Asymptotic distribution for latent root of covariance matrix under two-step monotone incomplete data

Shin-ichi, Tsukada
Meisei University, School of Education
2-1-1 Hodokubo
Hino Tokyo, 191-8506, Japan
E-mail: tsukada@ge.meisei-u.ac.jp

Tomoya, Yamada
Sapporo Gakuin University, Faculty of Economics
11 Bunkyodai
Ebetsu Hokkaido, 069-8555, Japan
E-mail: tomoya@sgu.ac.jp

1 Introduction

In multivariate analysis, there might be some incomplete sample. In this case, they cannot be analyzed by an existing method and we might analyze the complete data except incomplete sample. Because we throw information included incomplete sample in such analysis, we want to use that effectively. For analyzing these data, various statistical methods have been developed by Anderson and Olkin [1], Dempster, Laird and Rubin [4], Srivastava [10] and Little and Rubin [6].

In this paper, we treat an asymptotic distribution for a latent root of a covariance matrix under two-step monotone incomplete sample. Let $\boldsymbol{x}$ be distributed as $N_{p+q}(\boldsymbol{\mu}, \Sigma)$, $\boldsymbol{x}^{(1)}$ be the vector of the first p elements of $\boldsymbol{x}$ and $\boldsymbol{x}^{(2)}$ be the vector of the residual q elements of $\boldsymbol{x}$. Suppose we have N observations on $\boldsymbol{x}^{(1)}$, n observations on $\boldsymbol{x}^{(2)}$. Therefore we have the sample

$$\begin{pmatrix} \boldsymbol{x}_1^{(1)} \\ \boldsymbol{x}_1^{(2)} \end{pmatrix}, \begin{pmatrix} \boldsymbol{x}_2^{(1)} \\ \boldsymbol{x}_2^{(2)} \end{pmatrix}, \cdots \begin{pmatrix} \boldsymbol{x}_n^{(1)} \\ \boldsymbol{x}_n^{(2)} \end{pmatrix}, \begin{pmatrix} \boldsymbol{x}_{n+1}^{(1)} \\ * \end{pmatrix}, \cdots \begin{pmatrix} \boldsymbol{x}_N^{(1)} \\ * \end{pmatrix},$$

where the symbol $*$ denotes the missing sample.

Anderson and Olkin [1] consider a two-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. Kanda and Fujikoshi [5] investigate some fundamental properties of the MLE, and indicate that the MLE $\hat{\boldsymbol{\mu}}$ is unbiased but the MLE $\hat{\Sigma}$ 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 [2] derive a stochastic representation for the exact distribution of the MLE $\hat{\boldsymbol{\mu}}$ under two-step monotone sample. They obtain ellipsoidal confidence region for $\boldsymbol{\mu}$. Chang and Richards [3] deal with tests for the mean vector $\boldsymbol{\mu}$ and the covariance matrix Σ.

For principal component analysis, we consider the asymptotic distribution for the α-th largest latent root of the covariance matrix under two-step monotone incomplete sample. We represent the estimator for the covariance matrix which can be unitedly treated in Section 2. In Section 3, the asymptotic distribution for the α-th largest latent root is derived. The accuracy is investigated by numerical simulation in Section 4. Section 5 concludes the paper.

2 United estimator $\dot{S}$ for covariance matrix

For covariance matrix, there are some estimators which are the MLE, proposed by Kanda and Fujikoshi [5] and an unbiased estimator. Since we want to treat them unitedly, the united estimator $\dot{S}$ is composed in this section. Let $\bar{\boldsymbol{x}}_1^{(1)} = \sum_{\alpha=1}^{n} \boldsymbol{x}_\alpha^{(1)}/n$, $\bar{\boldsymbol{x}}_2^{(1)} = \sum_{\alpha=n+1}^{N} \boldsymbol{x}_\alpha^{(1)}/(N-n)$, $\bar{\boldsymbol{x}}^{(1)} = \sum_{\alpha=1}^{N} \boldsymbol{x}_\alpha^{(1)}/N$, $\bar{\boldsymbol{x}}^{(2)} = \sum_{\alpha=1}^{n} \boldsymbol{x}_\alpha^{(2)}/n$, $B = B_1 + B_2$,

$$
\begin{aligned}
A_{11,n} &= \sum_{\alpha=1}^{n} \left(\boldsymbol{x}_\alpha^{(1)} - \bar{\boldsymbol{x}}_1^{(1)}\right)\left(\boldsymbol{x}_\alpha^{(1)} - \bar{\boldsymbol{x}}_1^{(1)}\right)', \quad A_{12} = \sum_{\alpha=1}^{n} \left(\boldsymbol{x}_\alpha^{(1)} - \bar{\boldsymbol{x}}_1^{(1)}\right)\left(\boldsymbol{x}_\alpha^{(2)} - \bar{\boldsymbol{x}}^{(2)}\right)', \\
A_{21} &= \sum_{\alpha=1}^{n} \left(\boldsymbol{x}_\alpha^{(2)} - \bar{\boldsymbol{x}}^{(2)}\right)\left(\boldsymbol{x}_\alpha^{(1)} - \bar{\boldsymbol{x}}_1^{(1)}\right)', \quad A_{22} = \sum_{\alpha=1}^{n} \left(\boldsymbol{x}_\alpha^{(2)} - \bar{\boldsymbol{x}}^{(2)}\right)\left(\boldsymbol{x}_\alpha^{(2)} - \bar{\boldsymbol{x}}^{(2)}\right)', \\
B_1 &= \sum_{\alpha=n+1}^{N} \left(\boldsymbol{x}_\alpha^{(1)} - \bar{\boldsymbol{x}}_2^{(1)}\right)\left(\boldsymbol{x}_\alpha^{(1)} - \bar{\boldsymbol{x}}_2^{(1)}\right)', \quad B_2 = \frac{n(N-n)}{N}\left(\bar{\boldsymbol{x}}_1^{(1)} - \bar{\boldsymbol{x}}_2^{(1)}\right)\left(\bar{\boldsymbol{x}}_1^{(1)} - \bar{\boldsymbol{x}}_2^{(1)}\right)'.
\end{aligned}
$$

We could rewrite $A_{11,N} = \sum_{\alpha=1}^{N} \left(\boldsymbol{x}_\alpha^{(1)} - \bar{\boldsymbol{x}}^{(1)}\right)\left(\boldsymbol{x}_\alpha^{(1)} - \bar{\boldsymbol{x}}^{(1)}\right)'$ as follows:

$$
A_{11,N} = A_{11,n} + B_1 + B_2 = A_{11,n} + B.
$$

We could write estimators for the covariance matrix as follows:

$$
\begin{aligned}
\dot{S}_{11} &= \frac{1}{N-\delta} A_{11,N} = \frac{1}{N-\delta}\left(A_{11,n} + B\right), \\
\dot{S}_{12} &= \frac{1}{N-\delta} A_{11,N} A_{11,n}^{-1} A_{12} = \frac{1}{N-\delta}\left(A_{12} + B A_{11,n}^{-1} A_{12}\right), \\
\dot{S}_{22} &= \left(\frac{N}{N-\delta} - c_0\right)\frac{1}{n}\left(A_{22} - A_{21} A_{11,n}^{-1} A_{12}\right) \\
&\quad + \frac{1}{N-\delta} A_{21} A_{11,n}^{-1} A_{12} + \frac{1}{N-\delta} A_{21} A_{11,n}^{-1} B A_{11,n}^{-1} A_{12}.
\end{aligned}
$$

When $\delta = 0$ and $c_0 = 0$, $\dot{S}_{11}$, $\dot{S}_{12}$ and $\dot{S}_{22}$ are the maximum likelihood estimator (MLE) (See Chang and Richards [2].). They are the estimator (KFE) proposed by Kanda and Fujikoshi [5] when $\delta = 1$ and $c_0 = 0$, and they are the unbiased estimator (UBE) by Tsukada [11] when $\delta = 1$ and $c_0 = (N-n)\{(p+1)(p+2) - n\}/\{(N-1)(n-p-2)(n-p-1)\}$.

3 Asymptotic distribution for latent root of covariance matrix

Suppose that the population covariance matrix represents

$$
\Sigma = \Gamma \Lambda \Gamma',
$$

where Γ is an orthogonal matrix and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_{p+q})$ with $\lambda_1 > \cdots > \lambda_{p+q}$. The α-th largest latent root $\dot{l}_\alpha$ of $\dot{S}$ is expanded as follows:

$$
\dot{l}_\alpha = \lambda_\alpha + \dot{c}_{\alpha\alpha} + O_p(N^{-1}),
$$

where $\dot{C} = (\dot{c}_{ij}) = \Gamma'\left(\dot{S} - \Sigma\right)\Gamma$. We have to calculate the moments of $\dot{c}_{\alpha\alpha}$ to obtain the asymptotic distribution. From Proposition 4.1 by Chang and Richards [2],

$$
\begin{pmatrix} A_{11,n} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \sim W_{p+q}(n-1, \Sigma), \qquad B \sim W_p(N-n, \Sigma_{11})
$$

and they are mutually independent. Letting

$$
\begin{cases}
A_{11,n} = (n-1)\Sigma_{11} + \sqrt{n-1}\, U_{11,n}, \\
A_{12} = (n-1)\Sigma_{12} + \sqrt{n-1}\, U_{12}, \\
A_{22} = (n-1)\Sigma_{22} + \sqrt{n-1}\, U_{22},
\end{cases}
$$

and $B = (N-n)\Sigma_{11} + \sqrt{N-n}V$, we have the expansion of $c_{\alpha\alpha}$ as follows:

$$\begin{aligned} c_{\alpha\alpha} &= \frac{N}{(N-\delta)\sqrt{n-1}}\boldsymbol{\gamma}'_\alpha U \boldsymbol{\gamma}_\alpha - \frac{N-n}{(N-\delta)\sqrt{n-1}}\boldsymbol{\beta}'_\alpha(\mathbf{1}_2 \otimes U_{11,n})\boldsymbol{\beta}_\alpha \\ &\quad + \frac{\sqrt{N-n}}{N-\delta}\boldsymbol{\beta}'_\alpha(\mathbf{1}_2 \otimes V)\boldsymbol{\beta}_\alpha + O_p((N-\delta)^{-1}), \end{aligned}$$

where $\boldsymbol{\gamma}_\alpha$ is the α-th column of Γ and separated into p components and q components as $(\boldsymbol{\gamma}'_{1\alpha}, \boldsymbol{\gamma}'_{2\alpha})'$,

$$U = \begin{pmatrix} U_{11,n} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}, \quad \Psi = \begin{pmatrix} I_p & O \\ O & \Sigma_{21}\Sigma_{11}^{-1} \end{pmatrix}, \quad \mathbf{1}_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\beta}_\alpha = \Psi \boldsymbol{\gamma}_\alpha.$$

Since $A = (n-1)\Sigma + \sqrt{n-1}\,U \sim W_{p+q}(n-1, \Sigma)$ and $B = (N-n)\Sigma_{11} + \sqrt{N-n}\,V \sim W_p(N-n, \Sigma_{11})$, we have

(1) $$E[U] = O_{(p+q)\times(p+q)}, \qquad E[U_{11}] = E[V] = O_{p\times p}$$

and the expectation of the latent root $\dot{l}_\alpha$ is

(2) $$E[\dot{l}_\alpha] = \lambda_\alpha + O(N^{-1}).$$

Next we calculate an asymptotic variance for $\dot{l}_\alpha$. Since the expansion of $c^2_{\alpha\alpha}$ is

$$\begin{aligned} c^2_{\alpha\alpha} &= \frac{N^2}{(N-\delta)^2(n-1)}(\boldsymbol{\gamma}'_\alpha U \boldsymbol{\gamma}_\alpha)^2 + \frac{(N-n)^2}{(N-\delta)^2(n-1)}\{\boldsymbol{\beta}'_\alpha(\mathbf{1}_2 \otimes U_{11,n})\boldsymbol{\beta}_\alpha\}^2 \\ &\quad + \frac{N-n}{(N-\delta)^2}\{\boldsymbol{\beta}'_\alpha(\mathbf{1}_2 \otimes V)\boldsymbol{\beta}_\alpha\}^2 - \frac{2N(N-n)}{(N-\delta)^2(n-1)}(\boldsymbol{\gamma}'_\alpha U \boldsymbol{\gamma}_\alpha)\{\boldsymbol{\beta}'_\alpha(\mathbf{1}_2 \otimes U_{11,n})\boldsymbol{\beta}_\alpha\} \\ &\quad + \frac{2N\sqrt{N-n}}{(N-\delta)^2\sqrt{n-1}}(\boldsymbol{\gamma}'_\alpha U \boldsymbol{\gamma}_\alpha)\{\boldsymbol{\beta}'_\alpha(\mathbf{1}_2 \otimes V)\boldsymbol{\beta}_\alpha\} \\ &\quad - \frac{2(N-n)^{3/2}}{(N-\delta)^2\sqrt{n-1}}\{\boldsymbol{\beta}'_\alpha(\mathbf{1}_2 \otimes U_{11,n})\boldsymbol{\beta}_\alpha\}\{\boldsymbol{\beta}'_\alpha(\mathbf{1}_2 \otimes V)\boldsymbol{\beta}_\alpha\} + O_p((N-\delta)^{-3/2}), \end{aligned}$$

we need $E[(\boldsymbol{\gamma}'_\alpha U \boldsymbol{\gamma}_\alpha)^2]$, $E[\{\boldsymbol{\beta}'_\alpha(\mathbf{1}_2 \otimes U_{11,n})\boldsymbol{\beta}_\alpha\}^2]$, $E[\{\boldsymbol{\beta}'_\alpha(\mathbf{1}_2 \otimes V)\boldsymbol{\beta}_\alpha\}^2]$ and $E[(\boldsymbol{\gamma}'_\alpha U \boldsymbol{\gamma}_\alpha)\{\boldsymbol{\beta}'_\alpha(\mathbf{1}_2 \otimes U_{11,n})\boldsymbol{\beta}_\alpha\}]$. To calculate these expectations, we use the following lemma.

Lemma 3.1 *Suppose that A has $W_p(n, \Sigma)$ and M is a constant matrix,*

$$E[AMA] = p^2\Sigma M \Sigma + p\Sigma M'\Sigma + ptr(M\Sigma)\Sigma.$$

Therefore we obtain the following theorem.

Theorem 3.1 *Suppose the latent root λ_α of the covariance matrix Σ is distinct. The asymptotic distribution for the sample latent root $\sqrt{N-\delta}\,(\dot{l}_\alpha - \lambda_\alpha)$ of the sample covariance matrix $\dot{S}$ is the normal distribution with the mean 0 and the variance ϕ^2_E, where*

$$\phi^2_E = \frac{N^2}{(N-\delta)(n-1)}(2\lambda^2_\alpha) - \frac{(N-n)(N+1)}{(N-\delta)(n-1)}\left\{2(\lambda_\alpha - \boldsymbol{\gamma}'_{2\alpha}\Sigma_{22\cdot1}\boldsymbol{\gamma}_{2\alpha})^2\right\}.$$

Collorary 3.1 *When the sample size N is sufficiently large, the sample latent root $\sqrt{N-\delta}\,(\dot{l}_\alpha - \lambda_\alpha)$ is asymptotically distributed as the normal distribution with the mean 0 and the variance ϕ^2_A, where $\tau = (N-n)/N$ and*

$$\phi^2_A = \frac{1}{1-\tau}(2\lambda^2_\alpha) - \frac{\tau}{1-\tau}\left\{2(\lambda_\alpha - \boldsymbol{\gamma}'_{2\alpha}\Sigma_{22\cdot1}\boldsymbol{\gamma}_{2\alpha})^2\right\}.$$

Remark 1. When there is no missing sample, that is, $\tau = 0$, the asymptotic variance of the sample latent root is $2\lambda^2_\alpha$. This is the case of the complete sample and the asymptotic variance is corresponding to the result in complete sample.

Remark 2. The asymptotic distribution for the latent root using MLE and those for the latent root using KFE and UBE are different, but it is the same as those using three estimator when the sample size N is sufficiently large.

4 Numerical simulation

In this section we investigate the accuracy for the asymptotic distribution by numerical simulation. We treat the result for Collorary and assume that the population distribution is the normal distribution $N_8(\mathbf{0}, \Sigma)$, where

$$\Lambda = \mathrm{diag}(128, 64, 32, 16, 8, 4, 2, 1),$$

$$\Gamma = \begin{pmatrix} 0.647 & -0.719 & -0.217 & 0.007 & 0.082 & 0.084 & 0.057 & 0.026 \\ 0.235 & 0.471 & -0.802 & -0.247 & 0.003 & 0.087 & 0.088 & 0.057 \\ 0.251 & 0.080 & 0.393 & -0.834 & -0.256 & 0.003 & 0.087 & 0.084 \\ 0.267 & 0.118 & 0.011 & 0.363 & -0.843 & -0.256 & 0.003 & 0.082 \\ 0.284 & 0.161 & 0.065 & -0.009 & 0.363 & -0.834 & -0.247 & 0.007 \\ 0.302 & 0.209 & 0.130 & 0.065 & 0.011 & 0.393 & -0.802 & -0.217 \\ 0.321 & 0.262 & 0.209 & 0.161 & 0.118 & 0.080 & 0.471 & -0.719 \\ 0.342 & 0.321 & 0.302 & 0.284 & 0.267 & 0.251 & 0.235 & 0.647 \end{pmatrix},$$

$$\Sigma = \Gamma\Lambda\Gamma' = \begin{pmatrix} 88.24 & 3.413 & 14.12 & 16.00 & 15.59 & 14.56 & 13.27 & 11.78 \\ 3.413 & 42.90 & 3.194 & 9.774 & 11.45 & 11.78 & 11.67 & 11.30 \\ 14.12 & 3.194 & 25.07 & 6.175 & 10.05 & 11.35 & 11.90 & 12.16 \\ 16.00 & 9.774 & 6.175 & 18.06 & 9.278 & 11.80 & 13.02 & 13.84 \\ 15.59 & 11.45 & 10.05 & 9.278 & 16.06 & 12.49 & 14.62 & 16.14 \\ 14.56 & 11.78 & 11.35 & 11.80 & 12.49 & 17.02 & 16.49 & 18.96 \\ 13.27 & 11.67 & 11.90 & 13.01 & 14.62 & 16.49 & 20.52 & 22.29 \\ 11.78 & 11.30 & 12.16 & 13.85 & 16.14 & 18.96 & 22.29 & 27.14 \end{pmatrix}.$$

The number of simulation is a thousand hundred and $p = q = 4$. We establish $n = 200$ and change the missing ratio τ, i.e., the number of missing sample $N - n$. Table 1 denotes the result in the case of $\tau = 10.3\%$. The value in parentheses is a relative error, the underlined value denotes the estimator that the relative error is the smallest, and the notation x^y denotes the value $x \times 10^y$. The expectation of the latent root using MLE is closest to the population latent root when there is a positive bias. When there is a negative bias oppositely, the expectation of the latent root using UBE is closest. As for the variance, that of latent root using MLE is small though those of latent root using KFE and UBE are almost the same values and their relative errors are small. The result in the case of $\tau = 25.1\%$ represents Table 2. There is a similar tendency as Table 1. The relative error has decreased compared with Table 1 because total sample size N increase.

For the expectation and the variance, the value becomes small from the largest latent root to the smallest latent root but the relative error grows.

5 Conclusion

We obtain the asymptotic distribution for the latent root of the covariance matrix under two-step monotone incomplete sample. The validity of the result is shown by numerical simulation.

We find that the estimator using MLE is good for the large latent roots and the asymptotic variance of the estimator using MLE is smallest. Since the estimator of the latent root has a bias for the complete sample, we will correct the asymptotic bias in next step.

And it is found from Collorary that the asymptotic variance becomes small for the complete sample by using the missing sample.

As an application, we could asymptotically test the hypothesis that the population latent root is equivalent to a constant, i.e., $H_0 : \lambda_\alpha = \lambda_0(\text{constant})$, using that the criterion $\sqrt{N-\delta}\ (l_\alpha - \lambda_0)/\phi_A$ is asymptotically distributed as the standard normal distribution.

$N = 223,\ n = 200,\ N - n = 23,\ \tau = 10.3\%$							
	Expectation			Variance			
α	MLE	KFE	UBE	MLE	KFE	UBE	TRUE
1	128.4 (0.287%)	128.9 (0.739%)	129.0 (0.742%)	148.1 (-1.059%)	149.4 (-0.165%)	149.4 (-0.164%)	149.7
2	63.60 (-0.624%)	63.89 (-0.176%)	63.89 (-0.171%)	36.56 (-2.719%)	36.89 (-1.840%)	36.89 (-1.839%)	37.58
3	31.60 (-1.247%)	31.74 (-0.802%)	31.75 (-0.796%)	9.055 (-4.088%)	9.137 (-3.222%)	9.137 (-3.220%)	9.441
4	15.71 (-1.801%)	15.78 (-1.359%)	15.78 (-1.351%)	2.270 (-4.294%)	2.291 (-3.430%)	2.291 (-3.427%)	2.372
5	7.811 (-2.365%)	7.846 (-1.925%)	7.848 (-1.900%)	0.5951 (-5.053%)	0.6004 (-4.196%)	0.6006 (-4.172%)	0.6267
6	3.874 (-3.138%)	3.892 (-2.701%)	3.894 (-2.657%)	0.1494 (-6.614%)	0.1508 (-5.771%)	0.1509 (-5.689%)	0.1600
7	1.924 (-3.784%)	1.933 (-3.351%)	1.934 (-3.306%)	3.763^{-2} (-5.935%)	3.797^{-2} (-5.086%)	3.800^{-2} (-4.998%)	4.000^{-2}
8	0.9517 (-4.831%)	0.9560 (-4.403%)	0.9564 (-4.358%)	9.367^{-3} (-6.329%)	9.452^{-3} (-5.484%)	9.460^{-3} (-5.396%)	1.000^{-2}

Table 1: Expectaion and Variance (Collorary)

$N = 267,\ N - n = 200,\ n = 67,\ \tau = 25.1\%$							
	Expectation			Variance			
	MLE	KFE	UBE	MLE	KFE	UBE	TRUE
1	128.3 (0.264%)	128.8 (0.641%)	128.8 (0.650%)	127.9 (-1.123%)	128.9 (-0.379%)	128.9 (-0.374%)	129.4
2	63.66 (-0.533%)	63.90 (-0.159%)	63.91 (-0.147%)	32.06 (-2.083%)	32.30 (-1.345%)	32.31 (-1.341%)	32.74
3	31.66 (-1.049%)	31.78 (-0.677%)	31.79 (-0.663%)	8.029 (-3.224%)	8.090 (-2.495%)	8.090 (-2.490%)	8.297
4	15.76 (-1.507%)	15.82 (-1.137%)	15.82 (-1.120%)	2.030 (-3.484%)	2.045 (-2.757%)	2.045 (-2.751%)	2.103
5	7.828 (-2.149%)	7.858 (-1.781%)	7.862 (-1.720%)	0.5809 (-4.402%)	0.5853 (-3.682%)	0.5856 (-3.631%)	0.6077
6	3.875 (-3.121%)	3.890 (-2.757%)	3.894 (-2.651%)	0.1495 (-6.553%)	0.1506 (-5.849%)	0.1509 (-5.650%)	0.1600
7	1.924 (-3.784%)	1.932 (-3.422%)	1.934 (-3.313%)	3.763^{-2} (-5.932%)	3.791^{-2} (-5.223%)	3.800^{-2} (-5.010%)	4.000^{-2}
8	0.9517 (-4.831%)	0.9553 (-4.473%)	0.9563 (-4.366%)	9.367^{-3} (-6.328%)	9.438^{-3} (-5.622%)	9.459^{-3} (-5.410%)	1.000^{-2}

Table 2: Expectation and Variance (Collorary)

REFERENCES (RÉFERENCES)

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

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

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

[4] Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977).Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society*, Ser. B, **39**, 1–38.

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

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

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

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

[9] Sugiura, N. (1973). Derivatives of the characteristic root of a symmetric or a Hermitian matrix with two applications in multivariate analysis. *Comm. Statist.* **1**, 393–417.

[10] Srivastava, M. S. (1985). Multivariate data with missing observations, *communications in statistics*, theory and methods, **14**, 775–792.

[11] Tsukada, S. (2010). Unbiased estimators for the covariance matrix under a monotone incomplete sample, *Research bulletin of Meisei University, Education*, **1**, 63–83.