JOURNAL OF MULTIVARIATE ANALYSIS 32, 313–325 (1990)

A Class of Tests for a General Covariance Structure

HIROFUMI WAKAKI

Oita University, Oita, Japan

SHINTO EGUCHI

Shimane University, Shimane, Japan

AND

YASUNORI FUJIKOSHI

Hiroshima University, Hiroshima, Japan

Communicated by the Editors

Let S be a $p \times p$ random matrix having a Wishart distribution $W_p(n, n^{-1}\Sigma)$. For testing a general covariance structure $\Sigma = \Sigma(\xi)$, we consider a class of test statistics $T_h = n \inf \rho_h(S, \Sigma(\xi))$, where $\rho_h(\Sigma_1, \Sigma_2) = \sum_{j=1}^{p} h(\lambda_j)$ is a distance measure from Σ_1 to Σ_2, λ_i's are the eigenvalues of $\Sigma_1 \Sigma_2^{-1}$, and h is a given function with certain properties. This paper gives an asymptotic expansion of the null distribution of T_h up to the order n^{-1}. Using the general asymptotic formula, we give a condition for T_h to have a Bartlett adjustment factor. Two special cases are considered in detail when Σ is a linear combination or Σ^{-1} is a linear combination of given matrices.

1. INTRODUCTION

Let S be a $p \times p$ random matrix having a Wishart distribution $W_p(n, n^{-1}\Sigma)$. It is assumed that $n \geqslant p$, so that $S \in \Delta \equiv$ the set of all the $p \times p$ symmetric positive definite matrices, with probability one. We consder the problem of testing $H_0 : \Sigma \in \Delta_0$ against $H_1 : \Sigma \in \Delta - \Delta_0$, where Δ_0 is defined as

$$\Delta_0 = \{\Sigma(\xi); \xi \in \Xi\} \tag{1.1}$$

with an open set Ξ of R^q. It is assumed that

Received June 23, 1989.
AMS 1980 subject classifications: primary 62H10; secondary 62H15.
Key words and phrases: asymptotic expansion, class of test statistics, general covariance structure, linear structure, null distribution.

0047-259X/90 $3.00

A1. All the elements of $\Sigma(\xi)$ are known C^4-class functions on H_0, and the Jacobian matrix of $\Sigma(\xi)$ is of full rank.

Thus Δ_0 is a smooth subsurface with coordinates $\xi = (\xi^1, ..., \xi^q)'$ in the total space Δ. The hypothesis H_0 involves various covariance structures as special cases.

We consider a class of test statistics via minimization of the following divergence measures from S to Δ_0. Let h be a C^4-function on $(0, \infty)$ satisfying that

A2. $h(1) = 0$, $h_1 = 0$ and $h_2 = 1$,

A3. $h(\lambda) > 0$ for any $\lambda \neq 1$,

where h_r denotes the rth derivative of h at $\lambda = 1$. For arbitrary two matrices Σ_1 and Σ_2 in Δ we define a distance measure from Σ_1 to Σ_2 by

$$\rho_h(\Sigma_1, \Sigma_2) = \sum_{i=1}^{p} h(\lambda_i),$$

where λ_i's are the eigenvalues of $\Sigma_1 \Sigma_2^{-1}$. Note that $\rho_h(\Sigma_1, \Sigma_2) \geqslant 0$ with equality if and only if $\Sigma_1 = \Sigma_2$ because of A3. However, in general, ρ_h is non-symmetric and does not satisfy the triangle law. We consider a class of test statistics

$$T_h = n \inf_{\xi \in \Xi} \rho_h(S, \Sigma(\xi)) = n\rho_h(S, \Sigma(\hat{\xi}_h)), \tag{1.2}$$

where $\hat{\xi}_h$ is a minimizing point. For example, for $h(\lambda) = -\log \lambda + \lambda - 1$, ρ_h is the Kullback divergence and the corresponding statistic T_h is only based on the log-likelihood ratio criterion. Another typical example is $h(\lambda) = (\lambda - 1)^2/2$. We note that each T_h has parametrization-invariance, which property is common in methods via minimization or maximization (cf. Barndorff-Nielsen and Cox [5]). Swain [15] considered $\rho_h(S, \hat{\Sigma})$ as a class of factor analysis estimation procedures and showed that for every h satisfying A1 and A2, $\hat{\xi}_h$ is a consistent and asmptotically efficient estimator of ξ. Further, Eguchi [8] showed that $\hat{\xi}_h$ is second-order efficient if and only if

$$h_3 = -2. \tag{1.3}$$

These suggest that the asymptotic properties of T_h under the null hypotheses may be closely related with the local shape of h around $\lambda = 1$.

The main purpose of this paper is to extend an asymptotic distribution theory for T_h based on perturbation method and derive an asymptotic expansion of the null distribution of T_h up to the order n^{-1}. As a special result, it is shown that every T_h has asymptotically a chi-square distribution with $r=p(p+1)/2-q$ degrees of freedom. In general, a test statistic T_h is said to have a Bartlett adjustment factor in a strong sense if a modified statistics $T_h^*=mT_h$ satisfies

$$P(T_h^* \leqslant x \mid H_0)=P(\chi_r^2 \leqslant x)+o(n^{-1}), \tag{1.4}$$

where $m=O(1)$. We note that a Bartlett adjustment factor in a weak sense is determined by requiring only $E[T_h^*]=r+o(n^{-1})$ (For a recent discussion, see, e.g., Bandorff-Nielsen and Cox [5], MacCullagh and Cox [9]). Using our general expansion formula, it is shown that T_h has a Bartlett adjustment factor in a strong sense if and only if

$$h_3=-2 \qquad \text{and} \qquad h_4=6. \tag{1.5}$$

Consequently we see that the Bartlett adjustment factor is determined by only the local property of h. It is easily seen that $h(\lambda)=-\log\lambda-\lambda+1$ satisfies (1.5), and hence there exists a Bartlett adjustment factor for the likelihood ratio statistic.

It may be noted that asymptotic expansions of the distributions of T_h's in some special cases have been obtained by many authors (For example, Anderson [2–4], Muirhead [10], Nagao [11], Siotani, Hayakawa, and Fujikoshi [12], Sugiura [14]], etc.). An emphasis in this paper is put on an asymptotic distribution theory for T_h in a general case. In Section 2 we give stochastic expansions of $\hat{\xi}_h$ as well as T_h. In Section 3 we obtain an asymptotic expansion of the characteristic function of T_h which yields an asymptotic expansion of the null distribution of T_h up to the order n^{-1}. A key reduction in the expansion method is given in Lemma 3.2. As special cases, we consider the case that Σ is a linear combination or Σ^{-1} is a linear combination of given matrices. Some reductions are also given for the two cases.

2. Stochastic Expansion of T_h

Let ξ_0 be an arbitrary fixed point of Ξ. We shall derive a stochastic expansion of T_h at $\Sigma_0=\Sigma(\xi_0)$. For simplicity, let us denote as $\hat{\xi}=\hat{\xi}_h$, $\Sigma=\Sigma(\xi)$, $\Sigma_0=\Sigma(\xi_0)$, and $\hat{\Sigma}=\Sigma(\hat{\xi}_h)$. We shall expand T_h in terms of

$$V=\sqrt{n}\,\Sigma_0^{-1/2}(S-\Sigma_0)\Sigma_0^{-1/2} \tag{2.1}$$

which is $O_p(1)$. Some of differential-geometrical notions (for example, see Amari [1], Eguchi [8]) are used in the derivation of the expansion of T_h.

First we summarize the notations used in this paper. Let

$$\partial_a = \partial/\partial\xi^a, \qquad J_{ab\cdots} = \Sigma_0^{1/2}[\partial_a\partial_b\cdots\Sigma^{-1}]_0\Sigma_0^{1/2},$$
$$\hat{J}_{ab\cdots} = \Sigma_0^{1/2}[\partial_a\partial_b\cdots\Sigma^{-1}]_\wedge\Sigma_0^{1/2},$$

where $[\ ]_0$ and $[\ ]_\wedge$ denote the quantity $[\ ]$ evaluated at $\xi = \xi_0$ and $\hat{\xi}$, respectively. Noting that the log-likelihood function is

$$l(\xi) = \frac{n}{2}\{-\operatorname{tr} S\Sigma^{-1} + \log|\Sigma^{-1}|\} + \text{const},$$

we can write the score and the Fislher information matrix as

$$s_a = n^{-1/2}[\partial_a l(\xi)]_0 = -\tfrac{1}{2}\operatorname{tr} J_a V, \qquad a = 1, ..., q,$$

and

$$G = (g_{ab}), \quad g_{ab} = E(s_a s_b) = \tfrac{1}{2}\operatorname{tr} J_a J_b, \qquad a, b = 1, ..., q,$$

repectively. It follows from A1 that the information matrix G is non-singular. The exponential connection has coefficients

$$\Gamma_{ab,d} = E\{(\partial_a s_b)s_d\} = \tfrac{1}{2}\operatorname{tr} J_{ab}J_d,$$

with respect to coordinates ξ. As another version of J_{ab}, let

$$J_{[ab]} = \Sigma_0^{1/2}[\nabla_a\partial_b\Sigma^{-1}]_0\Sigma_0^{1/2} = J_{ab} - \tfrac{1}{2}J_c g^{cd}\operatorname{tr} J_d J_{ab},$$

where g^{ab} is the (a, b) element of G^{-1}, and

$$\nabla_a\partial_b = \partial_a\partial_b - g^{cd}\Gamma_{ab,d}\partial_c$$

with Einstein's summation convention. The summation convention is used throughout this paper. For example, $J_c g^{cd}$ means $\Sigma_{c=1}^{q} J_c g^{cd}$.

Considering the Taylor expansion of h, we have

$$\rho_h(S, \Sigma) = \operatorname{tr}\left[\frac{1}{2}(S\Sigma^{-1} - I_p)^2 + \frac{1}{3!}h_3(S\Sigma^{-1} - I_p)^3 + \frac{1}{4!}h_4(S\Sigma^{-1} - I_p)^4\right] + O(\operatorname{tr}(S\Sigma^{-1} - I_p)^4). \tag{2.3}$$

It is known (Swain [15]) that

$$\tilde{\xi}^a = \sqrt{n}\,(\hat{\xi}^a - \xi_0^a) \tag{2.4}$$

is asymptotically normal and hence $O_p(1)$. The Taylor expansion of $\Sigma(\xi)^{-1}$ yields

$$\sqrt{n}\,\Sigma_0^{1/2}(\hat{\Sigma}^{-1}-\Sigma_0^{-1})\Sigma_0^{1/2}=J_b\xi^b+\frac{1}{\sqrt{n}}J_{bc}\xi^b\xi^c+O_p(n^{-1}). \tag{2.5}$$

Let

$$\Lambda=\sqrt{n}\,\Sigma_0^{-1/2}(S\hat{\Sigma}^{-1}-I_p)\Sigma_0^{1/2}. \tag{2.6}$$

Then

$$\begin{aligned}\Lambda&=V+\left(I_p+\frac{1}{\sqrt{n}}V\right)\sqrt{n}\,\Sigma_0^{1/2}(\hat{\Sigma}^{-1}-\Sigma_0^{-1})\Sigma_0^{1/2}\\&=V+J_b\xi^b+\frac{1}{\sqrt{n}}\left(VJ_b\xi^b+\frac{1}{2}J_{bc}\xi^b\xi^c\right)+O_p(n^{-1}).\end{aligned} \tag{2.7}$$

Using (2.3) and (2.7), we obtain an expansion of T_h,

$$T_h=\operatorname{tr}\left[\frac{1}{2}\Lambda^2+\frac{1}{3!\sqrt{n}}h_3\Lambda^3+\frac{1}{4!\,n}h_4\Lambda^4\right]+O_p(n^{-3/2}). \tag{2.8}$$

In order to obtain an explicit expansion of T_h, it is necessary to obtain an expansion of ξ^a. The estimates $\hat{\xi}^a$, $a=1, ..., q$, satisfy the system of equations

$$[\partial_a\rho_h(S,\Sigma)]_\wedge=0, \qquad a=1, ..., q. \tag{2.9}$$

Using (2.3) it can be seen that $\hat{\xi}^a$'s satisfy

$$\operatorname{tr}[S[\partial_a\Sigma^{-1}]_\wedge\left\{S\hat{\Sigma}^{-1}-I_p+\frac{1}{2\sqrt{n}}h_3(S\hat{\Sigma}^{-1}-I_p)^2\right\}=O_p(n^{-1}),$$

or equivalently

$$\operatorname{tr}\left[\left(I_p+\frac{1}{\sqrt{n}}V\right)\hat{J}_a\left(\Lambda+\frac{1}{2\sqrt{n}}h_3\Lambda^2\right)\right]=O_p(n^{-1}). \tag{2.10}$$

Substituting (2.7) and

$$\hat{J}_a=J_a+\frac{1}{\sqrt{n}}J_{ab}\xi^b+O_p(n^{-1}) \tag{2.11}$$

into (2.10), it is seen that ξ^a's satisfy

$$\operatorname{tr}\{J_a(V+J_b\xi^b)\}+\frac{1}{\sqrt{n}}\operatorname{tr}\left[\tilde{h}_3 J_a(V+J_b\xi^b)^2+J_a\left(\frac{1}{2}J_{bc}-J_bJ_c\right)\xi^b\xi^c\right.$$
$$\left.+J_{ab}(V+J_c\xi^c)\xi^b\right]=O_p(n^{-1}),\qquad a=1,...,q, \tag{2.12}$$

where $\tilde{h}_3=1+\frac{1}{2}h_3$. The solution of ξ^a in (2.12) can be found in an expanded form

$$\xi^a=e^a+\frac{1}{\sqrt{n}}\varepsilon^a+O_p(n^{-1}). \tag{2.13}$$

In fact, substituting (2.13) into (2.12) we obtain

$$e^a=g^{ab}s_b,\qquad \varepsilon^a=-\tfrac{1}{2}g^{ab}\operatorname{tr}(J_bM+J_{bc}e^cW), \tag{2.14}$$

where

$$W=V+J_be^b,\qquad M=\tilde{h}_3W^2-J_bJ_ce^be^c+\tfrac{1}{2}J_{bc}e^be^c. \tag{2.15}$$

Hence, from (2.7), (2.8), and (2.14) we obtain an expansion of T_h given by

$$T_h=\frac{1}{2}\operatorname{tr}W^2+\frac{1}{\sqrt{n}}T_1(V)+\frac{1}{n}T_2(V)+O_p(n^{-3/2}), \tag{2.16}$$

where

$$\begin{aligned}
T_1(V)&=\tfrac{1}{2}\operatorname{tr}(J_{[ab]}W)e^ae^b-\operatorname{tr}(J_aJ_bW)e^ae^b+\operatorname{tr}(J_aW^2)e^a+\tfrac{1}{6}h_3\operatorname{tr}W^3,\\
T_2(V)&=\tfrac{1}{24}h_4\operatorname{tr}W^4-\tfrac{1}{4}\tilde{h}_3^2g^{ab}\operatorname{tr}(J_aW^2)\operatorname{tr}(J_bW^2)\\
&\quad+\tfrac{1}{2}\tilde{h}_3\operatorname{tr}(J_{[ab]}W^2)e^ae^b-\tilde{h}_3\operatorname{tr}(J_aJ_bW^2)e^ae^b\\
&\quad+\tfrac{1}{2}\tilde{h}_3g^{ab}\operatorname{tr}(J_bJ_cJ_d)\operatorname{tr}(J_aW^2)e^ce^d+\operatorname{tr}(J_aJ_bW^2)e^ae^b\\
&\quad-\tfrac{1}{4}g^{ab}\operatorname{tr}(J_{[ac]}W)\operatorname{tr}(J_{[bd]}W)e^ce^d+\tfrac{1}{2}\operatorname{tr}(J_aWJ_bW)e^ae^b\\
&\quad+\tfrac{1}{8}\{\operatorname{tr}(J_{[ab]}J_{[cd]})-4\operatorname{tr}(J_aJ_bJ_{[cd]})+4\operatorname{tr}(J_aJ_bJ_cJ_d)\\
&\quad-2g^{ef}\operatorname{tr}(J_aJ_bJ_e)\operatorname{tr}(J_cJ_dJ_f)\}e^ae^be^ce^d.
\end{aligned}$$

3. Asymptotic Expansion of the Null Distribution of T_h

We shall obtain an asymptotic expansion of the null distribution of T_h by formally inverting an asymptotic expansion of the characteristic function of T_h . the validity of the asymptotic expansions obtained by this method has been discussed under certain regularity conditions (see, e.g., Bhattacharya ad Ghosh [6], Chandra and Ghosh [7]). Our interest is how to evaluate the characteristic function of T_h up to the order n^{-1}. We can write the characteristic function of T_h as

$$\phi(t) = E[\exp(itT_h)] = E[\mathrm{etr}(\tfrac{1}{2}\theta W^2) T(V)] + O(n^{-3/2}), \tag{3.1}$$

where $\theta = it$ and $T(V)$ is defined by

$$T(V) = 1 + n^{-1/2}\theta T_1(V) + n^{-1}\{\tfrac{1}{2}\theta^2 T_1(V)^2 + \theta T_2(V)\} \tag{3.2}$$

with the expressions T_1 and T_2 in (2.16). The pdf of V is expressed as (see, e.g., Siotani, Hayakawa, and Fujikoshi [12, p. 160])

$$f(V) = f_0(V) Q(V) + O(n^{-3/2}), \tag{3.3}$$

where $f_0(V) = \{\pi^{p(p+1)/4} 2^{p(p+3)/4}\}^{-1} \mathrm{etr}(-\tfrac{1}{4}V^2)$,

$$\begin{aligned} Q(V) &= 1 + \frac{1}{\sqrt{n}} Q_1(V) + \frac{1}{n} Q_2(V), \\ Q_1(V) &= -\frac{1}{2}(p+1)\,\mathrm{tr}\, V + \frac{1}{6}\mathrm{tr}\, V^3, \\ Q_2(V) &= \frac{1}{2}\{Q_1(V)\}^2 - \frac{1}{24}p(2p^2+3p-1) + \frac{1}{4}(p+1)\,\mathrm{tr}\, V^2 - \frac{1}{8}\mathrm{tr}\, V^4. \end{aligned} \tag{3.4}$$

Therefore, we have

$$\phi(t) = \int a_p \mathrm{etr}(-\tfrac{1}{4}V^2 + \tfrac{1}{2}\theta W^2) Q(V) T(V)\, dV + O(n^{-32}), \tag{3.5}$$

where $dV = dv_{11}\, dv_{12} \cdots dv_{p-1,p}$ and $a_p = \{\pi^{p(p+1)/4} 2^{p(p+3)/4}\}^{-1}$.

We prepare some lemmas useful for reductions of (3.5). Noting that $G^{-1} = (g^{ab})$ exists, let

$$e^a = -\tfrac{1}{2} g^{ab}\,\mathrm{tr}(J_b V), \qquad U = -J_a e^a, \qquad \text{and} \qquad W = V - U. \tag{3.6}$$

Further, let

$$M = (\mathrm{vec} * (J_1), ..., \mathrm{vec} * (J_q)), \tag{3.7}$$

where for any $p \times p$ symmetric matrix $A = (a_{ij})$,

$$\mathrm{vec} * (A) = (a_{11}/\sqrt{2}, ..., a_{pp}/\sqrt{2}, a_{12}, ..., a_{p-1,p})'.$$

Noting that $\mathrm{vec} * (A)' \,\mathrm{vec} * (B) = \tfrac{1}{2}\mathrm{tr}\, AB$, we have the following lemma.

LEMMA 3.1. *Let* $P_M = M(M'M)^{-1}M'$. *Then*

$$\mathbf{e} = (e^1, ..., e^q)' = (M'M)^{-1}M' \operatorname{vec} * (V),$$

$$\operatorname{vec} * (U) = P_M \operatorname{vec} * (V),$$

$$\operatorname{vec} * (W) = (I_{p(p+1)/2} - P_M) \operatorname{vec} * (V).$$

LEMMA 3.2. *Let* θ *be any complex number whose real part is greater than* $-\frac{1}{2}$. *Then, for a function* $h(V)$ *of* V *and a function* $g(h, W)$ *of* U *and* W,

$$\int \operatorname{etr}(-\tfrac{1}{4}V^2 + \tfrac{1}{2}\theta W^2) h(V)\, g(U, W)\, dV$$

$$= (1-2\theta)^{-r/2} \int \operatorname{etr}(-\tfrac{1}{4}V^2) h(U + (1-2\theta)^{-1/2} W) \times g(U, (1-2\theta)^{-1/2} W)\, dV, \tag{3.8}$$

where $r = \frac{1}{2}p(p+1) - q$.

Proof. We shall show that (3.8) is obtained by considering the transformation $V \to \tilde{V}$ defined by

$$\tilde{V} = U + (1-2\theta)^{1/2} W. \tag{3.9}$$

Since $\operatorname{tr} UW = 2 \operatorname{vec} * (U)' \operatorname{vec} * (W) = 0$, we have

$$\operatorname{tr} \tilde{V}^2 = \operatorname{tr} V^2 - 2\theta \operatorname{tr} W^2.$$

Using Lemma 3.1 we can write (3.9) as

$$\operatorname{vec} * (\tilde{V}) = \{P_M + (1-2\theta)^{1/2} (I_{p(p+1)/2} - P_M)\} \operatorname{vec} * (V).$$

This implies that the inverse transformation is

$$\operatorname{vec} * (V) = \{P_M + (1-2\theta)^{-1/2} (I_{p(p+1)/2} - P_M)\} \operatorname{vec} * (\tilde{V})$$

or, equivalently,

$$V = \tilde{U} + (1-2\theta)^{-1/2} \tilde{W},$$

where $\tilde{U} = \frac{1}{2} J_a g^{ab} \operatorname{tr}(J_b \tilde{V})$ and $\tilde{W} = \tilde{V} - \tilde{U}$. Therefore, the Jacobian of the transformation (3.9) is

$$|P_M + (1-2\theta)^{-1/2} (I_{p(p+1)/2} - P_M)|$$

which equals $(1-2\theta)^{-r/2}$, since the characteristic roots of P_T are one or zero and $\operatorname{rank}(P_M) = q$. Further, it holds that $U = \tilde{U}$ and $W = (1-2\theta)^{-1/2} \tilde{W}$, since $\operatorname{vec} * (\tilde{U}) = P_M \operatorname{vec}(\tilde{V}) = P_M\{P_M + (1-2\theta)^{1/2} \times (I_{p(p+1)/2} - P_M)\} \operatorname{vec} * (V) = \operatorname{vec}(U)$. These complete the proof.

LEMMA 3.3. *Let V be a $p \times p$ symmetric random matrix with pdf $f_0(V)$ in* (3.3). *Let e^a, U, and W be the random variables defined by* (3.6). *Then*

(i) $\mathbf{e} = (e^1, ..., e^q)'$ *and W are independent,*

(ii) $\mathbf{e}$ *is distributied as $N_q(\mathbf{0}, G^{-1})$,*

(iii) $\operatorname{vec} * (U)$ *and* $\operatorname{vec} * (W)$ *are independently distributed as* $N_{p(p+1)/2}(\mathbf{0}, P_M)$ *and* $N_{p(p+1)/2}(\mathbf{0}, I_{p(p+1)/2} - P_M)$, *respectively.*

Proof. The results are easily obtained by using Lemma 3.1 and the fact that $\operatorname{vec} * (V)$ is distributed as $N_{P(P+1)/2}(\mathbf{0}, I_{p(p+1)/2})$.

Using Lemmas 3.2 and 3.3, we can write the characteristic function (3.5) as

$$\phi(t) = (1-2\theta)^{-r/2} E[Q(U + (1-2\theta)^{-1/2} W) \times T(U + (1-2\theta)^{-1/2} W)] + O(n^{-3/2}). \tag{3.10}$$

Here the expectation in (3.10) is taken with respect to the distribution of U (or $\mathbf{e}$) and W given in Lemma 3.3. After calculation of these expected values, we obtain

$$\phi(t) = (1-2\theta)^{-r/2} \left\{1 + \frac{1}{n} \sum_{j=0}^{3} c_j (1-2\theta)^{-j}\right\} + O(n^{-3/2}), \tag{3.11}$$

where the coefficients c_j's are given by

$$\begin{aligned} c_0 &= \tfrac{1}{72}\{-3p(2p^2+3p-1) - 9g^{abcd}K_{abcd} + g^{abcdef}K_{abc,def}\} \\ &\quad + \tfrac{1}{16} g^{ab}g^{cd}\{4K_{[ab]cd} - K_{[ab][cd]} + 2K_{[ac][bd]}\}, \\ c_1 &= -c_0 + \tilde{h}_3^2 C - (h_4 - 6)B + \tilde{h}_3 D, \\ c_2 &= -\tilde{h}_3^2(A + C) + (h_4 - 6)B - \tilde{h}_3 D, \qquad c_3 = \tilde{h}_3^2 A, \end{aligned} \tag{3.12}$$

and the coefficients $A, ..., D$ are given by

$$\begin{aligned} A &= \tfrac{1}{72}\{6p(4p^2+9p+7) - 36q(3p+4) - 9(p^2+2p+3)\, g^{ab}K_{a,b} \\ &\quad + 6(p+1)\, g^{abcd}K_{abc,d} + 18g^{abcd}K_{abcd} - g^{abcdef}K_{abc,def}\}, \\ B &= \tfrac{1}{48}\{p(p^2+5p+5) - 4q(2p+3) - 2g^{ab}K_{a,b} + g^{abcd}K_{abcd}\}, \\ C &= \tfrac{1}{12}\{p(4p^2+9p+7) - 12q(p+1) - 3g^{ab}g^{cd}K_{acbd} \\ &\quad - 2g^{ab}g^{cd}g^{ef}K_{ace,bdf}\}, \\ D &= -\tfrac{1}{6}p(p^2+3p+4) + q(2p+3) + \tfrac{1}{2} g^{ab}K_{a,b} - \tfrac{1}{4}(p+1)\, g^{ab}g^{cd}K_{abc,d} \\ &\quad - \tfrac{1}{2} g^{abcd}K_{abcd} + \tfrac{1}{36} g^{abcdef}K_{abc,def} - \tfrac{1}{4}(p+1)\, g^{ab}K_{[ab]} \\ &\quad + \tfrac{1}{4} g^{ab}g^{cd}K_{[ab]cd}. \end{aligned} \tag{3.13}$$

Here we use the following notations:

$$g^{abcd}=g^{ab}g^{cd}+g^{ac}g^{bd}+g^{ad}g^{bc},$$
$$g^{abcdef}=g^{ab}g^{cdef}+g^{ac}g^{bdef}+g^{ad}g^{bcef}+g^{ae}g^{bcdf}+g^{af}g^{bcde},$$
$$K_{abc\cdots}=\operatorname{tr}(J_aJ_bJ_c\cdots),\qquad K_{[ab]cd}=\operatorname{tr}(J_{[ab]}J_cJ_d),$$
$$K_{abc,def}=K_{abc}K_{def},\qquad \text{and so on.} \tag{3.14}$$

The formulae needed for the expectations are given in Appendix. By inverting the characteristic function term by term, we obtain an expansion of the null distribution of T_h as in the following theorem.

THEOREM 3.1. *Let T_h be the test statistic given by* (1.3) *with a function h satisfying* A2 *and* A3. *Suppose that a given covariance structure $\Sigma=\Sigma(\xi)$ satisfies* A1. *Then under the null hypothesis H_0 the distribution of T_h can be expanded for large n as*

$$P(T_h\leqslant x)=G_r(x)+\frac{1}{n}\sum_{j=0}^{3}c_jG_{r+2j}(x)+O(n^{-3/2}), \tag{3.15}$$

where $r=p(p+1)/2-q$, $G_k(\cdot)$ is the distribution function of χ^2-variable of k degrees of freedom and the coefficients c_j's are given by (3.12).

We note that all the terms in the coefficients are scalor functions, or independent of the parametrization. Consider a reparametrization of the model Δ_0 by $\tilde{\xi}=f(\xi)$, where f is a q-dimensional C^4-diffeomorphism. Then we can show that $g^{ab}K_{a,b}=\tilde{g}^{ab}\tilde{K}_{a,b}$, $g^{ab}g^{cd}K_{[ab]cd}=\tilde{g}^{ab}\tilde{g}^{cd}\tilde{K}_{[ab]cd}$, and so on, where the derivatives included in the right sides are evaluated at $\tilde{\xi}=\tilde{\xi}_0(=f(\xi_0))$.

THEOREM 3.2. *Under the same assumptions as in Theorem* 3.1 *it holds that T_h has a Bartlett adjustment factor in a sense of* (1.4) *if and only if the condition* (1.5) *is satisfied. Further, the Bartlett adjustment factor is given by $m=1+2c_0/(rn)$, i.e., it holds that*

$$P\left(\left(1+\frac{2c_0}{rn}\right)T_h\leqslant x\right)=G_r(x)+O(n^{-3/2}), \tag{3.16}$$

if the condition (1.5) *is satisfied.*

Proof. Let $m=1+b(n)/n$, *where* $b(n)=o(1)$. *Then, the characteristic function of mT_h can be expanded as*

$$\tilde{\phi}(t)=\phi(itm)=(1-2\theta)^{-r/2}\left[1+\frac{b(n)r}{2n}\left\{(1-2\theta)^{-1}-1\right.\right.$$
$$\left.\left.+\sum_{j=0}^{3}c_j(1-2\theta)^{-j}\right\}+O(n^{-3/2})\right].$$

Therefore, it is shown that T_h has a Bartlett adjustment factor if and only if

$$c_0 = -c_1 = \tfrac{1}{2} r b(n), \qquad c_2 = 0, \qquad c_3 = 0,$$

which is equivalent to $b(n) = 2c_0/r$, $h_3 = -2$, and $h_4 = 6$. This completes the proof.

When $h(\lambda) = -\log \lambda + \lambda - 1$, condition (1.5) is satisfied. So, the likelihood ratio test has a Bartlett adjustment factor. However, it may be noted that, in general, the adjustment factor c_0 depends on unknown parameter ξ. In practical use we need to use the adjustment factor $\hat{c}_0$ obtained from c_0 by replacing ξ by $\hat{\xi}$. It is interesting to obtain the condition such that c_0 does not depend on ξ or Δ_0.

4. Apllications

In this section, we consider two types of structures: (i) Σ is a linear combination of given matrices and (ii)Σ^{-1} is a linear combination of given matrices. It may be noted that these types of structures include many important structures as special cases (see, Anderson [2, 3]). The first structure (i) is

$$\Sigma = \xi^1 G_1 + \xi^2 G_2 + \cdots + \xi^q G_q, \tag{4.1}$$

where G_a's are given $p \times p$ symmetric matrices which are linearly independent, and ξ^a's are unknown such that Σ is positive definite. For applications of the general results in the preceding section, we have to prepare only two arrays of matrices J_a and J_{ab} which are easily calculated as

$$J_a = -\Sigma_0^{-1/2} G_a \Sigma_0^{-1/2}, \qquad a = 1, ..., q, \tag{4.2}$$

and

$$J_{ab} = J_a J_b + J_b J_a, \qquad a, b = 1, ..., q. \tag{4.3}$$

The second structure (ii) is

$$\Sigma^{-1} = \xi^1 G_1 + \xi^2 G_2 + \cdots + \xi^q G_q, \tag{4.4}$$

where G_a's are given $p \times p$ symmetric matrices which are linearly independent, and ξ^a's are unknown such as to make Σ positive definite. In this case J_{ab}'s are all 0 and

$$J_a = -\Sigma_0^{1/2} G_a \Sigma_0^{1/2}, \qquad a = 1, ..., q. \tag{4.5}$$

The asymptotic expansion formula in this case is much simpler than the one in the first case, (i).

We note that the sphericity structure $\Sigma = \sigma^2 I_p$ can be regarded as special cases of both covariance structures (i) and (ii). Since we can choose an arbitrary parametrization, we use

$$\Sigma = e^{\xi} I_p. \tag{4.6}$$

For the likelihood ratio test, the coefficients c_3 and c_4 are zero, since $\tilde{h}_3 = 0$ and $h_4 = 6$. In this case we must calculate only two terms $g^{abcd}K_{abcd}$ and $g^{abcdef}K_{abc,def}$ since the terms including $J_{[ab]}$ are all zero. It is easily seen that

$$g^{abcd}K_{abcd} = 12p^{-1} \quad \text{and} \quad g^{abcdef}K_{abc,def} = 120p^{-1}. \tag{4.7}$$

Therefore, as is well known, we have $c_0 = -c_1 = -\frac{1}{24}\{2p^3 + 3p - 1 - 4p^{-1}\}$.

Appendix

Let V be a $p \times p$ symmetric random matrix normal with pdf $f_0(V)$ in (3.3). Let $\mathbf{e} = (e^1, \ldots, e^q)'$ and W be the random vector and matrix defined by (3.6). Then, it holds that for any $p \times p$ matrices A and B,

$$E[e^a e^b] = g^{ab}, \qquad E[e^a e^b e^c e^d] = g^{abcd}, \qquad E[e^a e^b e^c e^d e^e e^f] = g^{abcdef},$$

$$E[\operatorname{tr}(AW)\operatorname{tr}(BW)] = 2\operatorname{tr}(AB) - g^{ab}\operatorname{tr}(AJ_a)\operatorname{tr}(BJ_b),$$

$$E[\operatorname{tr}(AWBW)] = \operatorname{tr} A \cdot \operatorname{tr} B + \operatorname{tr}(AB') - g^{ab}\operatorname{tr}(AJ_aBJ_b),$$

$$\begin{aligned}
E[\operatorname{tr}(AW^2)\operatorname{tr}(BW^2)] = {} & 2(p+2)\operatorname{tr}(A\bar{B}) + (p^2+2p+3)\operatorname{tr} A \cdot \operatorname{tr} B \\
& - (p+1)\, g^{ab}\{\operatorname{tr} A \cdot \operatorname{tr}(BJ_aJ_b) + \operatorname{tr} B \cdot \operatorname{tr}(AJ_aJ_b)\} \\
& - 4g^{ab}\{\operatorname{tr}(AJ_a\bar{B}J_b) + \operatorname{tr}(\bar{A}\bar{B}J_aJ_b)\} \\
& + g^{abcd}\operatorname{tr}(AJ_aJ_b)\operatorname{tr}(BJ_cJ_d),
\end{aligned}$$

$$\begin{aligned}
E[\operatorname{tr}(AW)\operatorname{tr}(W^3)] = {} & 6(p+1)\operatorname{tr} A - 6g^{ab}\operatorname{tr}(\bar{A}J_aJ_b) \\
& - 3(p+1)\, g^{ab}\operatorname{tr}(AJ_a)K_b + g^{abcd}\operatorname{tr}(AJ_a)K_{bcd},
\end{aligned}$$

$$E[\operatorname{tr}(W^4)] = p(2p^2+5p+5) - 4q(2p+3) - 2g^{ab}K_{a,b} + g^{abcd}K_{abcd},$$

$$\begin{aligned}
E[\{\operatorname{tr}(W^3)\}^2] = {} & 6p(4p^2+9p+7) - 36q(3p+4) \\
& - 9(p^2+2p+3)\, g^{ab}K_{a,b} + 6(p+1)\, g^{abcd}K_{abc,d} \\
& + 18g^{abcd}K_{abcd} - g^{abcdef}K_{abc,def},
\end{aligned}$$

where $\bar{A} = \frac{1}{2}(A + A')$. The expectations are obtained by using Lemma 3.3 and the fact that vec $*$ (V) is distributed as $N_{p(p+1)/2}(\mathbf{0}, I_{p(p+1)/2})$. The calculations can be simplified by using the properties such as

$$E[\operatorname{tr} W^2 \cdot \operatorname{tr} W^2] = E[\operatorname{tr} W^2 \cdot \operatorname{tr} \tilde{W}^2 + 2 \operatorname{tr} W\tilde{W} \cdot \operatorname{tr} W\tilde{W}],$$

where $\tilde{W}$ is a $p \times p$ symmetric random matrix having the same distribution W and being independent of W.

References

[1] Amari, S. (1982). Differential geometry of curved exponential families—Curvatures and information loss. *Ann. Statist.* **10** 357–387.

[2] Anderson, T. W. (1969). Statistical inference for covariance matrices with linear structure. In *Multivariate Analysis-II* (P. R. Krishnaiah, Ed.), pp. 55–66. Academic Press, New York.

[3] Anderson, T. W. (1970). Estimation of covariance matrices which are linear combinations or whose inverses are linear combinations of given matrices. In *Essays in Probability and Statistics* (R. C. Bose and others, Eds.), pp. 1–24. Univ. of North Carolina Press, Chapel Hill.

[4] Anderson, T. W. (1984). *An Introduction to Multivariate Analysis*, 2nd ed. Wiley, New York.

[5] Banadorff-Nielsen, O. E., and Cox, D. R. (1984). Bartlett adjustments to the likelihood ratio statistic and the distribution of the maximum likelihood estimator. *J. Roy. Statist. Soc. Ser. B* **46** 483–495.

[6] Bhattacharya, R. N., and Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

[7] Chandra, T. K., and Ghosh, J. K. (1979). Valid asymptotic expansions for the likelihood ratio statistic and other perturbed chi-square valiables. *Sankhyā Ser. A* **41** 22–47.

[8] Eguchi, S. (1985). A differential geometric approach to statistical inference on the basis of contrast functions. *Hiroshima Math. J.* **15** 341–391.

[9] McCullagh, P., and Cox, D. R. (1986). Invariants and likelihood ratio statistics. *Ann. Statist.* **14** 1419–1430.

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

[11] Nagao, H. (1973). On some test criteria for covariance matrix. *Ann. Statist.* **1** 700–709.

[12] Siotani, M., Hayakawa, T., and Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis.* Amer. Sci., Columbus, Ohio.

[13] Srivastava, J. N. (1966). On testing hypotheses regarding a class of covariance structures. *Psychometrika* **31** 147–164.

[14] Sugiura, N. (1969). Asymptotic expansions of the distributions of the likelihood-ratio criteria for covariance matrix. *Ann. Math. Statist.* **40** 2051–2063.

[15] Swain, A. J. (1975). A class of factor analysis estimation procedures with common asymptotic sampling properties. *Psychometrika* **40** 315–335.