N_1+N_2}\left(N_1\bar{\boldsymbol{x}}_2^{(1)} + N_2\bar{\boldsymbol{x}}_2^{(2)}\right) - F'\left(\frac{N_1\bar{\boldsymbol{x}}_1^{(1)} + N_2\bar{\boldsymbol{x}}_1^{(2)}}{N_1+N_2} - \hat{\boldsymbol{\mu}}_1\right), \\
\hat{\boldsymbol{\mu}}_3 &= \bar{\boldsymbol{x}}_3^{(1)} - G'\begin{pmatrix} \bar{\boldsymbol{x}}_1^{(1)} - \hat{\boldsymbol{\mu}}_1 \\ \bar{\boldsymbol{x}}_2^{(1)} - \hat{\boldsymbol{\mu}}_2 \end{pmatrix}, \\
\hat{\Sigma}_{11} &= \frac{1}{N}\left(W_{11}^{(1)} + W_{11}^{(2)} + W^{(3)}\right), \qquad \hat{\Sigma}_{12} = \hat{\Sigma}_{11}F, \\
\hat{\Sigma}_{22} &= \frac{1}{N_1+N_2}\left(W_{(12)(12)}^{(1)} + W^{(2)}\right)_{22\cdot 1} + F'\hat{\Sigma}_{11}F, \\
\hat{\Sigma}_{(12)3} &= \hat{\Sigma}_{(12)(12)}G, \qquad \hat{\Sigma}_{33} = \frac{1}{N_1}W_{33\cdot 12}^{(1)} + G'\hat{\Sigma}_{(12)(12)}G,
\end{aligned}$$

where

$$\begin{aligned}
F &= \left(W_{11}^{(1)} + W_{11}^{(2)}\right)^{-1}\left(W_{12}^{(1)} + W_{12}^{(2)}\right), \qquad G = \left(W_{(12)(12)}^{(1)}\right)^{-1}W_{(12)3}^{(1)}, \\
W^{(1)} &= n_1 S^{(1)}, \\
W^{(2)} &= n_2 S^{(2)} + \frac{N_1N_2}{N_1+N_2}\begin{pmatrix} \bar{\boldsymbol{x}}_1^{(1)} - \bar{\boldsymbol{x}}_1^{(2)} \\ \bar{\boldsymbol{x}}_2^{(1)} - \bar{\boldsymbol{x}}_2^{(2)} \end{pmatrix}\begin{pmatrix} \bar{\boldsymbol{x}}_1^{(1)} - \bar{\boldsymbol{x}}_1^{(2)} \\ \bar{\boldsymbol{x}}_2^{(1)} - \bar{\boldsymbol{x}}_2^{(2)} \end{pmatrix}', \\
W^{(3)} &= n_3 S^{(3)} + \frac{(N_1+N_2)N_3}{N}\left\{\bar{\boldsymbol{x}}^{(3)} - \frac{1}{N_1+N_2}\left(N_1\bar{\boldsymbol{x}}_1^{(1)} + N_2\bar{\boldsymbol{x}}_1^{(2)}\right)\right\} \\
&\qquad \times\left\{\bar{\boldsymbol{x}}^{(3)} - \frac{1}{N_1+N_2}\left(N_1\bar{\boldsymbol{x}}_1^{(1)} + N_2\bar{\boldsymbol{x}}_1^{(2)}\right)\right\}'.
\end{aligned}$$

The natural parameters in the conditional approach are defined by

$$\boldsymbol{\eta} = \begin{pmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \boldsymbol{\eta}_3 \end{pmatrix}, \quad \Delta = \begin{pmatrix} \Delta_{11} & \Delta_{12} & \Delta_{13} \\ \Delta_{21} & \Delta_{22} & \Delta_{23} \\ \Delta_{31} & \Delta_{32} & \Delta_{33} \end{pmatrix},$$

which are one to one correspondence to $(\boldsymbol{\mu}, \Sigma)$, where

$\Delta_{11} = \Sigma_{11}, \quad \Delta_{12} = \Delta_{21}' = \Sigma_{11}^{-1}\Sigma_{12}, \quad \Delta_{22} = \Sigma_{22\cdot 1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12},$

$\Delta_{(12)3} = \Delta_{3(12)}' = \Sigma_{(12)(12)}^{-1}\Sigma_{(12)3},$

$\Delta_{33} = \Sigma_{33\cdot 12} = \Sigma_{33} - \Sigma_{3(12)}\Sigma_{(12)(12)}^{-1}\Sigma_{(12)3}.$

The MLE $(\hat{\boldsymbol{\eta}}, \hat{\Delta})$ of $(\boldsymbol{\eta}, \Delta)$ are expressed as follows:

$$\hat{\boldsymbol{\eta}}_1 = \hat{\boldsymbol{\mu}}_1, \quad \hat{\boldsymbol{\eta}}_2 = \bar{\boldsymbol{x}}_2^{(1)} - \hat{\Delta}_{21}\bar{\boldsymbol{x}}_1^{(1)}, \quad \hat{\boldsymbol{\eta}}_3 = \bar{\boldsymbol{x}}_3^{(1)} - \hat{\Delta}_{3(12)} \begin{pmatrix} \bar{\boldsymbol{x}}_1^{(1)} \\ \bar{\boldsymbol{x}}_2^{(1)} \end{pmatrix},$$
$$\hat{\Delta}_{11} = \hat{\Sigma}_{11}, \quad \hat{\Delta}_{12} = \hat{\Delta}_{21}' = \left(W_{11}^{(1)} + W_{11}^{(2)}\right)^{-1}\left(W_{12}^{(1)} + W_{12}^{(2)}\right),$$
$$\hat{\Delta}_{22} = \frac{1}{N_1 + N_2}\left(W_{(12)(12)}^{(1)} + W^{(2)}\right)_{22\cdot 1},$$
$$\hat{\Delta}_{(12)3} = \left(W_{(12)(12)}^{(1)}\right)^{-1} W_{(12)3}^{(1)}, \quad \hat{\Delta}_{33} = \frac{1}{N_1} W_{33\cdot 12}^{(1)}.$$

The following lemma is published in Kanda and Fujikoshi [8] to calculate the expectation and variance of $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$.

Lemma 3.1. *Suppose that A, B and C is independently distributed as a Wishart distribution $W_p(\Sigma, n_1)$, $W_p(\Sigma, n_2)$ and $W_p(\Sigma, n_3)$, respectively. Let $n = n_1 + n_2 + n_3$ and let A, B, C and Σ be partitioned as in Lemma 2.1. Further, let $L = (A_{11} + B_{11})^{-1}(A_{12} + B_{12})$, and let*

$$D = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix},$$

where $D_{11} = A_{11} + B_{11} + C_{11}$, $D_{12} = D_{21}' = D_{11}L$, $D_{22} = \gamma A_{22\cdot 1} + L'D_{11}L$ and γ is a constant. Then we have

(i) $E[D_{11}] = n\Sigma_{11}$,

(ii) $E[D_{12}] = n\Sigma_{12}$,

(iii) if $n_1 + n_2 - p_1 - 1 > 0$, then
$$E[D_{22}] = \gamma(n - p_1)\Sigma_{22\cdot 1} + n\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} + p_1\left\{1 + \frac{n_3}{n_1 + n_2 - p_1 - 1}\right\}\Sigma_{22\cdot 1},$$

(iv) if $n_1 - p - 1 > 0$, then
$$E[trA^{-1}D] = p_1 + \gamma p_2 + \frac{(n_2 + n_3)p_1}{n_1 - p_1 - 1} + \frac{p_1 p_2}{n_1 - p_1 - 1}\left(\frac{n_2 + n_3}{n_1 - p_1 - 1} - \frac{n_3}{n_1 + n_2 - p_1 - 1}\right)$$

(v) if $n_1 - p - 1 > 0$, then
$$E\left[trA^{-1}\begin{pmatrix} I_{p_1} \\ L' \end{pmatrix}\Sigma_{11}\begin{pmatrix} I_{p_1} & L' \end{pmatrix}\right] = \frac{p_1}{n_1 - p_1 - 1} + \frac{p_1 p_2}{n_1 - p - 1}\left(\frac{1}{n_1 - p_1 - 1} - \frac{1}{n_1 + n_2 - p_1 - 1}\right).$$

Proof. The results (i), (ii) and (iii) are easily obtained from Lemma 2.1. For (iv) and (v), we evaluate the expectation by three steps; (1) $A_{22\cdot 1}$, (2) A_{12}, B_{12}, (3) A_{11}, B_{11}, C_{11} using the inverse matrix of the partition matrix. See Kanda and Fujikoshi [8] for details. □

The expectation of $\hat{\Sigma}$ is obtained as follows:

Theorem 3.1 (Kanda and Fujikoshi [8]). *Suppose that $N_1 > p$ and $N_1 > p_1 + p_2 + 2$. Then*

$$E\left[\hat{\Sigma}\right] = \frac{N-1}{N}\Sigma + \frac{1}{N}\begin{pmatrix} O & O & O \\ O & B_{22} & B_{23} \\ O & B_{32} & B_{33} \end{pmatrix}, \tag{7}$$

where $B_{22} = b_1\Sigma_{22\cdot 1}$, $B_{23} = B_{32}' = b_1\Sigma_{22\cdot 1}\Delta_{23}$, $B_{33} = b_1\Delta_{32}\Sigma_{22\cdot 1}\Delta_{23} + b_2\Sigma_{33\cdot 12}$,

$$\begin{aligned} b_1 &= -\frac{N_3\{N_1+N_2-(p_1+1)(p_2+2)\}}{N(N_1+N_2)(N_1+N_2-p_1-2)}, \\ b_2 &= -\frac{1}{N}\left[\frac{(N_2+N_3)\{N_1-(p_1+1)(p_1+p_2+2)\}}{N_1(N_1-p_1-p_2-2)} + \frac{p_2NN_2}{N_1(N_1+N_2)} \right. \\ &\quad \left. + \frac{p_1p_2N_3}{(N_1-p_1-p_2-2)(N_1+N_2-p_1-2)}\right]. \end{aligned}$$

Proof. To proof this theorem, it is used that $W^{(1)}$, $W^{(2)}$ and $W^{(3)}$ are independently distributed as $W_p(\Sigma, N_1-1)$, $W_{p_1+p_2}(\Sigma_{(12)(12)}, N_2)$ and $W_{p_1}(\Sigma_{11}, N_3)$, respectively. Lemma 2.1 and Lemma 3.1 are applied to evaluate the expectation. See Kanda and Fujikoshi [8] for details. □

3.2 Unbiased estimator of covariance matrix

From Theorem 3.1, the MLE $\hat{\Sigma}$ of the covariance matrix is biased as well as the 2-step monotone incomplete sample. In this section, we obtain the unbiased estimator for Σ in the 3-step monotone incomplete sample.

Theorem 3.2. *Let*

$$\begin{aligned} \tilde{\tilde{\Sigma}}_{22} &= \tilde{\Sigma}_{22} - c_0\hat{\Sigma}_{22\cdot 1}, \qquad \tilde{\tilde{\Sigma}}_{(12)(12)} = \begin{pmatrix} \tilde{\Sigma}_{11} & \tilde{\Sigma}_{12} \\ \tilde{\Sigma}_{21} & \tilde{\tilde{\Sigma}}_{22} \end{pmatrix}, \\ \tilde{\tilde{\Sigma}}_{(12)3} &= \tilde{\tilde{\Sigma}}_{(12)(12)}\left(W^{(1)}_{(12)(12)}\right)^{-1}W^{(1)}_{(12)3}, \quad \tilde{\tilde{\Sigma}}_{33} = c_1\hat{\Sigma}_{33} + c_2\hat{\Delta}_{32}\hat{\Sigma}_{22\cdot 1}\hat{\Delta}_{23} + c_3\hat{\Sigma}_{33\cdot 12}, \end{aligned}$$

where

$$\begin{aligned} c_0 &= \frac{Nb_1}{(N-1)(b_1+d_0)}, \\ c_1 &= \frac{N}{N-1} - \frac{N^2d_1(N_1-2)}{(N-1)(d_2-d_2N-2d_1N+d_1NN_1)}, \\ c_2 &= -\frac{b_1(d_1+d_2)(N_1+N_2)}{(N_1+N_2-p_1-1)\{d_2(N-1)-Nd_1(N_1-2)\}}, \\ c_3 &= \frac{d_1N}{d_2-d_2N-2d_1N+d_1NN_1}, \\ d_0 &= (N_1+N_2-1-p_1) + N_3\frac{N_1+N_2-2p_1-2p_2-2}{N_1+N_2-p_1-p_2-2}, \\ d_1 &= \frac{b_2(N_1+N_2-p_1-1)}{N(N_1+N_2)} - \frac{b_1p_2(N_1+N_2-p_1-p_2-2)}{N(N_1+N_2)(N_1-p_1-p_2-2)}, \\ d_2 &= \frac{b_1p_2(N_1+N_2-p_1-p_2-2)}{N(N_1+N_2)(N_1-p_1-p_2-2)}. \end{aligned}$$

Suppose that $N_1 > p$, $N_2+1 > p_1+p_2$, $N_3+1 > p_3$ *and* $N_1-p_1-p_2-2 > 0$. *Then we have*

$$E\left[\begin{pmatrix} \tilde{\tilde{\Sigma}}_{(12)(12)} & \tilde{\tilde{\Sigma}}_{(12)3} \\ \tilde{\tilde{\Sigma}}_{3(12)} & \tilde{\tilde{\Sigma}}_{33} \end{pmatrix}\right] = \Sigma.$$

Proof. Here, some typical expectations needed to evaluate the expectation of unbiased estimator are shown, and details are omitted. Applying Lemma 2.1, we get

$$E\left[\hat{\Sigma}_{21}\hat{\Sigma}_{11}^{-1}\hat{\Sigma}_{12}\right] = \left[\frac{N-1}{N} - \frac{1}{N}\left\{(N_1+N_2-1-p_1) + N_3\frac{N_1+N_2-2p_1-2p_2-2}{N_1+N_2-p_1-p_2-2}\right\}\right]\Sigma_{22} + \frac{1}{N}\left\{(N_1+N_2-1-p_1) + N_3\frac{N_1+N_2-2p_1-2p_2-2}{N_1+N_2-p_1-p_2-2}\right\}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}.$$

We can evaluate the expectation of $\tilde{\tilde{\Sigma}}_{22}$ using the above equation and Theorem 3.1.

The expectation of $\tilde{\tilde{\Sigma}}_{22}$ could be proved by a similar calculation for $\hat{\Sigma}_{(12)3}$ in Theorem 3.1.

$$\begin{aligned} E\left[\tilde{\tilde{\Sigma}}_{(12)3}\right] &= E\left[\tilde{\tilde{\Sigma}}_{(12)(12)}\left(W^{(1)}_{(12)(12)}\right)^{-1}W^{(1)}_{(12)3}\right] \\ &= E\left[\tilde{\tilde{\Sigma}}_{(12)(12)}\right]\Sigma^{-1}_{(12)(12)}\Sigma_{(12)3} \\ &= \Sigma_{(12)(12)}\Sigma^{-1}_{(12)(12)}\Sigma_{(12)3} = \Sigma_{(12)3}. \end{aligned}$$

Using Lemma 2.1 and Lemma 3.1, we get

$$\begin{aligned} E\left[\hat{\Sigma}_{33\cdot 12}\right] &= (N_1-2)\Sigma_{33} - \frac{b_1}{N}\Delta_{32}\Sigma_{22\cdot 1}\Delta_{23} \\ &\quad + \left\{\frac{1}{N_1}(N_1-p_1-p_2-1) - (N_1-2)\right\}\Sigma_{33\cdot 12}, \\ E\left[\hat{\Delta}_{32}\hat{\Sigma}_{22\cdot 1}\hat{\Delta}_{23}\right] &= \frac{(N_1+N_2-p_1-p_2-2)p_2}{(N_1+N_2)(N_1-p_1-p_2-2)}\Sigma_{33\cdot 12} + \frac{N_1+N_2-1-p_1}{N_1+N_2}\Delta_{32}\Sigma_{22\cdot 1}\Delta_{23}. \end{aligned}$$

The expectation of $\tilde{\tilde{\Sigma}}_{33}$ can be evaluated by the expectation of $\hat{\Sigma}_{33}$ in Theorem 3.1 and the above equations. □

We see that the estimator

$$\tilde{\tilde{\Sigma}} = \begin{pmatrix} \tilde{\tilde{\Sigma}}_{(12)(12)} & \tilde{\tilde{\Sigma}}_{(12)3} \\ \tilde{\tilde{\Sigma}}_{3(12)} & \tilde{\tilde{\Sigma}}_{33} \end{pmatrix}$$

is unbiased. As well as the 2-step incomplete sample, the effect of the unbiasedness is investigated by the numerical simulation in Section 4.

4 Numerical simulation

In this section, we evaluate the expectation of estimators by numerical simulation. Let $T = (t_{ij})$ be the estimator of covariance matrix, i.e., $\hat{\Sigma}$, $\tilde{\Sigma}$ and $\tilde{\tilde{\Sigma}}$. The expectation of the estimator T is evaluated by

$$\bar{T} = (\bar{t}_{ij}) = \frac{1}{n_s}\sum_{i=1}^{n_s} T_i,$$

where n_s is a number of simulation and T_i is the estimator in each simulation. We adopt

$$E = \sum_{\substack{i,j=1 \\ i\le j}}^{p} \left(\bar{t}_{ij} - \sigma_{ij}\right)^2$$

as the difference between $\bar{T}$ and $\Sigma = (\sigma_{ij})$. The number of simulation is a hundred thousand.

Let

$$P_p = \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{p-3} & \rho^{p-2} & \rho^{p-1} \\ \rho & 1 & \rho & \cdots & \rho^{p-4} & \rho^{p-3} & \rho^{p-2} \\ \rho^2 & \rho & 1 & & \rho^{p-5} & \rho^{p-4} & \rho^{p-3} \\ \vdots & \vdots & & \ddots & & \vdots & \vdots \\ \rho^{p-3} & \rho^{p-4} & \rho^{p-5} & & 1 & \rho & \rho^2 \\ \rho^{p-2} & \rho^{p-3} & \rho^{p-4} & \cdots & \rho & 1 & \rho \\ \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \cdots & \rho^2 & \rho & 1 \end{pmatrix}.$$

As the population covariance matrix, we adopt the following matrices

$$\begin{aligned} \Sigma_1 &= \Lambda_{1p} P_p \Lambda_{1p}, \qquad \text{(Case 1)} \\ \Sigma_2 &= \Lambda_{2p} P_p \Lambda_{2p}, \qquad \text{(Case 2)} \end{aligned}$$

where $\Lambda_{1p} = \mathrm{diag}(\sigma^{p-1}, \sigma^{p-2}, \ldots, \sigma^2, \sigma, 1)$ and $\Lambda_{2p} = \sigma I_p$, and assume that the population mean vector is $\mathbf{0}$.

4.1 2-step monotone incomplete sample

We assume that the population distribution is the 7-variate normal distribution. Let $p_1 = 4$ and $p_2 = 3$. We set $\rho = 0.15$ and $\sigma = \sqrt{2}$ in Case 1, $\rho = 0.5$ and $\sigma = 2$ in Case 2. The total sample size are 50, 100, 200, 500 and 1000. The missing rate $\tau(= N_2/N)$ are 0.2, 0.4, 0.6, 0.8. Table 1 represents the coefficient c_0 of the unbiased estimator. Table 2 - Table 6 represent the results in Case 1, and the results in Case 2 are represented in Table 7 - Table 11. In Table, E_1, E_2 and E_3 denote a partial error as follows:

$$E = \sum_{\substack{i,j=1 \\ i \le j}}^{p_1} (\bar{t}_{ij} - \sigma_{ij})^2 + \sum_{i=1}^{p_1} \sum_{j=p_1}^{p} (\bar{t}_{ij} - \sigma_{ij})^2 + \sum_{\substack{i,j=p_1+1 \\ i \le j}}^{p} (\bar{t}_{ij} - \sigma_{ij})^2 \equiv E_1 + E_2 + E_3,$$

which are the error for Σ_{11}, Σ_{12} and Σ_{22}, respectively, and the notation x^y denotes the value $x \times 10^y$.

As a matter of course, errors are small when the total sample size N is large. The error of $\tilde{\Sigma}$ and the unbiased estimator is smaller than that of the maximum likelihood estimator as a whole. The error of $\tilde{\Sigma}$ is close to that of the unbiased estimator in the case of $\tau = 0.2$, but the error of the unbiased estimator is smaller than that of $\tilde{\Sigma}$ in $\tau = 0.8$. There is a similar tendency in each Table, and the difference of errors is about one digit when $N = 1000$.

From these results, it is obvious that the unbiased estimator is more accurate. We recommend this estimator under the 2-step incomplete sample.

4.2 3-step monotone incomplete sample

We set the parameters of the population distribution as well as under the 2-step monotone incomplete sample. It assumes that the population distribution is the 9-variate normal distribution and $p_1 = p_2 = p_3 = 3$. We set $\rho = 0.15$ and $\sigma = \sqrt{2}$ in Case 1, $\rho = 0.5$ and $\sigma = 2$ in Case 2 as the 2-step monotone incomplete sample.

The missing rates (τ_1, τ_2, τ_3) are (0.77, 0.15, 0.08), (0.62, 0.25, 0.13), (0.52, 0.32, 0.16) and (0.46, 0.36, 0.18). The coefficients concerning the unbiased estimator is represented in Table 12. Table 13 - Table 15 represent the result in Case 1, and the results in Case 2 are represented in Table 16 - Table 18. In Table, from E_1 to E_6 denote a partial error as follows:

$$\begin{aligned} E &= \sum_{\substack{i,j=1\\ i\le j}}^{p_1} (\bar{t}_{ij}-\sigma_{ij})^2 + \sum_{i=1}^{p_1}\sum_{j=p_1+1}^{p_1+p_2} (\bar{t}_{ij}-\sigma_{ij})^2 + \sum_{i=1}^{p_1}\sum_{j=p_1+p_2+1}^{p} (\bar{t}_{ij}-\sigma_{ij})^2 \\ &\quad + \sum_{\substack{i,j=p_1+1\\ i\le j}}^{p_1+p_2} (\bar{t}_{ij}-\sigma_{ij})^2 + \sum_{i=p_1+1}^{p_1+p_2}\sum_{j=p_1+p_2+1}^{p} (\bar{t}_{ij}-\sigma_{ij})^2 + \sum_{\substack{i,j=p_1+p_2+1\\ i\le j}}^{p} (\bar{t}_{ij}-\sigma_{ij})^2 \\ &\equiv E_1+E_2+E_3+E_4+E_5+E_6. \end{aligned}$$

In this case, E_4, E_5 and E_6, which are the error for Σ_{22}, Σ_{23} and Σ_{33}, respectively, are expected to be improved by the correction. When $N = 100$ and $\tau_3 = 0.18$, E_6 is not improved and the total error E of the unbiased estimator is also not improved compared to the error of $\tilde{\Sigma}$. However, E_4, E_5 and E_6 are improved as the total sample size is large. Since the order of the coefficients b_1 is $O(N^{-1})$ and that of b_2 is $O(N^{-2})$ from Theorem 3.1, the improvement is small compared with the 2-step monotone incomplete sample. Even if coefficients c_2 and c_3 are looked, these are gotten. Coefficients c_2 and c_3 are small, and the improvement is also small.

Since the order of the coefficients c_0, c_1, c_2 and c_3 are $O(N^{-2})$, $O(1)$, $O(N^{-3})$ and $O(N^{-1})$, the unbiased estimator has asymptotically similar properties with the MLE.

5 Conclusion

We constitute the unbiased estimator for the covariance matrix Σ under the 2-step and the 3-step monotone incomplete sample. Numerical simulation show that the unbiased estimator improves the bias. The order of the bias under the 3-step monotone incomplete sample is $O(N^{-2})$ and that under the 2-step monotone incomplete sample is $O(N^{-1})$. Since the bias under the 3-step monotone incomplete sample is smaller than that under the 2-step monotone incomplete sample, the effect of a bias correction under the 3-step monotone incomplete sample is small.

Kanda and Fujikoshi [8] describe the unbiased estimator for Δ under the k-step monotone incomplete sample, but it seems that it may be difficult to constitute the unbiased estimator for Σ under the k-step monotone incomplete sample which will be a problem for the future. The order of the bias may be $O(N^{1-k})$ under the k-step monotone incomplete sample and the correction may not influence so much for the $k(\geq 4)$-step monotone incomplete sample.

Table 1: The coefficient c_0 concering the unbiase estimator

	$\tau=0.2$	$\tau=0.4$	$\tau=0.6$	$\tau=0.8$
$N=50$	-1.6660^{-3}	0.0000	2.7211^{-2}	6.5306^{-1}
$N=100$	-1.7957^{-3}	-4.0070^{-3}	-4.9474^{-3}	3.5915^{-2}
$N=200$	-1.0876^{-3}	-2.7358^{-3}	-5.3601^{-3}	-6.5634^{-3}
$N=500$	-4.7523^{-4}	-1.2435^{-3}	-2.6878^{-3}	-6.2174^{-3}
$N=1000$	-2.4391^{-4}	-6.4467^{-4}	-1.4243^{-3}	-3.5802^{-3}

Table 2: The case of $\Sigma=\Sigma_1$ and $N=50$

	$\tau=0.2$ $N_1=40, N_2=10$			$\tau=0.4$ $N_1=30, N_2=20$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	2.3552	6.7538^{-3}	6.7538^{-3}	2.3993	5.0516^{-3}	5.0516^{-3}
E_2	4.4861^{-4}	4.5769^{-5}	4.5769^{-5}	6.4901^{-4}	2.6455^{-4}	2.6455^{-4}
E_3	1.0782^{-2}	1.6567^{-4}	5.4178^{-5}	8.6972^{-3}	4.4607^{-6}	4.4607^{-6}
E	2.3665	6.9652^{-3}	6.8537^{-3}	2.4087	5.3206^{-3}	5.3206^{-3}
	$\tau=0.6$ $N_1=20, N_2=30$			$\tau=0.8$ $N_1=10, N_2=40$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	2.0713	5.3908^{-3}	5.3908^{-3}	2.3876	6.7461^{-3}	6.7461^{-3}
E_2	1.5612^{-3}	7.0194^{-4}	7.0194^{-4}	1.0291^{-3}	7.7547^{-4}	7.7547^{-4}
E_3	8.7614^{-5}	1.0631^{-2}	1.1884^{-4}	2.9159	3.3726	1.2930^{-1}
E	2.0730	1.6723^{-2}	6.2116^{-3}	5.3045	3.3801	1.3682^{-1}

Table 3: The case of $\Sigma=\Sigma_1$ and $N=100$

	$\tau=0.2$ $N_1=80, N_2=20$			$\tau=0.4$ $N_1=60, N_2=40$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	5.6580^{-1}	6.8098^{-4}	6.8098^{-4}	5.4891^{-1}	6.8554^{-4}	6.8554^{-4}
E_2	1.5406^{-4}	4.8578^{-5}	4.8578^{-5}	4.0253^{-4}	2.9986^{-4}	2.9986^{-4}
E_3	2.7944^{-3}	4.8758^{-5}	2.5336^{-6}	4.0823^{-3}	3.2906^{-4}	9.4919^{-6}
E	5.6875^{-1}	7.7832^{-4}	7.3210^{-4}	5.5340^{-1}	1.3145^{-3}	9.9488^{-4}
	$\tau=0.6$ $N_1=40, N_2=60$			$\tau=0.8$ $N_1=20, N_2=80$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	5.4923^{-1}	1.0290^{-3}	1.0290^{-3}	5.9817^{-1}	3.0913^{-3}	3.0913^{-3}
E_2	1.8299^{-4}	9.8078^{-5}	9.8078^{-5}	4.8292^{-4}	3.8359^{-4}	3.8359^{-4}
E_3	4.5719^{-3}	4.7673^{-4}	9.8867^{-6}	7.6517^{-3}	1.8194^{-2}	1.7465^{-4}
E	5.5399^{-1}	1.6038^{-3}	1.1369^{-3}	6.0630^{-1}	2.1669^{-2}	3.6495^{-3}

Table 4: The case of $\Sigma = \Sigma_1$ and $N = 200$

	$\tau = 0.2$, $N_1 = 160, N_2 = 40$			$\tau = 0.4$, $N_1 = 120, N_2 = 80$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	1.2097^{-1}	1.5502^{-3}	1.5502^{-3}	1.2247^{-1}	1.4501^{-3}	1.4501^{-3}
E_2	5.5458^{-5}	5.3448^{-5}	5.3448^{-5}	7.7793^{-5}	3.4600^{-5}	3.4600^{-5}
E_3	7.9178^{-4}	2.8016^{-5}	2.2151^{-6}	1.1935^{-3}	1.3518^{-4}	2.0943^{-6}
E	1.2182^{-1}	1.6316^{-3}	1.6058^{-3}	1.2374^{-1}	1.6199^{-3}	1.4868^{-3}
	$\tau = 0.6$, $N_1 = 80, N_2 = 120$			$\tau = 0.8$, $N_1 = 40, N_2 = 160$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	1.4658^{-1}	5.0190^{-4}	5.0190^{-4}	1.3462^{-1}	2.3597^{-4}	2.3597^{-4}
E_2	1.2721^{-4}	1.1830^{-4}	1.1830^{-4}	4.1906^{-4}	4.2582^{-4}	4.2582^{-4}
E_3	2.1185^{-3}	5.3706^{-4}	7.0709^{-6}	2.6179^{-3}	8.0144^{-4}	1.0151^{-5}
E	1.4883^{-1}	1.1573^{-3}	6.2726^{-4}	1.3766^{-1}	1.4632^{-3}	6.7194^{-4}

Table 5: The case of $\Sigma = \Sigma_1$ and $N = 500$

	$\tau = 0.2$, $N_1 = 400, N_2 = 100$			$\tau = 0.4$, $N_1 = 300, N_2 = 200$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	1.9044^{-2}	2.7188^{-4}	2.7188^{-4}	2.3433^{-2}	1.4041^{-4}	1.4041^{-4}
E_2	1.6410^{-5}	1.2222^{-5}	1.2222^{-5}	3.1681^{-5}	2.9937^{-5}	2.9937^{-5}
E_3	1.2724^{-4}	4.4094^{-6}	1.5043^{-7}	2.2540^{-4}	3.4036^{-5}	5.3136^{-7}
E	1.9188^{-2}	2.8852^{-4}	2.8426^{-4}	2.3690^{-2}	2.0438^{-4}	1.7088^{-4}
	$\tau = 0.6$, $N_1 = 200, N_2 = 300$			$\tau = 0.8$, $N_1 = 100, N_2 = 400$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	2.7359^{-2}	4.3671^{-4}	4.3671^{-4}	1.3704^{-2}	1.2472^{-3}	1.2472^{-3}
E_2	7.3274^{-5}	6.5046^{-5}	6.5046^{-5}	5.4416^{-5}	5.1057^{-5}	5.1057^{-5}
E_3	5.1529^{-4}	1.8331^{-4}	4.6796^{-6}	1.1903^{-3}	6.4227^{-4}	3.1489^{-6}
E	2.7947^{-2}	6.8506^{-4}	5.0643^{-4}	1.4949^{-2}	1.9405^{-3}	1.3014^{-3}

Table 6: The case of $\Sigma = \Sigma_1$ and $N = 1000$

	$\tau = 0.2$ $N_1 = 800, N_2 = 200$			$\tau = 0.4$ $N_1 = 600, N_2 = 400$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	4.6546^{-3}	5.5155^{-5}	5.5155^{-5}	4.6599^{-3}	1.6598^{-4}	1.6598^{-4}
E_2	1.2465^{-5}	1.1165^{-5}	1.1165^{-5}	1.6583^{-5}	1.7479^{-5}	1.7479^{-5}
E_3	2.1383^{-5}	8.6732^{-7}	2.2256^{-6}	6.0348^{-5}	1.0592^{-5}	9.8281^{-7}
E	4.6884^{-3}	6.7188^{-5}	6.8546^{-5}	4.7368^{-3}	1.9405^{-4}	1.8444^{-4}
	$\tau = 0.6$ $N_1 = 400, N_2 = 600$			$\tau = 0.8$ $N_1 = 200, N_2 = 800$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	5.1580^{-3}	2.3305^{-5}	2.3305^{-5}	5.6903^{-3}	5.4333^{-5}	5.4333^{-5}
E_2	1.6632^{-5}	1.6893^{-5}	1.6893^{-5}	8.8377^{-6}	1.1114^{-5}	1.1114^{-5}
E_3	1.1902^{-4}	3.9879^{-5}	1.8154^{-7}	4.0419^{-4}	2.4087^{-4}	1.2513^{-6}
E	5.2936^{-3}	8.0077^{-5}	4.0379^{-5}	6.1033^{-3}	3.0632^{-4}	6.6698^{-5}

Table 7: The case of $\Sigma = \Sigma_2$ and $N = 50$

	$\tau = 0.2$ $N_1 = 40, N_2 = 10$			$\tau = 0.4$ $N_1 = 30, N_2 = 20$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	3.0932^{-2}	4.2654^{-5}	4.2654^{-5}	3.2715^{-2}	4.6063^{-5}	4.6063^{-5}
E_2	2.4311^{-3}	2.6892^{-5}	2.6892^{-5}	2.8784^{-3}	3.1154^{-5}	3.1154^{-5}
E_3	2.5049^{-2}	8.4499^{-5}	2.7759^{-5}	2.3289^{-2}	6.4427^{-6}	6.4427^{-6}
E	5.8413^{-2}	1.5405^{-4}	9.7305^{-5}	5.8882^{-2}	8.3660^{-5}	8.3660^{-5}
	$\tau = 0.6$ $N_1 = 20, N_2 = 30$			$\tau = 0.8$ $N_1 = 10, N_2 = 40$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	3.0590^{-2}	7.0376^{-5}	7.0376^{-5}	3.0495^{-2}	5.6219^{-5}	5.6219^{-5}
E_2	2.5820^{-3}	2.7002^{-5}	2.7002^{-5}	1.9473^{-3}	2.6164^{-4}	2.6164^{-4}
E_3	2.9558^{-4}	2.0655^{-2}	1.6331^{-4}	6.4372	7.5186	3.0954^{-1}
E	3.3468^{-2}	2.0753^{-2}	2.6069^{-4}	6.4697	7.5190	3.0986^{-1}

Table 8: The case of $\Sigma = \Sigma_2$ and $N = 100$

	$\tau = 0.2$ $N_1 = 80, N_2 = 20$			$\tau = 0.4$ $N_1 = 60, N_2 = 40$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	8.7293^{-3}	4.4059^{-5}	4.4059^{-5}	8.9464^{-3}	6.7298^{-5}	6.7298^{-5}
E_2	9.3079^{-4}	4.2326^{-5}	4.2326^{-5}	9.5537^{-4}	3.7480^{-5}	3.7480^{-5}
E_3	7.8118^{-3}	1.7269^{-4}	5.5972^{-6}	1.0635^{-2}	7.8428^{-4}	3.5411^{-5}
E	1.7472^{-2}	2.5907^{-4}	9.1982^{-5}	2.0537^{-2}	8.8906^{-4}	1.4019^{-4}
	$\tau = 0.6$ $N_1 = 40, N_2 = 60$			$\tau = 0.8$ $N_1 = 20, N_2 = 80$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	8.1344^{-3}	3.0964^{-5}	3.0964^{-5}	7.7745^{-3}	5.1610^{-6}	5.1610^{-6}
E_2	6.4986^{-4}	2.4554^{-5}	2.4554^{-5}	5.8153^{-4}	2.2257^{-4}	2.2257^{-4}
E_3	1.1228^{-2}	9.4952^{-4}	1.1778^{-5}	1.5197^{-2}	3.9972^{-2}	3.7677^{-4}
E	2.0012^{-2}	1.0050^{-3}	6.7297^{-5}	2.3553^{-2}	4.0200^{-2}	6.0451^{-4}

Table 9: The case of $\Sigma = \Sigma_2$ and $N = 200$

	$\tau = 0.2$ $N_1 = 160, N_2 = 40$			$\tau = 0.4$ $N_1 = 120, N_2 = 80$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	1.9455^{-3}	1.9938^{-5}	1.9938^{-5}	1.9231^{-3}	9.4143^{-6}	9.4143^{-6}
E_2	2.4286^{-4}	1.0058^{-5}	1.0058^{-5}	1.5562^{-4}	1.0896^{-5}	1.0896^{-5}
E_3	2.1662^{-3}	7.9758^{-5}	6.5341^{-6}	2.5480^{-3}	1.8338^{-4}	3.5459^{-5}
E	4.3545^{-3}	1.0975^{-4}	3.6530^{-5}	4.6267^{-3}	2.0369^{-4}	5.5769^{-5}
	$\tau = 0.6$ $N_1 = 80, N_2 = 120$			$\tau = 0.8$ $N_1 = 40, N_2 = 160$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	1.9238^{-3}	6.3989^{-6}	6.3989^{-6}	1.9444^{-3}	9.1507^{-6}	9.1507^{-6}
E_2	3.5978^{-4}	5.3732^{-5}	5.3732^{-5}	1.4861^{-4}	4.8746^{-5}	4.8746^{-5}
E_3	5.4593^{-3}	1.3225^{-3}	1.3892^{-5}	6.9245^{-3}	2.1238^{-3}	1.0738^{-4}
E	7.7429^{-3}	1.3826^{-3}	7.4023^{-5}	9.0175^{-3}	2.1817^{-3}	1.6528^{-4}

Table 10: The case of $\Sigma = \Sigma_2$ and $N = 500$

	$\tau = 0.2$ $N_1 = 400, N_2 = 100$			$\tau = 0.4$ $N_1 = 300, N_2 = 200$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	3.2783^{-4}	3.7520^{-6}	3.7520^{-6}	2.6190^{-4}	5.4277^{-6}	5.4277^{-6}
E_2	2.2172^{-5}	4.4582^{-6}	4.4582^{-6}	4.4074^{-5}	5.8824^{-6}	5.8824^{-6}
E_3	3.0515^{-4}	6.5017^{-6}	1.1803^{-6}	5.9870^{-4}	8.8815^{-5}	2.6770^{-6}
E	6.5515^{-4}	1.4712^{-5}	9.3905^{-6}	9.0467^{-4}	1.0013^{-4}	1.3987^{-5}
	$\tau = 0.6$ $N_1 = 200, N_2 = 300$			$\tau = 0.8$ $N_1 = 100, N_2 = 400$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	3.0123^{-4}	2.4349^{-6}	2.4349^{-6}	3.5112^{-4}	3.0065^{-6}	3.0065^{-6}
E_2	3.8493^{-5}	8.8410^{-6}	8.8410^{-6}	4.0609^{-5}	2.1897^{-5}	2.1897^{-5}
E_3	9.5082^{-4}	2.5210^{-4}	4.4004^{-6}	3.2932^{-3}	1.8051^{-3}	3.3478^{-5}
E	1.2905^{-3}	2.6338^{-4}	1.5676^{-5}	3.6849^{-3}	1.8300^{-3}	5.8381^{-5}

Table 11: The case of $\Sigma = \Sigma_2$ and $N = 1000$

	$\tau = 0.2$ $N_1 = 800, N_2 = 200$			$\tau = 0.4$ $N_1 = 600, N_2 = 400$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	7.5113^{-5}	7.7355^{-7}	7.7355^{-7}	1.0676^{-4}	4.0446^{-6}	4.0446^{-6}
E_2	9.0140^{-6}	2.3944^{-6}	2.3944^{-6}	5.4728^{-6}	2.3510^{-6}	2.3510^{-6}
E_3	7.8483^{-5}	2.8415^{-6}	1.1896^{-6}	1.7051^{-4}	3.0781^{-5}	2.0496^{-6}
E	1.6261^{-4}	6.0094^{-6}	4.3575^{-6}	2.8275^{-4}	3.7176^{-5}	8.4452^{-6}
	$\tau = 0.6$ $N_1 = 400, N_2 = 600$			$\tau = 0.8$ $N_1 = 200, N_2 = 800$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	8.4669^{-5}	7.0901^{-7}	7.0901^{-7}	9.2591^{-5}	1.6870^{-6}	1.6870^{-6}
E_2	5.0390^{-6}	4.1228^{-6}	4.1228^{-6}	7.2248^{-6}	1.4301^{-5}	1.4301^{-5}
E_3	3.0369^{-4}	9.9279^{-5}	9.5260^{-7}	1.0045^{-3}	5.8618^{-4}	7.0857^{-6}
E	3.9340^{-4}	1.0411^{-4}	5.7844^{-6}	1.1043^{-3}	6.0217^{-4}	2.3073^{-5}

Table 12: The coefficients concering the unbiase estimator

	$N = 100$			
	$\tau_1 = 0.77$	$\tau_1 = 0.62$	$\tau_1 = 0.52$	$\tau_1 = 0.46$
b_1	-7.1964^{-4}	-1.2209^{-3}	-1.5431^{-3}	-1.7675^{-3}
b_2	-8.4203^{-3}	-1.7574^{-2}	-2.6588^{-2}	-3.3510^{-2}
c_0	-7.6174^{-6}	-1.2980^{-5}	-1.6453^{-5}	-1.8885^{-5}
c_1	1.0102	1.0103	1.0104	1.0105
c_2	7.6002^{-6}	1.2929^{-5}	1.6371^{-5}	1.8776^{-5}
c_3	-1.1440^{-6}	-2.9924^{-6}	-5.4469^{-6}	-7.8158^{-6}
	$N = 200$			
	$\tau_1 = 0.77$	$\tau_1 = 0.62$	$\tau_1 = 0.52$	$\tau_1 = 0.46$
b_1	-3.7162^{-4}	-6.5126^{-4}	-8.6474^{-4}	-9.9402^{-4}
b_2	-4.5380^{-3}	-9.2983^{-3}	-1.4210^{-2}	-1.9054^{-2}
c_0	-1.9105^{-6}	-3.3548^{-6}	-4.4614^{-6}	-5.1333^{-6}
c_1	1.0051	1.0051	1.0051	1.0051
c_2	1.9088^{-6}	3.3494^{-6}	4.4517^{-6}	5.1204^{-6}
c_3	-1.5073^{-7}	-3.8216^{-7}	-6.9837^{-7}	-1.0852^{-6}
	$N = 500$			
	$\tau_1 = 0.77$	$\tau_1 = 0.62$	$\tau_1 = 0.52$	$\tau_1 = 0.46$
b_1	-1.5910^{-4}	-2.7832^{-4}	-3.6176^{-4}	-4.2847^{-4}
b_2	-1.8620^{-3}	-3.8690^{-3}	-5.9270^{-3}	-8.0411^{-3}
c_0	-3.2174^{-7}	-5.6325^{-7}	-7.3252^{-7}	-8.6796^{-7}
c_1	1.0020	1.0020	1.0020	1.0020
c_2	3.2163^{-7}	5.6291^{-7}	7.3194^{-7}	8.6714^{-7}
c_3	-9.7614^{-9}	-2.5073^{-8}	-4.5647^{-8}	-7.1879^{-8}

Table 13: The case of $\Sigma = \Sigma_1$ and $N = 100$

	$\tau_1 = 0.77, \tau_2 = 0.15, \tau_3 = 0.08$ $N_1 = 77, N_2 = 15, N_3 = 8$			$\tau_1 = 0.62, \tau_2 = 0.25, \tau_3 = 0.13$ $N_1 = 62, N_2 = 25, N_3 = 13$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	8.5713	1.2795^{-2}	1.2795^{-2}	8.5902	8.5040^{-3}	8.5040^{-3}
E_2	6.6260^{-3}	9.8100^{-4}	9.8100^{-4}	7.4863^{-3}	3.4199^{-3}	3.4199^{-3}
E_3	2.4862^{-4}	2.5399^{-4}	2.5399^{-4}	4.6004^{-4}	4.6981^{-4}	4.6981^{-4}
E_4	1.5183^{-1}	6.1057^{-4}	5.9953^{-4}	1.5263^{-1}	8.1201^{-4}	7.9253^{-4}
E_5	1.4538^{-4}	8.0090^{-5}	8.0099^{-5}	1.8113^{-4}	9.5427^{-5}	9.5415^{-5}
E_6	2.4374^{-3}	1.2345^{-5}	9.8969^{-6}	3.1440^{-3}	1.0959^{-4}	9.3883^{-5}
E	8.7326	1.4733^{-2}	1.4719^{-2}	8.7541	1.3411^{-2}	1.3376^{-2}

	$\tau_1 = 0.52, \tau_2 = 0.32, \tau_3 = 0.16$ $N_1 = 52, N_2 = 32, N_3 = 16$			$\tau_1 = 0.46, \tau_2 = 0.36, \tau_3 = 0.18$ $N_1 = 46, N_2 = 36, N_3 = 18$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	9.1108	1.0779^{-2}	1.0779^{-2}	8.4984	9.4088^{-3}	9.4088^{-3}
E_2	7.4679^{-3}	2.2880^{-3}	2.2880^{-3}	5.1233^{-3}	1.9837^{-3}	1.9837^{-3}
E_3	7.2757^{-4}	7.4264^{-4}	7.4264^{-4}	1.3592^{-4}	1.3868^{-4}	1.3868^{-4}
E_4	1.7107^{-1}	2.1735^{-3}	2.1224^{-3}	1.8793^{-1}	4.7917^{-3}	4.7076^{-3}
E_5	1.6897^{-4}	5.3658^{-5}	5.3597^{-5}	1.1186^{-4}	8.2850^{-5}	8.2936^{-5}
E_6	2.4314^{-3}	1.3698^{-5}	7.2132^{-6}	1.4970^{-3}	6.6953^{-5}	9.3123^{-5}
E	9.2926	1.6050^{-2}	1.5993^{-2}	8.6932	1.6473^{-2}	1.6415^{-2}

Table 14: The case of $\Sigma = \Sigma_1$ and $N = 200$

	$\tau_1 = 0.77, \tau_2 = 0.15, \tau_3 = 0.08$ $N_1 = 154, N_2 = 31, N_3 = 15$			$\tau_1 = 0.62, \tau_2 = 0.25, \tau_3 = 0.13$ $N_1 = 125, N_2 = 50, N_3 = 25$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	1.9102	1.0243^{-2}	1.0243^{-2}	1.9483	1.7728^{-2}	1.7728^{-2}
E_2	2.6102^{-3}	7.4395^{-4}	7.4395^{-4}	4.6927^{-3}	2.1651^{-3}	2.1651^{-3}
E_3	1.0484^{-4}	1.0596^{-4}	1.0596^{-4}	3.3213^{-4}	3.3545^{-4}	3.3545^{-4}
E_4	4.1104^{-2}	3.7483^{-4}	3.7234^{-4}	4.3930^{-2}	6.9346^{-4}	6.8747^{-4}
E_5	5.8739^{-5}	2.5246^{-5}	2.5240^{-5}	3.8736^{-5}	2.9333^{-5}	2.9339^{-5}
E_6	7.7733^{-4}	2.6206^{-5}	2.5205^{-5}	9.7011^{-4}	6.9035^{-5}	6.5597^{-5}
E	1.9548	1.1519^{-2}	1.1515^{-2}	1.9983	2.1021^{-2}	2.1011^{-2}

	$\tau_1 = 0.52, \tau_2 = 0.32, \tau_3 = 0.16$ $N_1 = 105, N_2 = 63, N_3 = 32$			$\tau_1 = 0.46, \tau_2 = 0.36, \tau_3 = 0.18$ $N_1 = 91, N_2 = 73, N_3 = 36$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	2.3489	6.5720^{-3}	6.5720^{-3}	1.7915	2.2956^{-2}	2.2956^{-2}
E_2	3.9862^{-3}	2.8284^{-3}	2.8284^{-3}	2.0777^{-3}	4.7582^{-4}	4.7582^{-4}
E_3	5.0367^{-4}	5.0867^{-4}	5.0867^{-4}	3.2332^{-4}	3.2619^{-4}	3.2619^{-4}
E_4	4.1876^{-2}	5.7884^{-4}	5.7255^{-4}	5.1996^{-2}	2.1176^{-3}	2.1020^{-3}
E_5	2.8888^{-5}	1.7089^{-5}	1.7095^{-5}	4.8352^{-5}	3.2616^{-5}	3.2620^{-5}
E_6	1.1349^{-3}	1.1635^{-4}	1.0944^{-4}	1.1512^{-3}	1.2114^{-4}	1.1172^{-4}
E	2.3965	1.0621^{-2}	1.0608^{-2}	1.8471	2.6029^{-2}	2.6004^{-2}

Table 15: The case of $\Sigma = \Sigma_1$ and $N = 500$

	$\tau_1 = 0.77, \tau_2 = 0.15, \tau_3 = 0.08$ $N_1 = 385, N_2 = 77, N_3 = 38$			$\tau_1 = 0.62, \tau_2 = 0.25, \tau_3 = 0.13$ $N_1 = 312, N_2 = 125, N_3 = 63$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	3.6682^{-1}	1.3064^{-3}	1.3064^{-3}	4.2048^{-1}	6.4424^{-3}	6.4424^{-3}
E_2	7.2086^{-4}	2.7555^{-4}	2.7555^{-4}	1.6348^{-3}	1.0984^{-3}	1.0984^{-3}
E_3	3.6276^{-5}	3.6383^{-5}	3.6383^{-5}	2.1393^{-5}	2.1385^{-5}	2.1385^{-5}
E_4	8.5936^{-3}	3.7189^{-4}	3.7145^{-4}	6.7585^{-3}	8.6028^{-5}	8.5688^{-5}
E_5	1.4824^{-5}	1.1670^{-5}	1.1670^{-5}	2.1433^{-5}	1.8278^{-5}	1.8278^{-5}
E_6	1.4542^{-4}	8.3444^{-6}	8.2470^{-6}	1.8086^{-4}	1.8295^{-5}	1.7994^{-5}
E	3.7633^{-1}	2.0102^{-3}	2.0097^{-3}	4.2910^{-1}	7.6848^{-3}	7.6842^{-3}
	$\tau_1 = 0.52, \tau_2 = 0.32, \tau_3 = 0.16$ $N_1 = 263, N_2 = 158, N_3 = 79$			$\tau_1 = 0.46, \tau_2 = 0.36, \tau_3 = 0.18$ $N_1 = 227, N_2 = 182, N_3 = 91$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	3.4159^{-1}	2.0754^{-3}	2.0754^{-3}	3.1321^{-1}	1.6904^{-3}	1.6904^{-3}
E_2	5.7654^{-4}	4.2281^{-4}	4.2281^{-4}	7.3338^{-4}	4.5996^{-4}	4.5996^{-4}
E_3	6.6311^{-5}	6.6808^{-5}	6.6808^{-5}	1.7167^{-4}	1.7222^{-4}	1.7222^{-4}
E_4	8.4985^{-3}	3.6298^{-4}	3.6201^{-4}	7.5166^{-3}	1.8423^{-4}	1.8343^{-4}
E_5	2.5358^{-5}	1.9243^{-5}	1.9242^{-5}	2.7417^{-5}	3.5778^{-5}	3.5783^{-5}
E_6	2.3043^{-4}	3.6530^{-5}	3.5884^{-5}	3.6004^{-4}	9.7182^{-5}	9.5750^{-5}
E	3.5098^{-1}	2.9837^{-3}	2.9821^{-3}	3.2202^{-1}	2.6398^{-3}	2.6375^{-3}

Table 16: The case of $\Sigma = \Sigma_2$ and $N = 100$

	$\tau_1 = 0.77, \tau_2 = 0.15, \tau_3 = 0.08$ $N_1 = 77, N_2 = 15, N_3 = 8$			$\tau_1 = 0.62, \tau_2 = 0.25, \tau_3 = 0.13$ $N_1 = 62, N_2 = 25, N_3 = 13$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	5.6992^{-3}	4.3468^{-6}	4.3468^{-6}	5.5498^{-3}	2.5515^{-6}	2.5515^{-6}
E_2	8.0570^{-4}	8.0390^{-6}	8.0390^{-6}	7.5371^{-4}	6.9750^{-6}	6.9750^{-6}
E_3	2.0056^{-5}	3.3042^{-6}	3.3042^{-6}	3.8078^{-5}	1.5022^{-5}	1.5022^{-5}
E_4	6.7607^{-3}	5.9516^{-5}	5.8891^{-5}	7.6276^{-3}	1.4471^{-4}	1.4276^{-4}
E_5	7.1622^{-4}	4.0700^{-6}	4.0560^{-6}	1.1488^{-3}	7.4377^{-5}	7.3918^{-5}
E_6	6.9978^{-3}	7.2472^{-5}	6.2263^{-5}	7.8238^{-3}	1.7704^{-4}	1.4358^{-4}
E	2.1000^{-2}	1.5175^{-4}	1.4090^{-4}	2.2942^{-2}	4.2067^{-4}	3.8480^{-4}
	$\tau_1 = 0.52, \tau_2 = 0.32, \tau_3 = 0.16$ $N_1 = 52, N_2 = 32, N_3 = 16$			$\tau_1 = 0.46, \tau_2 = 0.36, \tau_3 = 0.18$ $N_1 = 46, N_2 = 36, N_3 = 18$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	5.5300^{-3}	5.7587^{-6}	5.7587^{-6}	6.2705^{-3}	1.9415^{-5}	1.9415^{-5}
E_2	6.5024^{-4}	7.1888^{-6}	7.1888^{-6}	9.1932^{-4}	7.2074^{-5}	7.2074^{-5}
E_3	2.5360^{-5}	4.5874^{-6}	4.5874^{-6}	6.9734^{-5}	3.7721^{-5}	3.7721^{-5}
E_4	7.3846^{-3}	1.1557^{-4}	1.1337^{-4}	6.4526^{-3}	2.8703^{-5}	2.7546^{-5}
E_5	7.8756^{-4}	1.0424^{-5}	1.0286^{-5}	8.5117^{-4}	2.5815^{-5}	2.5570^{-5}
E_6	5.7747^{-3}	2.4847^{-6}	4.6793^{-6}	4.7457^{-3}	9.0484^{-5}	1.3295^{-4}
E	2.0153^{-2}	1.4601^{-4}	1.4587^{-4}	1.9309^{-2}	2.7421^{-4}	3.1527^{-4}

Table 17: The case of $\Sigma = \Sigma_2$ and $N = 200$

	$\tau_1 = 0.77, \tau_2 = 0.15, \tau_3 = 0.08$ $N_1 = 154, N_2 = 31, N_3 = 15$			$\tau_1 = 0.62, \tau_2 = 0.25, \tau_3 = 0.13$ $N_1 = 125, N_2 = 50, N_3 = 25$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	1.3172^{-3}	3.0482^{-6}	3.0482^{-6}	1.5286^{-3}	3.9085^{-6}	3.9085^{-6}
E_2	1.1327^{-4}	9.6543^{-6}	9.6543^{-6}	1.5139^{-4}	4.8057^{-6}	4.8057^{-6}
E_3	1.0501^{-5}	5.7283^{-6}	5.7283^{-6}	1.2202^{-5}	4.0873^{-6}	4.0873^{-6}
E_4	1.6467^{-3}	1.5969^{-5}	1.5894^{-5}	1.5014^{-3}	8.8032^{-6}	8.7637^{-6}
E_5	2.1861^{-4}	7.8565^{-6}	7.8424^{-6}	1.4576^{-4}	1.1198^{-5}	1.1222^{-5}
E_6	1.8503^{-3}	3.7714^{-5}	3.5942^{-5}	2.4977^{-3}	1.5669^{-4}	1.4815^{-4}
E	5.1566^{-3}	7.9970^{-5}	7.8110^{-5}	5.8371^{-3}	1.8950^{-4}	1.8094^{-4}

	$\tau_1 = 0.52, \tau_2 = 0.32, \tau_3 = 0.16$ $N_1 = 105, N_2 = 63, N_3 = 32$			$\tau_1 = 0.46, \tau_2 = 0.36, \tau_3 = 0.18$ $N_1 = 91, N_2 = 73, N_3 = 36$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	1.3199^{-3}	5.1892^{-6}	5.1892^{-6}	1.2382^{-3}	1.0546^{-5}	1.0546^{-5}
E_2	2.0370^{-4}	3.8904^{-6}	3.8904^{-6}	1.0131^{-4}	1.5965^{-5}	1.5965^{-5}
E_3	1.6309^{-5}	1.3782^{-5}	1.3782^{-5}	1.3552^{-5}	1.8392^{-5}	1.8392^{-5}
E_4	1.8531^{-3}	3.2215^{-5}	3.1901^{-5}	2.0717^{-3}	7.4252^{-5}	7.3717^{-5}
E_5	1.7451^{-4}	1.7754^{-5}	1.7765^{-5}	2.7890^{-4}	2.3734^{-5}	2.3652^{-5}
E_6	2.6373^{-3}	1.9450^{-4}	1.8008^{-4}	3.2341^{-3}	3.8091^{-4}	3.5375^{-4}
E	6.2048^{-3}	2.6733^{-4}	2.5261^{-4}	6.9378^{-3}	5.2380^{-4}	4.9603^{-4}

Table 18: The case of $\Sigma = \Sigma_2$ and $N = 500$

	$\tau_1 = 0.77, \tau_2 = 0.15, \tau_3 = 0.08$ $N_1 = 385, N_2 = 77, N_3 = 38$			$\tau_1 = 0.62, \tau_2 = 0.25, \tau_3 = 0.13$ $N_1 = 312, N_2 = 125, N_3 = 63$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	2.5263^{-4}	5.3491^{-6}	5.3491^{-6}	2.5447^{-4}	1.4414^{-6}	1.4414^{-6}
E_2	4.5003^{-5}	4.4900^{-6}	4.4900^{-6}	4.2039^{-5}	2.3561^{-6}	2.3561^{-6}
E_3	1.7122^{-6}	1.3157^{-6}	1.3157^{-6}	1.7782^{-5}	1.3895^{-5}	1.3895^{-5}
E_4	2.3169^{-4}	3.2733^{-6}	3.2741^{-6}	3.3219^{-4}	1.0166^{-5}	1.0143^{-5}
E_5	1.9808^{-5}	2.1733^{-6}	2.1748^{-6}	3.5683^{-5}	3.9415^{-6}	3.9404^{-6}
E_6	3.1839^{-4}	9.3630^{-6}	9.2111^{-6}	4.4177^{-4}	4.5398^{-5}	4.4745^{-5}
E	8.6925^{-4}	2.5964^{-5}	2.5815^{-5}	1.1239^{-3}	7.7197^{-5}	7.6520^{-5}

	$\tau_1 = 0.52, \tau_2 = 0.32, \tau_3 = 0.16$ $N_1 = 263, N_2 = 158, N_3 = 79$			$\tau_1 = 0.46, \tau_2 = 0.36, \tau_3 = 0.18$ $N_1 = 227, N_2 = 182, N_3 = 91$		
	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$	$\hat{\Sigma}$	$\tilde{\Sigma}$	$\tilde{\tilde{\Sigma}}$
E_1	2.4498^{-4}	2.0588^{-6}	2.0588^{-6}	2.2725^{-4}	3.4642^{-6}	3.4642^{-6}
E_2	2.7179^{-5}	7.2557^{-6}	7.2557^{-6}	3.5448^{-5}	2.7797^{-6}	2.7797^{-6}
E_3	3.8775^{-6}	4.8747^{-6}	4.8747^{-6}	5.3194^{-6}	4.9778^{-6}	4.9778^{-6}
E_4	3.1407^{-4}	9.0516^{-6}	9.0267^{-6}	2.7605^{-4}	3.3394^{-6}	3.3218^{-6}
E_5	3.6256^{-5}	2.5382^{-6}	2.5361^{-6}	3.3265^{-5}	3.2579^{-6}	3.2569^{-6}
E_6	5.7703^{-4}	8.1516^{-5}	7.9921^{-5}	6.5251^{-4}	1.1270^{-4}	1.1019^{-4}
E	1.2034^{-3}	1.0730^{-4}	1.0567^{-4}	1.2298^{-3}	1.3052^{-4}	1.2799^{-4}

References

[1] Anderson, T. W. (1957). Maximum likelihood estimators for a multivariate normal distribution when some observations are missing, *J. Amer. Statist. Assoc.*, **52**, 200–203.

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

[3] Chang, W. -Y. and Richards, D. St. P. (2009). Finite-sample inference with monotone incomplete multivariate normal data, I, *J. Multi. Anal.*, **100**, 1883–1899.

[4] Chang, W. -Y. and Richards, D. St. P. (2010). Finite-sample inference with monotone incomplete multivariate normal data, II, *J. Multi. Anal.*, **101**, 603–620.

[5] Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977).Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Statist. Soc.*, Ser. B, **39**, 1–38.

[6] Fujisawa, H. (1995). A note on the maximum likelihood estimators for multivariate normal distribution with monotone data, *Commun. Statist. Theory Meth.*, **24**, 1377–1382.

[7] Hao, J. and Krishnamoorthy, K. (2001). Inferences on a normal covariance matrix and generalized variance with monotone missing data, *J. Multi. Anal.*, **78**, 62–82.

[8] Kanda, T. and Fujikoshi, Y. (1998). Some basic properties of the MLE's for a multivariate normal distribution with monotone missing data, *Amer. J. Math. Statist.*, **18**, 161–190.

[9] Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data, 2nd ed.*, Wiley-Interscience.

[10] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*, Wiley, New York.

[11] Provost, S. B. (1991). Inference about covariances under missing values, *Environmetrics*, **2**, 487-496.

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

[13] Srivastava, M. S. (1985). Multivariate data with missing observations, *Commun. Statist. Theory Meth.*, **14**, 775–792.