Applied Mathematical Sciences, Vol. 3, 2009, no. 54, 2695 - 2702

Testing a Normal Covariance Matrix for Small Samples with Monotone Missing Data

Evelina Veleva

Rousse University "A. Kanchev"
Department of Numerical Methods and Statistics
Bulgaria, 7017 Rousse, 8 Studentska str., room 1.424
eveleva@ru.acad.bg

Abstract

We consider samples with monotone missing data, drawn from a normal population to test if the covariance matrix is equal to a given positive definite matrix. We propose an imputation procedure for the missing data and give the exact distribution of the corresponding likelihood ratio test statistic from the classical complete case.

Mathematics Subject Classification: 62H10, 62H15

Keywords: monotone missing data, Bellman gamma distribution, covariance matrix, hypotheses testing

1 Introduction

The problem of missing data is an important applied problem, because missing values are encountered in many practical situations (see [4]). Two commonly used approaches to analyze incomplete data are the likelihood based approach and the multiple imputation. The imputation method is to impute the missing data to form complete data and then use the standard methods available for the complete data analysis. For a good exposition of the imputation procedures and validity of imputation inferences in practice, we refer to [6].

In this paper we consider samples with monotone missing data pattern. Let $(X_1, \dots, X_n)^t$ be a random vector with multivariate normal distribution $N_n(\mu, \Sigma)$,where the mean vector μ and the covariance matrix Σ are unknown. Suppose that we have $k_1 + \cdots + k_n$ independent observations, k_1 of which are on $(X_1, \dots, X_n)^t$, k_2 - on $(X_2, \dots, X_n)^t$ and so on, k_n on the random variable X_n. Assume that $k_j \geq 0$ and $m_j = k_1 + \cdots + k_j > j$, $j = 1, \dots, n$. The data can be written in the following pattern, known as a monotone pattern

$$\begin{array}{ccccccccc}
x_{1,1} & \cdots & x_{1,m_1} & & & & & & \\
\vdots & & \vdots & & & & & & \\
x_{n-1,1} & \cdots & x_{n-1,m_1} & x_{n-1,m_1+1} & \cdots & x_{n-1,m_{n-1}} & & & \\
x_{n,1} & \cdots & x_{n,m_1} & x_{n,m_1+1} & \cdots & x_{n,m_{n-1}} & \cdots & x_{n,m_n} &
\end{array} . \tag{1}$$

In the literature on inference for μ and Σ, it is noticeable that the exact distributions of $\hat{\mu}$ and $\hat{\Sigma}$, the maximum likelihood estimators of μ and Σ, have remained unknown. This problem is basic to inference with incomplete data when large samples are infeasible or impractical (see [1]). In [1] the authors initiate a program of research on inference for μ and Σ with the goal of deriving explicit results analogous to those existing in the classical complete case.

In this paper we consider hypotheses $H_0 : \Sigma = \Sigma_0$ against $H_a : \Sigma \neq \Sigma_0$, propose an imputation procedure for the missing data in (1) and give the exact distribution of the corresponding likelihood ratio test statistic from the classical complete case.

2 Preliminary Notes

We shall use the following known propositions, which can be found in [5].

Proposition 2.1 *Let the $n \times 1$ random vector $\mathbf{x}$ be normally distributed according to $\mathbf{x} \sim N_n(\mu, \Sigma)$, then the $m \times 1$ random vector $\mathbf{y}$, obtained by the linear transformation $\mathbf{y} = \mathrm{A}\,\mathbf{x} + \mathrm{c}$, where A denotes an $m \times n$ matrix of constants with full rank m and c an $m \times 1$ vector of constants, has the normal distribution $\mathbf{y} \sim N_m(\mathrm{A}\mu + \mathrm{c}, \mathrm{A}\Sigma\mathrm{A}^t)$.*

Proposition 2.2 *Let the $n \times 1$ random vector $\mathbf{x}$ be normally distributed according to $\mathbf{x} \sim N_n(\mu, \Sigma)$, then $\mathbf{x}^t\mathrm{A}\,\mathbf{x}$, where A is an $n \times n$ matrix of constants has noncentral chi-square distribution $\chi'^2(rank\,\mathrm{A}, \mu^t\mathrm{A}\mu)$, if and only if the matrix $\mathrm{A}\Sigma$ is idempotent.*

Proposition 2.3 *If A is idempotent, then $rank\,\mathrm{A} = tr\mathrm{A}$.*

Proposition 2.4 *If the $n \times n$ matrix A with $rank\mathrm{A} = r$ is idempotent, then $\mathrm{I}_n - \mathrm{A}$ is also idempotent with $rank(\mathrm{I}_n - \mathrm{A}) = n - r$.*

Proposition 2.5 *Let the $n \times 1$ random vector $\mathbf{x}$ be normally distributed according to $\mathbf{x} \sim N_n(\mu, \Sigma)$, then the linear form $\mathrm{A}\mathbf{x}$ and the quadratic form $\mathbf{x}^t\mathrm{B}\mathbf{x}$ with the positive definite or positive semidefinite matrix B are independent, if and only if $\mathrm{A}\Sigma\mathrm{B} = 0$.*

Let A be a real square matrix of order n. Let α and β be nonempty subsets of the set $N_n = \{1, \ldots, n\}$. By $\mathrm{A}[\alpha, \beta]$ we denote the submatrix of A, composed of the rows with numbers from α and the columns with numbers from β. When $\beta \equiv \alpha$, $\mathrm{A}[\alpha, \alpha]$ is denoted simply by $\mathrm{A}[\alpha]$.

The Bellman gamma distribution is a matrix variate distribution, which is a generalization of the Wishart and the matrix gamma distributions. The next definition is given in [3]. By $\Gamma_n^*(a_1, \ldots, a_n)$ is denoted the generalized multivariate gamma function, $\Gamma_n^*(a_1, \ldots, a_n) = \pi^{n(n-1)/4} \prod_{j=1}^{n} \Gamma\left(a_j - \frac{1}{2}(j-1)\right)$, for $a_j > \frac{1}{2}(j-1)$, $j = 1, \ldots, n$.

Definition 2.6 *A random positive definite matrix* $\mathbf{U}$ *($n \times n$) is said to follow Bellman gamma type I distribution, denoted by* $\mathbf{U} \sim BG_n^I(a_1, \ldots, a_n; \mathrm{C})$*, if its probability density function is given by*

$$\frac{\prod_{i=1}^{n} (\det \mathrm{C}[\{i, \ldots, n\}])^{a_i - a_{i-1}}}{\Gamma_n^*(a_1, \ldots, a_n)} \frac{(\det \mathrm{U})^{a_n - (n+1)/2}}{\prod_{i=2}^{n} (\det \mathrm{U}[\{1, \ldots, i-1\}])^{a_i - a_{i-1}}} \, etr(-\mathrm{CU}),$$

where C *($n \times n$) is a positive definite constant matrix,* $a_0 = 0$ *and* $a_j > \frac{1}{2}(j-1)$*,* $j = 1, \ldots, n$.

3 Main Results

Let us replace the missing values in (1) by zero and denote the obtained matrix by $\mathbf{X}$,

$$\mathbf{X} = \begin{pmatrix} x_{1,1} & \cdots & x_{1,m_1} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ x_{n-1,1} & \cdots & x_{n-1,m_1} & \cdots & x_{n-1,m_{n-1}} & 0 & \cdots & 0 \\ x_{n,1} & \cdots & x_{n,m_1} & \cdots & x_{n,m_{n-1}} & x_{n,m_{n-1}+1} & \cdots & x_{n,m_n} \end{pmatrix} \quad (2)$$

Theorem 3.1 *Let the matrix* $\mathbf{X}$*, defined by (2) presents the observations (1) on a random vector* $(X_1, \ldots, X_n)^t$ *with multivariate standard normal distribution* $N_n(0, \mathrm{I}_n)$*. Then the matrix* $\mathbf{W} = \mathbf{X}\mathbf{X}^t$ *has Bellman gamma type I distribution* $BG_n^I\left(\frac{m_1}{2}, \ldots, \frac{m_n}{2}; \frac{1}{2}\mathrm{I}_n\right)$.

Proof. Let $\mathbf{W} = (w_{i,j})$ be the matrix $\mathbf{W} = \mathbf{X}\mathbf{X}^t$ and let $\mathbf{w}_i = (w_{1,i}, \ldots, w_{i-1,i})^t$, $i = 2, \ldots, n$. Let us denote by $\mathbf{z}_i$ the vector of available observations on X_i, i.e. $\mathbf{z}_i = (x_{i,1}, \ldots, x_{i,m_i})^t$, $i = 1, \ldots, n$. Let $\mathbf{X}_i$ be the matrix $\mathbf{X}_i = \mathbf{X}[\{1, \ldots, i\}, \{1, \ldots, m_{i+1}\}]$, $i = 1, \ldots, n-1$. The distribution of $\mathbf{z}_i$ is $N_{m_i}(0, \mathrm{I}_{m_i})$, $i = 1, \ldots, n$. Since $\mathbf{w}_i = \mathbf{X}_{i-1}\mathbf{z}_i$ and $\mathbf{X}_{i-1}\mathbf{X}_{i-1}^t =$

$\mathbf{W}[\{1,\dots,i-1\}]$, according to Proposition 2.1 the distribution of $\mathbf{w}_i$, $i = 2,\dots,n$ is $N_{i-1}(0,\mathbf{W}[\{1,\dots,i-1\}])$ under the condition that the matrix $\mathbf{X}_{i-1}$ has fixed elements. It is easy to see that

$$\mathbf{W}[\{1,\dots,i\}] = \begin{pmatrix} \mathbf{W}[\{1,\dots,i-1\}] & \mathbf{w}_i \\ \mathbf{w}_i^t & w_{i,i} \end{pmatrix}.$$

Let

$$w_{(i)} = w_{i,i} - \mathbf{w}_i^t(\mathbf{W}[\{1,\dots,i-1\}])^{-1}\mathbf{w}_i, \tag{3}$$

i.e.

$$w_{(i)} = \frac{\det \mathbf{W}[\{1,\dots,i\}]}{\det \mathbf{W}[\{1,\dots,i-1\}]},\ i = 2,\dots,n. \tag{4}$$

For $i = 2,\dots,n$, $w_{(i)}$ can be written in the form

$$w_{(i)} = \mathbf{z}_i^t\,\mathbf{z}_i - \mathbf{z}_i^t\,\mathbf{X}_{i-1}^t(\mathbf{X}_{i-1}\mathbf{X}_{i-1}^t)^{-1}\mathbf{X}_{i-1}\mathbf{z}_i = \mathbf{z}_i^t[\mathrm{I}_{m_i} - \mathbf{X}_{i-1}^t(\mathbf{X}_{i-1}\mathbf{X}_{i-1}^t)^{-1}\mathbf{X}_{i-1}]\mathbf{z}_i.$$

Since

$$\mathbf{X}_{i-1}[\mathrm{I}_{m_i} - \mathbf{X}_{i-1}^t(\mathbf{X}_{i-1}\mathbf{X}_{i-1}^t)^{-1}\mathbf{X}_{i-1}] = \mathbf{X}_{i-1} - \mathbf{X}_{i-1} = 0,$$

according to Proposition 2.5, $\mathbf{w}_i$ is independent of $w_{(i)}$ under the condition that the matrix $\mathbf{X}_{i-1}$ has fixed elements. Additionally, since the matrix $\mathbf{X}_{i-1}^t(\mathbf{X}_{i-1}\,\mathbf{X}_{i-1}^t)^{-1}\mathbf{X}_{i-1}$ is idempotent, from Proposition 2.3 we have that

$$rank(\mathbf{X}_{i-1}^t(\mathbf{X}_{i-1}\mathbf{X}_{i-1}^t)^{-1}\mathbf{X}_{i-1}) = tr(\mathbf{X}_{i-1}^t(\mathbf{X}_{i-1}\mathbf{X}_{i-1}^t)^{-1}\mathbf{X}_{i-1}) = tr(\mathbf{I}_{i-1}) = i-1.$$

Hence, according to Proposition 2.4 the matrix $\mathrm{I}_{m_i} - \mathbf{X}_{i-1}^t(\mathbf{X}_{i-1}\mathbf{X}_{i-1}^t)^{-1}\mathbf{X}_{i-1}$ is also idempotent with the rank $m_i - i + 1$. Now, applying Proposition 2.2 we get that $w_{(i)}$ has chi-square distribution $\chi^2\,(m_i - i + 1)$, again under the condition that the matrix $\mathbf{X}_{i-1}$ has fixed elements. The conditional distributions of $w_{(i)}$ and $\mathbf{w}_i$ depend only on the elements of the matrix $\mathbf{W}[\{1,\dots,i-1\}]$. Therefore the joint density of $w_{1,1}$, $w_{(2)}$, $\mathbf{w}_2$, $\dots$, $w_{(n)}$, $\mathbf{w}_n$ will have the form

$$\frac{w_{1,1}^{\frac{m_1}{2}-1}e^{-\frac{1}{2}w_{1,1}}}{2^{\frac{m_1}{2}}\Gamma\left(\frac{m_1}{2}\right)}\prod_{i=2}^{n}\left(\frac{w_{(i)}^{\frac{m_i-i+1}{2}-1}e^{-\frac{1}{2}w_{(i)}}}{2^{\frac{m_i-i+1}{2}}\Gamma\left(\frac{m_i-i+1}{2}\right)}\frac{e^{-\frac{1}{2}w_i^t(W[\{1,\dots,i-1\}])^{-1}w_i}}{(2\pi)^{\frac{i-1}{2}}\left(\det W[\{1,\dots,i-1\}]\right)^{\frac{1}{2}}}\right).$$

By the transformation of the variables $w_{(i)}$ into $w_{i,i}$ by means of (3) with $\det \mathrm{J} = 1$ we get the distribution of the elements $w_{i,j}$ of $\mathbf{W}$, which using (4) can be written in the form

$$\frac{1}{2^{\frac{m_1+\cdots+m_n}{2}}\Gamma_n^*\left(\frac{m_1}{2},\dots,\frac{m_n}{2}\right)}\frac{(\det W)^{[m_n-(n+1)]/2}}{\prod\limits_{i=2}^{n}(\det W[\{1,\dots,i-1\}])^{(m_i-m_{i-1})/2}}\,etr\left(-\frac{1}{2}W\right).$$

Consequently, according to Definition 2.1 $\mathbf{W} \sim BG_n^I\left(\frac{m_1}{2}, \dots, \frac{m_n}{2}; \frac{1}{2}\mathrm{I}_n\right)$. $\square$

Let $\mathbf{z}_i$ be the vector of available observations on X_i in (1), i.e. $\mathbf{z}_i = (x_{i,1}, \dots, x_{i,m_i})^t$, $i = 1, \dots, n$. Let us denote by $\bar{\mathbf{z}}_i$ the mean of the elements of $\mathbf{z}_i$, $i = 1, \dots, n$. Consider the data matrix

$$\mathbf{X} = \begin{pmatrix} x_{1,1} & \cdots & x_{1,m_1} & \cdots & \bar{\mathbf{z}}_1 & \bar{\mathbf{z}}_1 & \cdots & \bar{\mathbf{z}}_1 \\ \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ x_{n-1,1} & \cdots & x_{n-1,m_1} & \cdots & x_{n-1,m_{n-1}} & \bar{\mathbf{z}}_{n-1} & \cdots & \bar{\mathbf{z}}_{n-1} \\ x_{n,1} & \cdots & x_{n,m_1} & \cdots & x_{n,m_{n-1}} & x_{n,m_{n-1}+1} & \cdots & x_{n,m_n} \end{pmatrix}, \tag{5}$$

in which we substitute $\bar{\mathbf{z}}_1, \dots, \bar{\mathbf{z}}_{n-1}$ for the missing values in $1, \dots, n-1$'th row respectively of the data in (1).

Theorem 3.2 *Let the matrix* $\mathbf{X}$*, defined by (5) presents the observations (1) on a random vector* $(X_1, \dots, X_n)^t$ *with multivariate normal distribution* $N_n(\mu, \mathrm{I}_n)$*. Let* $\mathbf{x}_i$*,* $i = 1, \dots, m_n$ *be the column vectors of the matrix* $\mathbf{X}$*,* $\bar{\mathbf{x}}$ *be the vector* $\bar{\mathbf{x}}$=$(\bar{z}_1, \dots, \bar{z}_n)^t$ *and* $\mathbf{S}$ *be the matrix*

$$\mathbf{S} = \sum_{i=1}^{m_n} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^t. \tag{6}$$

Then $\bar{\mathbf{x}}$ *and* $\mathbf{S}$ *are independent,* $\bar{\mathbf{x}} \sim N_n(\mu, diag(m_1^{-1}, \dots, m_n^{-1}))$ *and* $\mathbf{S}$ *has Bellman gamma distribution* $BG_n^I\left(\frac{m_1-1}{2}, \dots, \frac{m_n-1}{2}; \frac{1}{2}\mathrm{I}_n\right)$.

Proof. Let Θ be an orthogonal $m_n \times m_n$ matrix of the form:

$$\Theta = \begin{pmatrix} \frac{1}{\sqrt{m_n}} & \theta_{1,2} & \cdots & \theta_{1,m_1} & \theta_{1,m_1+1} & \cdots & \theta_{1,m_2} & \cdots & \theta_{1,m_n} \\ \frac{1}{\sqrt{m_n}} & \theta_{2,2} & \cdots & \theta_{2,m_1} & \theta_{1,m_1+1} & \cdots & \theta_{1,m_2} & \cdots & \theta_{1,m_n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ \frac{1}{\sqrt{m_n}} & \theta_{m_1,2} & \cdots & \theta_{m_1,m_1} & \theta_{1,m_1+1} & \cdots & \theta_{1,m_2} & \cdots & \theta_{1,m_n} \\ \frac{1}{\sqrt{m_n}} & 0 & \cdots & 0 & \theta_{m_1+1,m_1+1} & \cdots & \theta_{m_1+1,m_2} & \cdots & \theta_{1,m_n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ \frac{1}{\sqrt{m_n}} & 0 & \cdots & 0 & \theta_{m_2,m_1+1} & \cdots & \theta_{m_2,m_2} & \cdots & \theta_{1,m_n} \\ \frac{1}{\sqrt{m_n}} & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & \theta_{1,m_n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ \frac{1}{\sqrt{m_n}} & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & \theta_{m_n,m_n} \end{pmatrix}.$$

The all elements in the first column of Θ are equal to $\frac{1}{\sqrt{m_n}}$. In columns from 2 to m_1, the last $m_n - m_1$ elements are equal to zero. In columns from $m_j + 1$

to m_{j+1}, $j = 1, \dots, n-1$ the last $m_n - m_{j+1}$ elements are equal to zero and the first m_j elements are equal to each other. It can be easily checked that such orthogonal matrix exists and even is not unique. Let $\mathbf{Y} = (y_{i,j})$ be the matrix $\mathbf{Y} = \mathbf{X}\,\Theta$. Then $\mathbf{Y}\,\mathbf{Y}^t = \mathbf{X}\,\Theta\Theta^t\mathbf{X}^t = \mathbf{X}\,\mathbf{X}^t$. Let us denote by $\mathbf{y}_i$, $i = 1, \dots, m_n$ the column vectors of $\mathbf{Y}$. It is easy to see that $\mathbf{y}_1 = \sqrt{m_n}\,\bar{\mathbf{x}}$. Therefore the matrix $\mathbf{S}$, defined by (6) can be written in the form

$$\mathbf{S} = \sum_{i=1}^{m_n} \mathbf{x}_i\mathbf{x}_i^t - m_n\bar{\mathbf{x}}\bar{\mathbf{x}}^t = \mathbf{X}\,\mathbf{X}^t - \mathbf{y}_1\mathbf{y}_1^t = \mathbf{Y}\,\mathbf{Y}^t - \mathbf{y}_1\mathbf{y}_1^t = \sum_{i=2}^{m_n} \mathbf{y}_i\mathbf{y}_i^t. \tag{7}$$

Since the matrix Θ is orthogonal, $\theta_1^t\,\theta_i = 0$ for $i = 2, \dots, m_n$, where θ_i, $i = 1, \dots, m_n$ denote the column vectors of Θ. Therefore, the sum of elements of θ_i equals to zero, $i = 2, \dots, m_n$. Hence it follows that $E(\mathbf{y}_i) = 0$, $i = 2, \dots, m_n$. It can be checked that in the matrix $\mathbf{Y}$ the all elements, lying on the places of the missing data in (1) are equal to zero.

Let $y_{k,s}$, $s \neq 1$ be a nonzero element of $\mathbf{Y}$. Then $1 < s \leq m_k$, therefore the last $m_n - m_k$ elements of the vector θ_s are equal to zero. Let the vector θ_s has m nonzero elements, then

$$y_{k,s} = x_{k,1}\theta_{1,s} + \cdots + x_{k,m}\theta_{m,s}. \tag{8}$$

The distribution of $y_{k,s}$ as a linear combination of independent normal distributed random variables is also normal. The variance of $y_{k,s}$ is

$$Var(y_{k,s}) = \theta_{1,s}^2 Var(x_{k,1}) + \cdots + \theta_{m,s}^2 Var(x_{k,m}) = \theta_{1,s}^2 + \cdots + \theta_{m,s}^2 = 1.$$

For $k = 1, \dots, n$

$$y_{k,1} = \sqrt{m_n}\bar{\mathbf{z}}_k = \frac{\sqrt{m_n}}{m_k}(x_{k,1} + \cdots + x_{k,m_k}). \tag{9}$$

Consequently $y_{k,1}$ is also a linear combination of the observations in the k'th row of (1) and hence is normally distributed,

$$E(y_{k,1}) = \frac{\sqrt{m_n}}{m_k}\left[E(x_{k,1}) + \cdots + E(x_{k,m_k})\right] = \sqrt{m_n}\mu_k,$$

where μ_k is the k'th coordinate of the mean vector μ and

$$Var(y_{k,1}) = \frac{m_n}{m_k^2}\left[Var(x_{k,1}) + \cdots + Var(x_{k,m_k})\right] = \frac{m_n}{m_k}.$$

Since $(X_1, \dots, X_n)^t \sim N_n(\mu, \mathrm{I}_n)$, the vectors $\mathbf{z}_1, \dots, \mathbf{z}_n$ are independent. Consequently, if $k_1, \dots, k_p$ are different integers from the interval $[1, n]$ and $y_{k_1,s_1}, \dots, y_{k_p,s_p}$ are nonzero elements of the matrix $\mathbf{Y}$, then they are independent. Hence for $i = 1, \dots, m_n$ the nonzero elements of the vector $\mathbf{y}_i$ have joint multivariate normal distribution. For $i = 2, \dots, m_n$ it is $N_{n-j}(0, \mathrm{I}_{n-j})$, where j is the biggest integer, such that $m_j < i$ $(m_0 = 1)$. Since $\mathbf{y}_1 = \sqrt{m_n}\,\bar{\mathbf{x}}$, the distribution of $\bar{\mathbf{x}}$ is $N_n(\mu, diag(m_1^{-1}, \dots, m_n^{-1}))$. From (8) and (9) it follows that if $y_{k,s_1}, \dots, y_{k,s_p}$ are arbitrary nonzero elements of $\mathbf{Y}$, then they have joint multivariate normal distribution. Moreover, (8) and (9) can be written in the form

$$y_{k,s} = (x_{k,1} - \mu_k)\theta_{1,s} + \cdots + (x_{k,m} - \mu_k)\theta_{m,s},\ s \neq 1,$$

$$y_{k,1} = (x_{k,1} - \mu_k)\frac{\sqrt{m_n}}{m_k} + \cdots + (x_{k,m_k} - \mu_k)\frac{\sqrt{m_n}}{m_k} + \mu_k\sqrt{m_n}.$$

Hence for $s > 1$

$$\begin{aligned} Cov(y_{k,1}, y_{k,s}) = E(y_{k,1}y_{k,s}) = E(x_{k,1} - \mu_k)^2\theta_{1,s}\frac{\sqrt{m_n}}{m_k} + \cdots \\ + E(x_{k,m} - \mu_k)^2\theta_{m,s}\frac{\sqrt{m_n}}{m_k} = (\theta_{1,s} + \cdots + \theta_{m,s})\frac{\sqrt{m_n}}{m_k} = 0 \end{aligned}$$

and for $1 < s < q$

$$\begin{aligned} Cov(y_{k,s}, y_{k,q}) = E(y_{k,s}y_{k,q}) = E(x_{k,1} - \mu_k)^2\theta_{1,s}\theta_{1,q} + \cdots \\ + E(x_{k,m} - \mu_k)^2\theta_{m,s}\theta_{m,q} = \theta_{1,s}\theta_{1,q} + \cdots + \theta_{m,s}\theta_{m,q} = 0. \end{aligned}$$

Therefore $y_{k,s_1}, \dots, y_{k,s_p}$ are independent and hence the vectors $\mathbf{y}_1, \dots, \mathbf{y}_{m_n}$ are independent. Consequently from (7) and Theorem 3.1 the Theorem follows. □

For the data in (1), let us consider the hypotheses $H_0 : \Sigma = \Sigma_0$ against $H_a : \Sigma \neq \Sigma_0$, where Σ_0 is an arbitrary positive definite matrix of size n. It is shown in [4], that the testing problem is invariant under a suitable transformation of the data in (1) and without loss of generality we can assume that $\Sigma_0 = \mathrm{I}_n$, where I_n is the identity matrix of size n.

It is easy to see that under $H_0 : \Sigma = \mathrm{I}_n$, the maximum likelihood estimations for the missing values in the j'th row in (1) are equal to $\bar{\mathbf{z}}_j$, $j = 1, \dots, n$. The matrix $\frac{1}{(m_n-1)}\mathbf{S}$ is actually the empirical covariance matrix, obtained from the data matrix (5).

Let us consider the modified likelihood ratio test statistic λ^*,

$$\lambda^* = (e/m_n)^{nm_n/2}(\det \mathbf{S})^{(m_n-1)/2}e^{-(tr\mathbf{S})/2},$$

which is unbiased in the classical case of fully observed data matrix (see [2]).

Theorem 3.3 *Under $H_0 : \Sigma = \mathrm{I}_n$, λ^* is distributed as the product*

$$K\,(\zeta_1 \dots \zeta_{n-1})^{(m_n-1)/2}\eta_1 \dots \eta_n,$$

where K is the constant $K = (e/m_n)^{nm_n/2}\,2^{n(m_n-1)/2}$, $\zeta_1, \dots, \zeta_{n-1}$, $\eta_1, \dots, \eta_n$ are mutually independent random variables, ζ_j has Beta distribution Beta ($(m_{j+1}-j-1)/2, j/2)$, $j = 1, \dots, n-1$ and $\eta_j \sim \xi_j^{(m_n-1)/2} e^{-\xi_j}$, where ξ_j is Gamma distributed $G((m_j-1)/2, 1)$, $j = 1, \dots, n$.

The proof of Theorem 3.3 will appear in a subsequent paper "Stochastic representations of the Bellman gamma distribution".

Since Theorem 3.3 give the exact distribution of the test statistic λ^*, the test procedure, suggested in this paper is proper for small samples.

References

[1] W. Chang and D. St. P. Richards, Finite - sample inference with monotone incomplete multivariate normal data I, *J. Multivariate Anal.,* In Press, Corrected Proof, Available online 14 May 2009.

[2] N.C. Giri, *Multivariate Statistical Analysis,* Marcel Dekker Inc., New York, 2004.

[3] A.K. Gupta and D.K. Nagar, *Matrix variate distributions,* Chapman & Hall/CRC, 2000.

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

[5] K. R. Koch, *Estimation and Hypothesis Testing in Linear Models,* Springer - Verlag, Berlin, 1999.

[6] J.A. Little and D.B. Rubin, *Statistical Analysis With Missing Data,* 2nd edition, Wiley - Interscience, Hoboken, NJ, 2002.

Received: May, 2009