LIKELIHOOD RATIO TESTS OF CERTAIN HYPOTHESES
ABOUT PRINCIPAL COMPONENTS

Yasunori Fujikoshi*

1. Introduction

Let a random vector $x$ have a $p$-variate normal distribution with mean vector $\mu$ and covariance matrix $\Sigma$. Let $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_p \geq 0$ be the latent roots of $\Sigma$ and $\gamma_1, \gamma_2, \cdots, \gamma_p$ be a set of the corresponding orthonormal latent vectors of $\Sigma$. Then

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

where $\Gamma = (\gamma_1, \cdots, \gamma_p)$ and $\Lambda = \text{diag}(\delta_1, \cdots, \delta_p)$. The $i$-th principal component of $x$ is defined by $\xi_i = \gamma_i'(x - \mu)$. The hypotheses considered here are

$H_{1k}: \gamma_1 = \gamma_{10}, \cdots, \gamma_k = \gamma_{k0}$ and $\delta_1 = \delta_{10}, \cdots, \delta_k = \delta_{k0}$,

$H_k: \gamma_1 = \gamma_{10}, \cdots, \gamma_k = \gamma_{k0}$ and $\delta_1 = \cdots = \delta_p = \delta$,

where $\gamma_{10}, \cdots, \gamma_{k0}$ are given orthonormal vectors and $\delta_{i0}$'s are given values such that $\delta_{10} \geq \cdots \geq \delta_{k0} > 0$. In each case the alternatives are $K_i: \Sigma > 0$. We also consider to test the hypothesis $H_i$ against the alternatives $K_i$: $\delta_1 = \cdots = \delta_p = \delta$.

Anderson [2] has given an asymptotic test for testing $H_i$ against $K_i$ under the assumption that $\gamma_{10}$ corresponds to a latent root of multiplicity one. Kshirsagar [4] considered a goodness of fit test for a hypothetical principal component in certain situations and gave exact tests for testing $H_i$ against $K_i$ when $\delta$ is known and for testing $H_i$ against $K_i$. The procedures hitherto proposed for these hypotheses are different from ones by the likelihood ratio (=LR) method. The purpose of this paper is to derive LR tests for these hypotheses. The method given here will be useful for deriving LR tests of certain hypotheses on latent roots and vectors of certain matrices. The test statistics are given in closed forms. The distributions of these statistics are under investigation and will be treated in a later paper.

2. Some lemmas

The following Lemma 1 has been proved in Anderson [2, p. 131].

Lemma 1. Let $X = \text{diag}(x_1, \cdots, x_p)$ and $Y = \text{diag}(y_1, \cdots, y_p)$, where $0 < x_1 \leq x_2 \leq \cdots \leq x_p < \infty$ and $\infty > y_1 > y_2 > \cdots > y_p > 0$. Then

$$\min_{\theta \in \mathcal{O}(p)} \text{tr} X \theta' \theta = \sum_{i=1}^p x_i y_i,$$

where $\mathcal{O}(p)$ stands for the set of all orthogonal matrices of order $p$.

Let

$$G(x_1, \cdots, x_p) = \left| \hat{x}_1 x_1 e^{-\hat{x}_1 x_1} \right|$$

and

$$\hat{x}_j = a_j/d_j \quad \text{for } j = 1, 2, \cdots, p.$$
Lemma 2. If $0 < a_1 \leq \cdots \leq a_p < \infty$ and $\infty > d_1 > \cdots > d_p > 0$, then the maximum $\lambda$ is given by

$$\lambda = \max_{0 < c \leq a_1, \cdots, x_p < \infty} G(x_1, \cdots, x_p)$$

where $c$ is any fixed positive number.

Proof. This result is obtained by using the property that $x_j e^{-x_j^2}$ is monotone increasing on $(0, x_j]$ and monotone decreasing on $[x_j, \infty)$.

Lemma 3. If $a_1 = a_2 = \cdots = a_p = a, d_1 > 0$ and $d_2 > \cdots > d_p > 0$, then the maximum $\lambda$ of $G(x_1, \cdots, x_p)$ with respect to $x_1, \cdots, x_p$ such that $0 < x_1 \leq \cdots \leq x_p < \infty$ is given by

$$\lambda = \max_{0 < c \leq \hat{x}_1, \cdots, \hat{x}_p} G(\hat{x}_1, \cdots, \hat{x}_p), \quad \text{if } b_2 \leq d_i,$$

$$= \max_{d_2 > \cdots > d_p > 0} G(\hat{y}_2, \cdots, \hat{y}_p), \quad \text{if } d_1 < b_1.$$ 

where $b=\hat{d}_i, b_j=(j-1)d_j - \sum_{t=1}^{j-1} d_i$ for $j=2, \cdots, p$ and $\hat{y}_j = a(1/j) \sum_{t=1}^{j-1} d_i^{-1}$ for $j=2, \cdots, p$.

Proof. From the proof of Lemma 2 we have an inequality

$$\lambda \leq G(\hat{x}_1, \cdots, \hat{x}_p).$$

If $b_2 \leq d_i$, then $x_1 \leq x_2 \leq \cdots \leq x_p$ and hence the equality holds in (2.6). If $b_1 \leq d_i < b_1$, then we have

$$\lambda \leq \max_{0 < c \leq \hat{x}_1, \cdots, \hat{x}_p} \max_{0 < c \leq \hat{x}_1, \cdots, \hat{x}_p} G(x_1, \cdots, x_p)$$

and $\hat{y}_1 \leq \hat{x}_1 \leq \cdots \leq x_p$. This implies that $\lambda$ in the case of $b_1 \leq d_i < b_1$ is given by $G(\hat{y}_2, \hat{y}_3, \hat{x}_1, \cdots, \hat{x}_p)$.

3. LR tests

Let a sample of size $N(N > p), x_1, x_2, \cdots, x_N$ be available for the testing problems considered in this paper. It is easy to see that the LR statistics can be found by considering

$$f(\Sigma) = |\Sigma|^{-n/2} \exp \left( -\frac{1}{2} \Sigma^{-1} S \right),$$

where $S = \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})'$, $\bar{x} = \frac{1}{N} \sum_{i=1}^{n} x_i$, and $\exp$ denotes the exponential of trace. Further it is well known [Anderson (1), Lemma 3.2.2] that

$$\max_{\hat{\Sigma}} f(\Sigma) = f(\hat{\Sigma}),$$

where $N \hat{\Sigma} = S$. First we derive the LR criterion $\lambda_1$ for testing $H_{1k}$ against $K_1$, which is given by

$$\lambda_1 = \max_{H_{1k}} f(\Sigma)/f(\hat{\Sigma}).$$

Put

$$\Gamma_0 = (\gamma_1, \cdots, \gamma_k), \quad \Delta_0 = \text{diag} (\delta_1, \cdots, \delta_k),$$

$$\Delta = \text{diag} (\delta_{k+1}, \cdots, \delta_p) \quad \text{and} \quad \Gamma_1 = (\gamma_{k+1}, \cdots, \gamma_p).$$
We may characterize $I'$ as

$$I' = I'_{x\theta}, \quad \theta \in \mathbb{R}^{p-k},$$

where $I'_{x\theta}$ is a fixed matrix such that $I'_{x\theta}'I'_{x\theta} = I_{p-k}$ and $I'_{x\theta}'I'_{x\theta} = 0$. Here $I_k$ is the identity matrix of order $k$. Then, under $H_{k}$, $f(\Sigma)$ can be expressed as

\begin{equation}
|d_{10}|^{-N/4} \text{etr} \left( -\frac{1}{2} d_{10}^{-1} I'_{x\theta} S \Gamma_{10}^{-1} \right) |d_{1}|^{-N/4} \text{etr} \left( -\frac{1}{2} d_{1}^{-1} \theta' \left( \Gamma_{10}' S \Gamma_{10} \right) \theta \right).
\end{equation}

The maximum over $\theta \in \mathbb{R}^{p-k}$ is obtained from Lemma 1 and, after that, the maximum over $d_{1}$ subject to $d_{k} \geq d_{k+1} \geq \cdots \geq d_{p} > 0$ is obtained from Lemma 2. Therefore we have

**Theorem 1.** The LR test criterion for testing $H_{k}$ against $K_{1}$ is given by

\begin{equation}
l_{k} = \begin{cases} T(\delta_{k+1}, \delta_{k+2}, \delta_{p}), & \text{if } \delta_{k+1} \leq \delta_{k}, \\
T(\delta_{k}, \delta_{k}, \delta_{p}), & \text{if } \delta_{k} < \delta_{k+1} \leq \delta_{k+1}, \quad j = 2, \ldots, p-k, \\
T(\delta_{k}, \delta_{k}, \delta_{k}, \delta_{k}), & \text{if } \delta_{k} > \delta_{p},
\end{cases}
\end{equation}

where

\begin{equation}
T(\delta_{k+1}, \delta_{k+2}, \delta_{p}) = \left| \frac{1}{N} S \right|^{N/2} \text{etr} \left( -\frac{1}{2} d_{10}^{-1} I'_{x\theta} S \Gamma_{10}^{-1} \right) \times \prod_{j=k+2}^{p} \delta_{j}^{N/2} e^{-d_{j} / (2d_{j})}
\end{equation}

and

\begin{equation}
\delta_{j} = d_{j} / N, \quad j = k+1, \ldots, p.
\end{equation}

Here $d_{k+1} > d_{k+2} > \cdots > d_{p} > 0$ are the latent roots of $I'_{x\theta} S \Gamma_{10}$. We note that

\begin{equation}
T(\delta_{k+1}, \delta_{k+2}, \delta_{p}) = \left| \frac{1}{N} S \right|^{N/2} \text{etr} \left( -\frac{1}{2} d_{10}^{-1} I'_{x\theta} S \Gamma_{10}^{-1} \right) \times \prod_{j=k+1}^{p} \delta_{j}^{N/2} e^{-d_{j} / (2d_{j})}
\end{equation}

and this statistic is the LR criterion [Gupta [3], Mallows [5]] for testing the hypothesis,

$H_{k}:$ $k$ given orthogonal and unit vectors $\gamma_{1}, \ldots, \gamma_{k}$ are latent vectors of $\Sigma$ and the corresponding roots are $\delta_{1}, \ldots, \delta_{k}$.

Next let us derive the LR test criterion for testing $H_{k}$ against $K_{1}$. Under $H_{k}$, we have

\begin{equation}
f(\Sigma) = \delta_{k}^{-N/2} e^{-d_{1} / (2d_{1})} \times |d_{10}|^{-N/4} \text{etr} \left( -\frac{1}{2} d_{10}^{-1} I'_{x\theta} S \Gamma_{10}^{-1} \right) = g(\Delta, \theta),
\end{equation}

where $d_{1} = \gamma_{1} S \gamma_{1}, \Delta = \text{diag}(\delta_{1}, \ldots, \delta_{p}), \theta \in \mathbb{R}^{p-1}$ and $I_{x\theta}$ is a fixed matrix such that $I'_{x\theta}'I_{x\theta} = I_{p-1}$ and $I'_{x\theta}'I_{x\theta} = 0$. Using Lemma 1 we get

\begin{equation}
\max_{H_{k}} f(\Sigma) = \max_{H_{k}} \max_{\theta \in \mathbb{R}^{p-1}} g(\Delta, \theta)
\end{equation}

where

\begin{equation}
\delta_{k} \geq \cdots \geq d_{p} \geq 0
\end{equation}

are the latent roots of $I'_{x\theta} S \Gamma_{10}$. Here we used the fact that $d_{1} > \cdots > d_{p} > 0$ with probability one. Applying Lemma 3 to the final expression of (3.9), we have

**Theorem 2.** The LR criterion $\lambda_{k}$ for testing $H_{k}$ against $K_{1}$ is given by
where

\[ T_k = |S| \left( \prod_{j=1}^{p} d_j \right)^{k} \left( \prod_{j=1}^{p} d_j \right)^{j} \left( \prod_{j=1}^{p} d_j \right)^{p-k}, \quad k = 1, 2, \ldots, p, \]

with the convention that \( \prod_{j=1}^{p} d_j = 1 \), and \( b_j = d_j - \sum_{i=j}^{p} (d_i - d_j) \) for \( j = 2, \ldots, p-1 \).

We note that

\[ T^{1/2} = |S|^{1/2} (\tau_0^T S \tau_0) \left( I_{10}^T S I_{10} \right)^{1/2} \]

and this statistic is the LR criterion [Gupta [3], Mallows [5]] for testing the hypothesis, \( H_0: \) given unit vector \( \tau_0 \) is the latent vector of \( \Sigma \).

Theorem 3. The LR criterion \( \lambda_3 \) for testing \( H_0 \) against \( K_1 \) is given as follows:

When \( \dot{\delta} \) is unknown,

\[ \lambda_3^{(1)} = \begin{cases} \frac{|W|}{w_{11}} \left( \sum_{i=1}^{p} w_{ii} \right)^{p-1}, & \text{if } w_{11} \geq 1 \sum_{i=1}^{p} w_{ii}, \\ \frac{|W|}{w_{11}} \left( \sum_{i=1}^{p} w_{ii} \right)^{p-1}, & \text{if } w_{11} < 1 \sum_{i=1}^{p} w_{ii}. \end{cases} \]

When \( \dot{\delta} \) is known,

\[ \lambda_3^{(2)} = \begin{cases} \frac{|W|}{w_{11}(N\dot{\delta})^{p-1}} \exp \left( \frac{-1}{w_{11}} \sum_{i=1}^{p} w_{ii} \right), \quad \text{if } w_{11} \geq N\dot{\delta}, \\ \frac{|W|}{w_{11}(N\dot{\delta})^{p-1} \text{tr} W}, \quad \text{if } w_{11} < N\dot{\delta}. \end{cases} \]

where \( W = (w_{ii}) = I'_{10} S I_{10}, \quad I'_{10} = (\tau_0^T, I_{10}) \) and \( I_{10} \) is a fixed matrix such that \( I_{10}^T S \tau_0 = 0 \).

Proof. Under \( H_0 \), we can write \( f(\Sigma) \) as

\[ g(\dot{\delta}, \dot{\delta}) = \dot{\delta}^{N/2} \dot{\delta}^{-(p-1)N/2} \exp \left( \frac{-1}{w_{11}} \sum_{i=1}^{p} w_{ii} \right). \]

Here \( 0 < \dot{\delta}_1 \leq \dot{\delta} < \infty \). By the similar consideration as in Lemmas 2 and 3 we find that

\[ \max_{\dot{\delta}_1 \geq \dot{\delta}} \frac{g(\dot{\delta}_1, \dot{\delta})}{g(\dot{\delta}, \dot{\delta})} = \begin{cases} 1, & \dot{\delta}_1 \geq \dot{\delta}, \\ \dot{\delta}_1, & \dot{\delta}_1 < \dot{\delta}, \end{cases} \]

and

\[ \max_{\dot{\delta}_1} \frac{g(\dot{\delta}_1, \dot{\delta})}{g(\dot{\delta}_1, \dot{\delta})} = \begin{cases} 1, & \dot{\delta}_1 \geq \dot{\delta}, \\ \dot{\delta}_1, & \dot{\delta}_1 < \dot{\delta}, \end{cases} \]

where \( \dot{\delta}_1 = w_{11}/N, \dot{\delta}_2 = ((p-1)N)^{-1} \sum_{i=1}^{p} w_{ii} \) and \( \dot{\delta} = \text{tr} W/(pN) \). From the above results and (3.2) we obtain the desired result.

Let \( d_1 > d_2 > \cdots > d_p > 0 \) be the latent roots of \( S \). When \( \dot{\delta} \) is unknown, from Anderson [2, p. 131] we have

\[ \max_{\dot{\delta}_1} f(\Sigma) = (d_1/N)^{-N/2} ((N(p-1))^{-1} \sum_{i=1}^{p} d_i)^{-(p-1)/2} e^{-pN/2}. \]

This result and (3.15) imply the following theorem:

Theorem 4. The LR criterion \( \lambda_4 \) for testing \( H_0 \) against \( K_1 \) when \( \dot{\delta} \) is unknown is given by
Likelihood Ratio Tests of Certain Hypotheses

(3.18) \[ \lambda_i = \begin{cases} \frac{d_1 \left( \frac{1}{p-1} \sum_i d_i \right)^{p-1}}{w_{ii} \left( \frac{1}{p-1} \sum_i w_{ii} \right)^{p-1}}, & \text{if } w_{ii} \geq \frac{1}{p-1} \sum_i w_{ii}, \\ \frac{d_1 \left( \frac{1}{p-1} \sum_i d_i \right)^{p-1}}{\frac{1}{p} \text{tr } W^p}, & \text{if } w_{ii} < \frac{1}{p-1} \sum_i w_{ii}, \end{cases} \]

with the same notations as in Theorem 3. Here \( d_1 > d_2 > \cdots > d_p > 0 \) are the latent roots of \( W \).

To carry out the tests given in this paper we need to derive the null distributions of the LR criterion \( \lambda_j (j=1, 2, 3, 4) \), which are under investigation.

References