AISM (2006) 58: 101–114
DOI 10.1007/s10463-005-0010-z

Arjun K. Gupta · Jin Xu

On some tests of the covariance matrix under general conditions

Received: 7 April 2004 / Revised: 8 November 2004 / Published online: 11 February 2006

Abstract We consider the problem of testing the hypothesis about the covariance matrix of random vectors under the assumptions that the underlying distributions are nonnormal and the sample size is moderate. The asymptotic expansions of the null distributions are obtained up to $n^{-1/2}$. It is found that in most cases the null statistics are distributed as a mixture of independent chi-square random variables with degree of freedom one (up to $n^{-1/2}$) and the coefficients of the mixtures are functions of the fourth cumulants of the original random variables. We also provide a general method to approximate such distributions based on a normalization transformation.

Keywords Covariance matrix · Test statistic · Characteristic function · Canonical correlation · Multiple correlation coefficient

1 Introduction

Testing hypotheses about the covariance matrix is an important aspect of statistical inference in the multivariate analysis besides the testing hypothesis about the mean vector. In the literature, these tests have been studied extensively under the assumption of normality. See, for example, Anderson (1984), Muirhead (1982) for a complete treatment. It is found that for the hypothesis considered below the likelihood ratio tests are null robust (Kariya and Kim, 1997) within the elliptical family (Fang and Zhang, 1990). However, in the general situation, where the underlying distributions are not necessarily normal and only the existence of certain moments is assumed, the robustness property of likelihood ratio test may not hold. For instance, Kano (1995), Fujikoshi (1997, 2002a,b), Gupta et al. (2005)

A.K. Gupta (✉) J. Xu
Department of Mathematics and Statistics,
Bowling Green State University,
Bowling Green, OH 43403, USA

study the asymptotic distributions of some test statistics about the mean vector under nonnormality. It is shown that the high order nonzero cumulants, especially the skewness, have significant effect on the distributions of these statistics. Then it is important to make (finiteness) corrections to obtain more precise approximations to the true distributions. In this paper, we will derive the asymptotic expansions of some test statistics about the covariance matrix up to order of $n^{-1/2}$ and therefore extend the results of Gupta et al. (1975), Gupta (1977), Muirhead Waternaux (1980), Gupta and Tang (1984), Muirhead (1985), Gupta and Nagar (1988), Tang and Gupta (1990) to the general cases. The tests under consideration are (1) testing that a covariance matrix equals a specified matrix; (2) testing equality of k covariance matrices; (3) the sphericity test; (4) testing uncorrelation of two sets of variables; (5) testing canonical correlation coefficients.

In Sect. 2, we introduce some notation and assumptions. The main results will be presented in Sect. 3. A general method of approximating the mixture of independent chi-square random variables of degree of freedom one is introduced in Sect. 4. Some simulation results are demonstrated in Sect. 4 as well.

2 Notation and assumptions

We adopt the notation from Fujikoshi (2002a). Let $\boldsymbol{y} = (y_1, \dots, y_p)'$ be a p-variate random vector with the mean $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)'$ and the covariance matrix $\Sigma = (\sigma_{ab})$. The kth cumulant of $\boldsymbol{y}$ are denoted by $\kappa_{i_1 \dots i_k}$ (Stuart and Ord, 1987). Then the second, third and fourth order cumulants can be expressed as

$$\begin{aligned} \kappa_{ab} = \sigma_{ab}, \quad \kappa_{abc} &= E[(y_a-\mu_a)(y_b-\mu_b)(y_c-\mu_c)], \\ \kappa_{abcd} &= E[(y_a-\mu_a)(y_b-\mu_b)(y_c-\mu_c)(y_d-\mu_d)] \\ &\quad -(\sigma_{ab}\sigma_{cd}+\sigma_{ac}\sigma_{bd}+\sigma_{ad}\sigma_{bc}). \end{aligned}$$

The multivariate kurtosis introduced by Mardia (1970) is denoted as $\kappa_4^{(1)} = \sum_{a,b,c,d} \kappa_{abcd}\sigma^{ab}\sigma^{cd}$, where $\Sigma^{-1} = (\sigma^{ab})$ and the summation is over all possible combinations of indices a, b, c, d such that $1 \le a, b, c, d \le p$. And the following cumulant functions are used in the asymptotic expansions of the characteristic functions of the sample covariance matrix:

$$\begin{aligned} m_{ab;cd} &= \kappa_{abcd} + \sigma_{ac}\sigma_{bd} + \sigma_{ad}\sigma_{bc}, \\ m_{ab;cd;ef} &= \kappa_{abcdef} + \sum_{[12]} \kappa_{acef}\sigma_{bd} + \sum_{[4]} \kappa_{ace}\kappa_{bdf} + \sum_{[8]} \sigma_{ac}\sigma_{be}\sigma_{df}. \end{aligned}$$

Here $\sum_{[j]}$ means the sum for all j possible combinations. Further, if $\boldsymbol{y}$ has been standardized such that $\boldsymbol{\mu} = \boldsymbol{0}$ and $\Sigma = I_p$, where I_p is the identity matrix of dimension p, then $\kappa_4^{(1)} = \sum_{a,b} \kappa_{aabb}$.

Secondly, let $q = p(p+1)/2$; for any $p \times p$ symmetric matrix C, define a $q \times 1$ vector by $\text{vech}C = (c_{11}, c_{12}, \dots, c_{1p}, c_{22}, \cdots, c_{pp})'$. Let $\tilde{\boldsymbol{e}} = \text{vech}I_p$. We will use these notation frequently later.

Thirdly, we need the following assumptions which allow an expansion with a remainder $o(n^{-\eta/2})$.

A1: $E(\|\boldsymbol{y}\|^{2(\eta+2)}) < +\infty$;

A2: The characteristic function of $\boldsymbol{x} = (y_1, \ldots, y_p, y_1^2, y_1 y_2, \ldots, y_p^2)'$ satisfies the Cramér's condition, i.e., $\limsup_{\|\boldsymbol{t}\|\to\infty} |E(\exp(i\boldsymbol{t}'\boldsymbol{x}))| < 1$.

The restriction of the second assumption is based on one of the validity conditions (Bhattacharya and Ghosh, 1978; Hall, 1992) for asymptotic expansions. Note that the Cramér's condition is satisfied if $\boldsymbol{y}$ is a continuous type random vector.

3 Main results

3.1 Testing that a covariance matrix equals a specified matrix

Let $\boldsymbol{y}_i, i = 1, \ldots, n$ be independent identically distributed (i.i.d.) samples from $\boldsymbol{y}$. Consider testing the null hypothesis $H_0 : \Sigma = \Sigma_0$. Without loss of generality, we may assume that $\Sigma_0 = I$. Under the assumption of normality, the likelihood ratio test statistic is given by $\Lambda = (e/n)^{pn/2}\mathrm{etr}(-A/2)|A|^{n/2}$, where $A = \sum_{i=1}^n (\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\boldsymbol{y}_i - \bar{\boldsymbol{y}})'$, and $\mathrm{etr}(\cdot)$ is the abbreviation of the expression $\exp\{\mathrm{tr}(\cdot)\}$. It has been shown that the likelihood ratio test above is biased. An unbiased test can be obtained with a slight modification of Λ as

$$\Lambda_1 = e^{p(n-1)/2}\mathrm{etr}\left(-\frac{(n-1)\boldsymbol{S}}{2}\right)|\boldsymbol{S}|^{(n-1)/2},$$

where $\boldsymbol{S} = A/(n-1)$, the sample covariance. The null and nonnull distributions of Λ_1 under normality as well as the null distribution under the elliptical family have been obtained. See Muirhead (1982) for a summary. We shall extend these results to more general situations where only the existence of certain moments is assumed.

Let $\boldsymbol{V} = \sqrt{n}(\boldsymbol{S} - I)$. Under the null hypothesis, we can expand the modified test statistic as follows,

$$T_1 = -2\,\log\,\Lambda_1 = \frac{1}{2}\mathrm{tr}\boldsymbol{V}^2 - \frac{1}{3\sqrt{n}}\mathrm{tr}\boldsymbol{V}^3 + \mathrm{O}_p(n^{-1}).$$

Then the characteristic function of T_1, denoted by $C_{T_1}(t)$, can be written as

$$C_{T_1}(t) = C_1(t) + C_2(t) + \mathrm{O}(n^{-1}),$$

where $C_1(t) = \mathrm{E}\big[\exp(it\frac{1}{2}\mathrm{tr}\boldsymbol{V}^2)\big]$, $C_2(t) = \mathrm{E}\big[\exp(it\frac{1}{2}\mathrm{tr}\boldsymbol{V}^2)(-\frac{it}{3\sqrt{n}}\mathrm{tr}\boldsymbol{V}^3)\big]$.

To compute the expectations, we need some lemmas.

Lemma 3.1 *Suppose that* $\boldsymbol{y}$ *has the sixth moment. Then the characteristic function of* $\boldsymbol{V}$ *can be expanded as*

$$C_{\boldsymbol{V}}(T) = \exp\left\{\frac{i^2}{2}\sum_{a\le b,c\le d} m_{ab;cd}t_{ab}t_{cd}\right\}$$
$$\times\left[1 + \frac{i^3}{6\sqrt{n}}\sum_{a\le b,c\le d,e\le f} m_{ab;cd;ef}t_{ab}t_{cd}t_{ef} + \mathrm{o}(n^{-1/2})\right], \quad (1)$$

where $T = (\frac{1}{2}(1+\delta_{ab})t_{ab})$, $t_{ab} = t_{ba}$; δ_{ab} *is the Kronecker delta, i.e.,* $\delta_{aa} = 1$ *and* $\delta_{ab} = 0$ *for* $a \neq b$.

Proof See Fujikoshi (2002a) Lemma 3.1. □

Lemma 3.2 *Suppose that $\boldsymbol{y}$ has the sixth moment and the characteristic function of $\tilde{\boldsymbol{y}} = \mathrm{vech}(\boldsymbol{y}\boldsymbol{y}')$ satisfies*

$$\int_{R^q} |C_{\tilde{\boldsymbol{y}}}(\boldsymbol{t})|^r \mathrm{d}\boldsymbol{t} < \infty$$

for some $r \geq 1$. Then the probability density function of $\boldsymbol{V}$ can be expanded as

$$g(\boldsymbol{V}) = g_0(\boldsymbol{V})\Big[1 + \frac{1}{6\sqrt{n}} \sum_{a\leq b, c\leq d, e\leq f} m_{ab;cd;ef}\Big(\boldsymbol{e}'_{ab}\Lambda\boldsymbol{v}\boldsymbol{e}'_{cd}\Lambda\boldsymbol{v}\boldsymbol{e}'_{ef}\Lambda\boldsymbol{v} \\ -\boldsymbol{e}'_{ab}\Lambda\boldsymbol{v}\lambda_{cd;ef} - \boldsymbol{e}'_{cd}\Lambda\boldsymbol{v}\lambda_{ab;ef} - \boldsymbol{e}'_{ef}\Lambda\boldsymbol{v}\lambda_{ab;cd}\Big) + \mathrm{o}(n^{-1/2})\Big], \quad (2)$$

where

$$g_0(\boldsymbol{V}) = (2\pi)^{-\frac{q}{2}}|M|^{-\frac{1}{2}} \exp\{-\frac{1}{2}\boldsymbol{v}'M^{-1}\boldsymbol{v}\};$$

M and Λ are $q \times q$ matrices defined by $M = (m_{ab;\,cd})$ and $\Lambda = (\lambda_{ab;\,cd}) = M^{-1}$, $1 \leq a \leq b \leq p$, $1 \leq c \leq d \leq p$; $\boldsymbol{v} = \mathrm{vech}\boldsymbol{V}$, and $\boldsymbol{e}_{ab}$ is obtained from $\boldsymbol{v}$ by replacing all entries by 0 but v_{ab} by 1.

Proof Formally invert Eq. (1). □

Using the notation introduced in Lemma 3.2, we first write $\frac{1}{2}\mathrm{tr}\boldsymbol{V}^2 = \frac{1}{2}\sum_{i\leq j} v_{ij}^2 (2-\delta_{ij}) = \boldsymbol{v}'\Gamma\boldsymbol{v}$, where $\Gamma = I_q - \frac{1}{2}\mathrm{diag}\tilde{\boldsymbol{e}}$. Secondly, noticing that the second term inside the bracket of the density of $\boldsymbol{V}$, Eq. (2), is a polynomial of odd degree in $\boldsymbol{v}$, we have

$$C_1(t) = \int e^{it\boldsymbol{v}'\Gamma\boldsymbol{v}} g_0(\boldsymbol{V})\mathrm{d}\boldsymbol{v} + \mathrm{o}(n^{-1/2}) = \prod_{h=1}^{q}(1-2it\lambda_h)^{-\frac{1}{2}} + \mathrm{o}(n^{-1/2}),$$

where λ_h, $h = 1, \ldots, q$ are the characteristic roots of $M\Gamma$. Similarly, since $\mathrm{tr}\boldsymbol{V}^3$ is a polynomial of odd degree in $\boldsymbol{v}$, $C_2(t)$ is of order $\mathrm{o}(n^{-1/2})$. We summarize our results in the following proposition.

Proposition 3.1 *Under the conditions given in Lemma 3.2, the characteristic function of the test statistic T_1 under null hypothesis can be expanded as $C_{T_1}(t) = \prod_{h=1}^{q}(1-2it\lambda_h)^{-\frac{1}{2}} + \mathrm{o}(n^{-1/2})$, where λ_h are the characteristic roots of $M\Gamma$. Hence, $T_1 \overset{d}{=} \sum_{h=1}^{q}\lambda_h\chi_h^2(1)$, where $\chi_h^2(1)$ are independent chi-square random variables with one degree of freedom.*

Remark 1 Throughout the remaining part of this paper, we will use this notation $U \overset{d}{=} V$ to indicate $|P(U \leq x) - P(V \leq x)| = \mathrm{o}(n^{-1/2})$ uniformly in x.

Remark 2 It is noted that the third cumulants, hence the skewness, has no effect on this test since M is a function of the fourth cumulants of the original random vector. (See also Ito 1968).

Recently, Yanagihara et al. (2004) obtained the asymptotic expansion of the null distribution of T_1 up to the order n^{-1}. They showed that the term of the order n^{-1} is a linear mixture of series of chi-square distributions, and that the coefficients of this linear mixture are functions of high order (up to eighth) cumulants of the original population. Then in practice we will have two problems to deal with. First, how to estimate the coefficients or equivalently the high order cumulants? Second, how to approximate the mixture of chi-square distributions? It is known that the error of the empirical (moment) estimator is of order $n^{-1/2}$. Then the n^{-1} terms in the expansion will be dominated by the estimation error if the plug-in method is used. For this reason, we confine our derivations up to order $n^{-1/2}$ in this paper. We will propose a novel approach to the approximation problem in Sect. 4.

3.2 Testing equality of k covariance matrices

Suppose that there are i.i.d. samples from k (≥ 2) populations. Let $\boldsymbol{y}_j^{(i)}$ denote the jth sample from ith population, $j = 1, \ldots, n_i$, $i = 1, \ldots, k$. Assume the mean and the covariance of $\boldsymbol{y}_j^{(i)}$ are respectively $\boldsymbol{\mu}^{(i)}$ and Σ_i. Consider testing the null hypothesis $H_0 : \Sigma_1 = \Sigma_2 = \cdots = \Sigma_k$. The modified likelihood ratio statistic for the unbiased test under normality turns out to be

$$\Lambda_2 = \frac{\prod_{i=1}^{k} |\boldsymbol{S}^{(i)}|^{\frac{n_i-1}{2}}}{|\boldsymbol{S}|^{\frac{n-k}{2}}}$$

where $\boldsymbol{S}^{(i)} = \sum_{j=1}^{n_i}(\boldsymbol{y}_j^{(i)} - \bar{\boldsymbol{y}}^{(i)})(\boldsymbol{y}_j^{(i)} - \bar{\boldsymbol{y}}^{(i)})'/(n_i - 1)$, $\boldsymbol{S} = \sum_{i=1}^{k}(n_i - 1)\boldsymbol{S}^{(i)}/(n - k)$, $n = \sum_{i=1}^{k} n_i$.

The exact null distribution of Λ_2 was obtained by Gupta and Tang (1984). The asymptotic approximation can be found in Muirhead (1982, p. 309). It is seen that under the null hypothesis Λ_2 is invariant under the transformation $\Sigma^{-1/2}(\boldsymbol{y}_j^{(i)} - \boldsymbol{\mu}_i)$, assuming that Σ is the common covariance matrix. Therefore, without loss of generality, we may assume $\Sigma = I_p$ when we derive the asymptotic null distribution. Let $\boldsymbol{V}^{(i)} = \sqrt{n_i}(\boldsymbol{S}^{(i)} - I_p)$, $\rho_i = \sqrt{n_i/n}$, $\boldsymbol{V}_w = \rho_1 \boldsymbol{V}^{(1)} + \cdots + \rho_k \boldsymbol{V}^{(k)}$, $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_k)'$.

In addition to the Assumptions A1 and A2, suppose that n_i satisfies

A3: $\rho_i^{-1} = \mathrm{O}(1) \quad \text{as} \quad n \to \infty.$

Next, let $\boldsymbol{v}^{(i)} = \mathrm{vech}\boldsymbol{V}^{(i)}$, $\boldsymbol{v}_w = \mathrm{vech}\boldsymbol{V}_w$, $\boldsymbol{v}_k = (\boldsymbol{v}^{(1)'}, \boldsymbol{v}^{(2)'}, \ldots, \boldsymbol{v}^{(k)'})'$, and let $\boldsymbol{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)'$ be a $k \times 1$ vector with all entries zeros but the ith one. Then we have $\boldsymbol{v}^{(i)} = (\boldsymbol{e}_i' \otimes I_m)\boldsymbol{v}_k$ and $\boldsymbol{v}_w = (\boldsymbol{\rho}' \otimes I_m)\boldsymbol{v}_k$. The notation '$\otimes$' is the usual Kronecker product.

Lemma 3.3 *Under the conditions given in Lemma 3.2, the asymptotic density function of $\boldsymbol{v}_k$ can be obtained as follows.*

$$g(\boldsymbol{v}_k) = g_0(\boldsymbol{v}_k)\Bigg[1 + \frac{1}{6\sqrt{n}}\sum_{i=1}^{k}\sum_{a\le b, c\le d, e\le f}\frac{1}{\rho_i}m^{(i)}_{ab;cd;ef}$$
$$\times\Big(\boldsymbol{e}'_{ab}\Lambda^{(i)}\boldsymbol{v}^{(i)}\cdot\boldsymbol{e}'_{cd}\Lambda^{(i)}\boldsymbol{v}^{(i)}\cdot\boldsymbol{e}'_{ef}\Lambda^{(i)}\boldsymbol{v}^{(i)}$$
$$-\boldsymbol{e}'_{ab}\Lambda^{(i)}\boldsymbol{v}^{(i)}\lambda^{(i)}_{cd;ef} - \boldsymbol{e}'_{cd}\Lambda^{(i)}\boldsymbol{v}^{(i)}\lambda^{(i)}_{ab;ef} - \boldsymbol{e}'_{ef}\Lambda^{(i)}\boldsymbol{v}^{(i)}\lambda^{(i)}_{ab;cd}\Big)$$
$$+\mathrm{o}(n^{-1/2})\Bigg],$$

where

$$g_0(\boldsymbol{v}_k) = (2\pi)^{-\frac{kq}{2}}\prod_{i=1}^{k}|M^{(i)}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\boldsymbol{v}_k{}'M_k^{-1}\boldsymbol{v}_k\right\},$$

and $M_k = \mathrm{diag}\{M^{(1)}, \ldots, M^{(k)}\}$. *The superscript* (i) *indicates the quantity is from the* i*th population.*

Proof Note that $g(\boldsymbol{v}_k) = \prod_{i=1}^{k} g(\boldsymbol{V}^{(i)})$. Then make use of Lemma 3.2 and simplify. □

Return to the test statistic. Define $T_2 = -2\log\Lambda_2$. It can be expanded as

$$T_2 = \frac{1}{2}\sum_{i=1}^{k}\mathrm{tr}\boldsymbol{V}^{(i)^2} - \frac{1}{2}\mathrm{tr}\boldsymbol{V}_w^2 - \frac{1}{3\sqrt{n}}\sum_{i=1}^{k}\mathrm{tr}\frac{\boldsymbol{V}^{(i)^3}}{\rho_i} + \frac{1}{3\sqrt{n}}\mathrm{tr}\boldsymbol{V}_w^3 + \mathrm{o}_p(n^{-1/2}).$$

Using the notation introduced before, we can write $\frac{1}{2}\sum_{i=1}^{k}\mathrm{tr}\boldsymbol{V}^{(i)^2} - \frac{1}{2}\mathrm{tr}\boldsymbol{V}_w^2 = \boldsymbol{v}_k{}'\Omega_k\boldsymbol{v}_k$, where $\Omega_k = \Omega\otimes\Gamma$, $\Omega = I_k - \boldsymbol{\rho}\boldsymbol{\rho}'$. It is noted that the rank of Ω_k is $(k-1)q$. Finally proceeding in the same way as in the derivation of the asymptotic characteristic function of T_1, we obtain the representation of T_2.

Proposition 3.2 *Suppose that the conditions of Lemma 3.2 hold for each population, then the characteristic function of the test statistic* T_2 *under the null hypothesis can be expanded as* $C_{T_2}(t) = \prod_{h=1}^{(k-1)q}(1 - 2it\lambda_h)^{-1/2} + \mathrm{o}(n^{-1/2})$. *Hence,* $T_2 \stackrel{d}{=} \sum_{h=1}^{(k-1)q}\lambda_h\chi_h^2(1)$, *where* λ_h, $i = 1, 2, \ldots, (k-1)q$ *are the nonzero characteristic roots of* $M_k\Omega_k$.

So again the asymptotic distribution of the test statistic T_2 is a mixture of independent chi-square random variables. However, the approximation is actually up to $n^{-1/2}$ unlike what is stated in Muirhead (1982).

3.3 The sphericity test

The sphericity hypothesis is termed as $H_0 : \Sigma = \lambda I_p$, where $\lambda(>0)$ is unspecified. Given an i.i.d. sample of size n, the (unbiased) likelihood ratio test statistic under normality is given by

$$\Lambda_3 = \frac{|\boldsymbol{S}|}{(\frac{1}{p}\mathrm{tr}\boldsymbol{S})^p}.$$

Its asymptotic distribution under normality can be found, for example, in Muirhead (1982, p. 344). The locally best invariant (LBI) test is obtained by Sugiura (1972) and shown to be null and nonnull robust with respect to the elliptical family (Kariya and Kim, 1997). We shall find the asymptotic distribution of the test statistic, $T_3 = -(n-1)\log\,\Lambda_3$, under general condition up to $n^{-1/2}$.

let $\boldsymbol{V} = \sqrt{n}(\frac{\boldsymbol{S}}{\lambda} - I)$, then the asymptotic characteristic function and the density function of $\boldsymbol{V}$ can be shown identical to Eqs. (1) and (2), respectively, provided the assumptions of Lemma 3.2 hold. Next we expand T_3 as

$$T_3 = \frac{1}{2}\mathrm{tr}\boldsymbol{V}^2 - \frac{1}{2p}(\mathrm{tr}\boldsymbol{V})^2 + \frac{1}{\sqrt{n}}\left[\frac{1}{3p^2}(\mathrm{tr}\boldsymbol{V})^3 - \frac{1}{3}\mathrm{tr}\boldsymbol{V}^3\right] + \mathrm{o}_p(n^{-1/2}),$$

and write $\frac{1}{2}\mathrm{tr}\boldsymbol{V}^2 - \frac{1}{2p}(\mathrm{tr}\boldsymbol{V})^2 = \boldsymbol{v}'\Gamma_1\boldsymbol{v}$, where $\boldsymbol{v} = \mathrm{vech}\boldsymbol{V}$, $\Gamma_1 = I_q - \frac{1}{2}\mathrm{diag}\tilde{\boldsymbol{e}} - \frac{1}{2p}\tilde{\boldsymbol{e}}\tilde{\boldsymbol{e}}'$. It is noted that Γ_1 and Γ are slightly different and the rank of Γ_1 is $q-1$. Finally, we get

Proposition 3.3 *Under the conditions of Lemma 3.2, the characteristic function of the test statistic T_3 under the null hypothesis can be expanded as $C_{T_3}(t) = \prod_{h=1}^{q-1}(1-2it\lambda_h)^{-\frac{1}{2}} + \mathrm{o}(n^{-1/2})$. Hence, $T_3 \overset{d}{=} \sum_{h=1}^{q-1}\lambda_h\chi_h^2(1)$, where λ_h, $h = 1, 2, \ldots, q-1$ are the nonzero characteristic roots of $M\Gamma_1$.*

3.4 Testing uncorrelation of two sets of variables

Partition $\boldsymbol{y} = (\boldsymbol{y}_1', \boldsymbol{y}_2')'$ with $\boldsymbol{y}_1 : p_1 \times 1$, $\boldsymbol{y}_2 : p_2 \times 1$ $(p_1 \le p_2)$ and the covariance matrix correspondingly as $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{pmatrix}$. Consider the null hypothesis $H_0 : \Sigma_{12} = 0$. Under normality, H_0 is equivalent to testing the independence between $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$. The likelihood ratio test statistic is found to be

$$\Lambda_4 = |I - \boldsymbol{S}_{21}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}|^{n/2},$$

where $\boldsymbol{S}$ is the sample covariance matrix partitioned in the same way as Σ (Muirhead, 1982, p. 542). Note that Λ_4 is invariant under the transformation $\boldsymbol{y} \to \begin{pmatrix} B_{11} & \boldsymbol{0} \\ \boldsymbol{0} & B_{22} \end{pmatrix}(\boldsymbol{y} - \boldsymbol{c})$, where B_{ii} $(p_i \times p_i)$, $i = 1, 2$, are nonsingular real matrices and $\boldsymbol{c}$ is an arbitrary constant vector in $\mathbb{R}^p$. Without loss of generality, we may assume that $\boldsymbol{\mu} = \boldsymbol{0}$ and $\Sigma = \begin{pmatrix} I_{p_1} & \Sigma_{12} \\ \Sigma_{12}' & I_{p_2} \end{pmatrix}$ when we derive the asymptotic distribution of the test statistic, $T_4 = -2\log\Lambda_4$.

Let $\boldsymbol{V} = \sqrt{n}(\boldsymbol{S} - \Sigma)$. Under the null hypothesis, the asymptotic distribution of $\boldsymbol{V}$ is given by Eq. (2). On the other hand, we can expand

$$T_4 = \mathrm{tr}\boldsymbol{V}_{12}\boldsymbol{V}_{21} - \frac{1}{3\sqrt{n}}(\mathrm{tr}\boldsymbol{V}^3 - \mathrm{tr}\boldsymbol{V}_{11}^3 - \mathrm{tr}\boldsymbol{V}_{22}^3) + \mathrm{o}_p(n^{-1/2}),$$

and write $\mathrm{tr}\boldsymbol{V}_{12}\boldsymbol{V}_{21} = \boldsymbol{v}'\Gamma_4\boldsymbol{v}$, where $\boldsymbol{v} = \mathrm{vech}\boldsymbol{V}$, $\Gamma_4 = \sum_{1\le i\le p_1, p_1+1\le j\le p}\boldsymbol{e}_{ij}\boldsymbol{e}_{ij}'$, and $\boldsymbol{e}_{ij}$ is defined in Lemma 3.2. Note that Γ_4 is a diagonal matrix with $[(a-1)(p-a/2+1)+b]^{\text{th}}$ diagonal element equal 1 for $1 \le a \le p_1$, $p_1+1 \le b \le p$ and others equal 0. (Obviously the rank of Γ_4 is $p_1 p_2$.) Then Let M_4^* be the nonzero block from $M\Gamma_4$, i.e., $M_4^* = (m_{ab,cd})$ with $1 \le a, c \le p_1,\ p_1 + 1 \le b, d \le p$. Finally we get

Proposition 3.4 *Under the conditions of Lemma 3.2, the characteristic function of the test statistic T_4 under the null hypothesis can be expanded as $C_{T_4}(t) = |I_{p_1 p_2} - 2itM_4^*|^{-1/2} + o(n^{-1/2})$. Hence, $T_4 \stackrel{d}{=} \sum_{h=1}^{p_1 p_2} \lambda_h \chi_h^2(1)$, where λ_h, $h = 1, 2, \dots, p_1 p_2$ are the characteristic roots of M_4^*.*

3.5 Testing canonical correlation coefficients

Let the random vector $\boldsymbol{y}$ and its mean and covariance matrix be partitioned in the same fashion as in Sect. 3.4. The canonical correlation is the measure of correlation structure between $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ after being reduced to the simplest form possible by means of linear transformations of $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$. Let $\rho_1^2, \dots, \rho_{p_1}^2$ $(1 \ge \rho_1^2 \ge \cdots \ge \rho_{p_1}^2 \ge 0)$ be the characteristic roots of the matrix $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. Then their positive square roots with $1 \ge \rho_1 \ge \cdots \ge \rho_{p_1} \ge 0$ are called the population canonical correlation coefficients. Their sample versions are denoted by $r_1, r_2, \dots, r_{p_1}$. [Anderson (1984, Chapter 12) or Muirhead (1982, Chapter 11)].

The hypothesis of interest is $H_{p_1-k} : \rho_{k+1} = \cdots = \rho_{p_1} = 0$ $(\rho_k > 0)$ for $k = 0, 1, \dots, p_1 - 1$. When $k = 0$, it is equivalent to testing $H_0 : \Sigma_{12} = \mathbf{0}$ which has been treated in the previous section. The motivation for this test is to reduce the dimensionality between two sets of variables. In other words, we try to determine the number of useful canonical variables which represent all the information. The likelihood ratio statistic for H_{p_1-k} under normality is given by

$$\Lambda_{p_1-k} = \prod_{i=k+1}^{p_1} (1 - r_i^2).$$

The asymptotic distribution of the test statistic, defined by $T_{p_1-k} = -n \log \Lambda_{p_1-k}$, has been investigated by Bartlett (1938, 1947), Lawley (1959), Fujikoshi (1976), Glyun and Muirhead (1978), among others. Muirhead Waternaux (1980) derive the asymptotic distribution of r_i^2 and the asymptotic null distribution of T_{p_1-k} as well. Their result is summarized in the following proposition via our notation.

Proposition 3.5 *Assume that the population has finite fourth order cumulants, then when the null hypothesis H_{p_1-k} is true, the limiting characteristic function of the test statistic T_{p_1-k} is $C_{T_{p_1-k}}(t) = |I_{(p_1-k)(p_2-k)} - 2itM_{p_1-k}^*|^{-1/2}$, where $M_{p_1-k}^* = (m_{ab,cd})$ with $k + 1 \le a, c \le p_1,\ p_1 + k + 1 \le b, d \le p$. Hence, $T_{p_1-k} \xrightarrow{L} \sum_{h=1}^{(p_1-k)(p_2-k)} \lambda_h \chi_h^2(1)$, where λ_h, $h = 1, 2, \dots, (p_1 - k)(p_2 - k)$ are the characteristic roots of $M_{p_1-k}^*$. (Here $\xrightarrow{L}$ stands for convergence in distribution.)*

Remark 3 Although proposition 3.5 obtains the limiting null distribution of T_{p_1-k}, we conjecture that it is also the asymptotic distribution up to order $O(n^{-1/2})$ as for the other test statistics considered earlier. Further investigation is needed. Nevertheless, if the population is from an elliptical distribution, then it is found that both the limiting null distribution and the limiting nonnull distribution of T_{p_1-k} possess a neat form (Muirhead Waternaux, 1980).

As an important special case when $p_1 = 1$ and $p_2 = p - 1$, the canonical correlation coefficient is reduced to the multiple correlation coefficient. Since it has its own interest, we will study its asymptotic distribution.

First write $\Sigma = \begin{pmatrix} \sigma_{11} & \boldsymbol{\sigma}'_{12} \\ \boldsymbol{\sigma}_{12} & \Sigma_{22} \end{pmatrix}$, $\boldsymbol{\sigma}_{12} : (p-1) \times 1$. The multiple correlation coefficient between y_1 and $\boldsymbol{y}_2$ is defined to be $\bar{R} = (\boldsymbol{\sigma}'_{12}\Sigma_{22}^{-1}\boldsymbol{\sigma}_{12}/\sigma_{11})^{1/2}$. Because of invariance, it can be assumed without loss of generality that $\Sigma = \begin{pmatrix} 1 & P' \\ P & I_{p-1} \end{pmatrix}$, where $P = (\bar{R}, 0, \cdots, 0)'$. Let $\boldsymbol{S}$ be the sample covariance matrix from a sample of size n and denote the sample multiple correlation coefficient by $R^2 = \boldsymbol{s}'_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{s}_{12}/s_{11}$. The asymptotic distribution of R^2 under normality can be found in Muirhead (1982, p. 172). Next we will derive the asymptotic distribution R^2 under nonnormality.

Let $u_{11} = \sqrt{n}(s_{11} - 1)/(1-\bar{R}^2)$, $\boldsymbol{u}_{12} = \sqrt{n}(1-\bar{R}^2)^{-1/2}(I_{p-1}-PP')^{-1/2}(\boldsymbol{s}_{12}-P)$ and $\boldsymbol{U}_{22} = \sqrt{n}(I_{p-1} - PP')^{-1/2}(\boldsymbol{S}_{22} - I_{p-1})(I_{p-1} - PP')^{-1/2}$. Then we can expand

$$R^2 = \bar{R}^2 + \frac{1}{\sqrt{n}}\bar{R}(1 - \bar{R}^2)(2u_{12} - \bar{R}u_{11} - \bar{R}u_{22}) + \mathrm{O}_p(n^{-1}),$$

where u_{12} and u_{22} are the first component of $\boldsymbol{u}_{12}$ and the (1,1)th component of $\boldsymbol{U}_{22}$, respectively.

When $\bar{R} \neq 0, 1$, define $T_5 = \sqrt{n}(R^2 - \bar{R}^2)/2\bar{R}(1 - \bar{R}^2)$. Then the asymptotic distribution of T_5 can be obtained along the same course as in Muirhead (1982, p. 179) where he derives it for the elliptical family. At last we can get $T_5 \xrightarrow{L} N(0, \sigma_R^2)$, where $\sigma_R^2 = \boldsymbol{\alpha}' M_5^* \boldsymbol{\alpha}/(1 - \bar{R}^2)^2$, $\boldsymbol{\alpha} = (-\bar{R}/2,\ 1,\ -\bar{R})'$ and $M_5^* = (m_{ab;\,cd})$ with $1 \le a \le b \le 2,\ 1 \le c \le d \le 2$.

For testing the hypothesis $\bar{R} = 0$, we define $T_6 = nR^2$. Under the null hypothesis, it can be expanded as

$$T_6 = \boldsymbol{u}'_{12}\boldsymbol{u}_{12} + \mathrm{O}_p(n^{-1/2}).$$

Note that $\boldsymbol{u}_{12} = \sqrt{n}\boldsymbol{s}_{12}$ after replacing $\bar{R}$ by 0 in the definition. Then under nonnormality we obtain the asymptotic null distribution of T_6 as $\sum_{h=1}^{p-1} \lambda_h \chi_h^2(1)$, where $\lambda_h, h = 1, 2, \ldots, p-1$ are the characteristic roots of $M_6^* = (m_{1a;\,1b})$ with $2 \le a \le b \le p$.

A special function of $\bar{R}$ is $\bar{\theta} = \bar{R}^2/(1 - \bar{R}^2)$ (Muirhead, 1985) with the sample version defined as $\theta = R^2/(1 - R^2)$. It can be verified that $\sqrt{n}(\theta - \bar{\theta})/2\bar{\theta}$ and $n\theta$ have the same asymptotic distributions as T_5 and T_6, respectively, provided the same assumptions hold.

3.6 Additional remarks

We have studied the asymptotic expansions of the null distributions of some statistics for testing the hypothesis about the covariance matrix. It is found that in most cases the test statistics are distributed, up to $\mathrm{O}(n^{-1/2})$, as quadratic forms in normal variables, or equivalently as mixtures of independent chi-square variables with degree of freedom one. It is also found that the coefficients of the mixtures are functions of the fourth cumulants of the original random variables. They are expressed in terms of the characteristic roots of certain matrix in the corresponding quadratic forms. (In the following section, we will see that only the trace of some powers of the matrix are needed to obtain the approximation.) These facts imply that the likelihood ratio test statistics obtained under normality are robust to the departure of the skewness to a certain level ($\mathrm{O}(n^{-1/2})$) but are sensitive to the change of the kurtosis unlike those for testing about the mean vectors.

To close this section, we would conjuncture there be some new test statistics (criteria) which will accommodate the effect of nonnormality so that their distributions will be the same as under normality up to a higher order (n^{-1} at least). Further investigation is needed.

4 Approximation and simulation

As seen previously most test statistics under null hypothesis are distributed as a mixture of independent chi-square random variables with degree of freedom one. Many authors have studied both the exact and the approximate distributions of this type of statistics, e.g., Jensen and Solomon (1972), David (1977), Solomon and Stephens (1977), Konishi at al. (1988) among others. For a review, see Johnson and Kotz (1970, Chapter 29), Mathai and Provost (1992). Some recent developments include Kuonen (1999), Lu and King (2002), etc. A simple approximation to this type of mixture is the Satterthwaite method, which approximates the mixture distribution by a distribution $\beta\chi^2(\nu)$ such that the first two moments agree. In our cases, the parameters β and ν can be easily determined (up to $n^{-1/2}$) as $\beta = \mathrm{tr}(M^*)^2/\mathrm{tr}M^*$ and $\nu = (\mathrm{tr}M^*)^2/\mathrm{tr}(M^*)^2$, where M^* is the matrix of the corresponding quadratic form.

The paper by Konishi at al. (1988) derives higher order asymptotic expansions of the distributions to yield satisfactory accuracy. Based on this, we will further devise a transformation that converts the null test statistic to a standard normal variable with a controlled error. It thereby provides a more applicable way to test the hypothesis in practice. Some Monte Carlo simulations are presented in the second subsection.

4.1 Approximation

Let T be a generic statistic such that $T = \sum_{j=1}^{k} \lambda_j \chi_j^2(1)$. Let $m_r = \sum_{j=1}^{k} \lambda_j^r$, $r = 1, 2, \ldots$ Assume that $w_s = m_s/m_1 = \mathrm{O}(1)$, for $s = 2, 3, \ldots$ Then an approximation of the distribution of T via transformation is obtained by Konishi at al. (1988) as follows.

Theorem 4.1 *Let* $W = \sqrt{m_1}[(\frac{T}{m_1})^h - 1 - \frac{1}{m_1}h(h-1)w_2]/\sqrt{2h^2 w_2}$. *Then the asymptotic expansion of the distribution of* W *is given by*

$$P(W \leq x) = \Phi(x) - m_1^{-1} a_2(x)\phi(x) + \mathrm{O}(m_1^{-3/2}),$$

where $h = 1 - \frac{2m_1 m_3}{3m_2^2}$, $\Phi(x)$ *and* $\phi(x)$ *are the standard normal distribution function and density function, respectively, and*

$$a_2(x) = w_2^{-3}\left[w_3\left(-\frac{2}{3}w_3 + \frac{2}{3}w_2^2\right)H_1(x) + \left(\frac{1}{2}w_4 w_2 - \frac{20}{27}w_3^2 + \frac{2}{9}w_3 w_2^2\right)H_3(x)\right].$$

Here $H_j(x)$ *is the Hermite polynomial of degree* j.

Remark 4 The theorem obtained by Konishi at al. (1988) actually gives the asymptotic expansion for a general quadratic form which is equivalent to T with noncentral chi-square variables. And the expansion is derived up to order m_1^{-3}. Here we ignore the terms higher than m_1^{-1} because the remaining terms are quite involved. More importantly it is because the magnitude of the remaining terms would be comparable to the errors brought into the m_1^{-1} term when we replace the coefficients by their estimators.

Next we shall apply the normalization transformation derived by Xu and Gupta (2005).

Corollary 4.1 *Under the assumptions of Theorem 4.1, we have*

$$|W| - \frac{1}{m_1}a_2(|W|) \stackrel{d}{=} |Z|, \tag{3}$$

where $Z \sim N(0, 1)$. *Here the equality in distribution is up to order* $\mathrm{O}(m_1^{-1})$.

Finally, the null hypothesis is rejected if the LHS of Eq. (3) is greater than $z_{\alpha/2}$, where $z_{\alpha/2}$ is the upper $\alpha/2$ percentile of the standard normal distribution and α is the nominal size of the test.

4.2 Simulation

Consider the null hypothesis $H_0 : \Sigma = \Sigma_0$ as an example to see the performance of the asymptotic expansion and the approximation discussed above. The other cases can be studied in a similar fashion.

To take nonnormality into account, we consider the following model. Let $\boldsymbol{y}$ be a four dimensional random vector with independent components as follows:

(a) y_1: standard normal distribution;
(b) y_2: standardized skew t distribution with degrees of freedom 6 and 4 (Jones and Faddy, 2003);
(c) y_3: standard Laplace distribution;
(d) y_4: standardized χ^2 distribution with degrees of freedom 12.

(Then the kurtosis of the chosen model is $\kappa_4^{(1)} = 5.7228$.) The Monte Carlo studies can be summarized as follows. First, generate a sample of size n from the model. Second, compute T_1, W (defined in Theorem 4.1), and the normalized test statistic from Eq. (3) replace the coefficients m_i and w_i by their estimators, i.e., we estimate the matrix M by its sample version. Third, determine the empirical size of the test after 10,000 repetitions. Meanwhile, we shall compute the empirical size of the two-sided test by the Satterthwaite method.

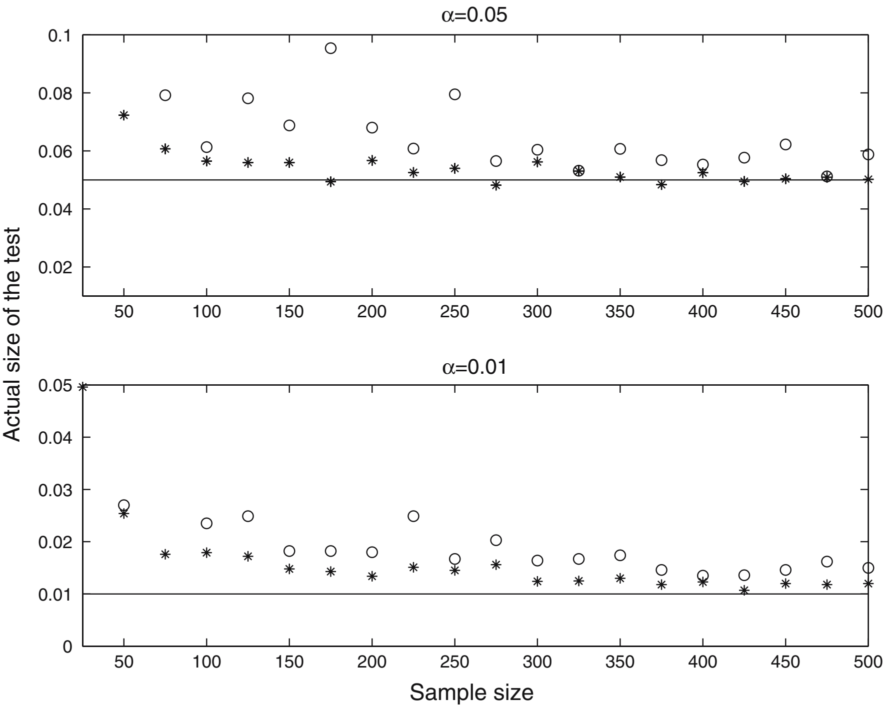

Fig. 1 Actual sizes of the normalized test statistic (denoted by *asterisk*) in comparison to those by the Satterthwaite method (denoted by *circle*)

The results are displayed in Fig. 1. From that, it can be seen that our approximation yields satisfactory results compared with the Satterthwaite approximation even for the moderate sample sizes.

Acknowledgements The second author is supported by a non-service fellowship at Bowling Green State University.

References

Anderson, T. W. (1984). *An Introduction to Multivariate Analysis* (2nd ed.). New York: John Wiley & Sons.

Bartlett, M. S. (1938). Further aspects of the theory of multiple regression. *Mathematical Proceedings of the Cambridge Philosophical Society*, *34*, 33–40.

Bartlett, M. S. (1947). Multivariate analysis. *Journal of the Royal Statistical Society* (Suppl.), *9*, 176–190.

Bhattacharya, R. N., & Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion. *Annals of Statistics*, *6*, 434–451; Corrigendum, ibid. *8* (1980).

David, A. W. (1977). A differential equation approach to linear combinations of independent chi-squares. *Journal of the American Statistical Association*, *72*, 212–214.

Fang, K., & Zhang, Y. (1990). *Generalized Multivariate Analysis*. New York: Springer-Verlag.

Fujikoshi, Y. (1976). Asymptotic expansions for the distributions of some multivariate tests. In P. R. Krishinaiah (Ed.). *Multivariate Analysis,* Vol. 4, (pp. 55–71). Amsterdam: North Holland.

Fujikoshi, Y. (1997). An asymptotic expansion of the distribution of Hotelling's T^2-statistic. *Journal of Multivariate Analysis*, *61*, 187–193.
Fujikoshi, Y. (2002a). Asymptotic expansions for the distributions for multivariate basic statistics and one-way MANOVA tests under nonnormality. *Journal of Statistical Planning and Inference*, *108*, 263–282.
Fujikoshi, Y. (2002b). Some recent results on asymptotic expansions of multivariate test statistics for mean vectors under non-normality. *Calcutta Statistical Association Bulletin*, *52* (Special 4th Triennial Proceedings Volume), Nos. 205–208, 1–46.
Gupta, A. K. (1977). On the distribution of sphericity test criterion in the multivariate Gaussian distribution. *Australian & New Zealand Journal of Statistics*, *19*, 202–205.
Gupta, A. K., Chattopadhyay, A. K., & Krishnaiah, P. R. (1975). Asymptotic distributions of the determinants of some random matrices. *Communications in Statistics*, *4*, 33–47.
Gupta, A. K., & Nagar, D. K. (1988). Asymptotic expansion of the nonnull distribution of likelihood ratio statistic for testing multisample sphericity. *Communications in Statistics. Theory and Methods*, *17*, 3145–3155.
Gupta, A. K., & Tang, J. (1984). Distribution of likelihood ratio statistic for testing equality of covariance matrices of multivariate Gaussian models. *Biometrika*, *71*, 555–559.
Gupta, A. K., Xu, J., & Fujikoshi, Y. (2005). An asymptotic expansion of the distribution of Rao's U-statistic under a general condition. *Journal of Multivariate Analysis* (to appear).
Hall, P. (1992). *The Bootstrap and Edgeworth Expansions*. New York: Springer-Verlag.
Ito, K. (1968). On the effect of heteroscedasticity and nonnormality upon some multivariate test procedures. In P. R. Krishinaiah (Ed.). *Multivariate Analysis,* Vol. *2*, (pp. 87–120). New York: Academic Press.
Jensen, D. R., & Solomon, H. (1972). A Gaussian approximation to the distribution of a definite quadratic form. *Journal of the American Statistical Association*, *67*, 898–902.
Johnson, N. I., & Kotz, S. (1970). *Continuous Univariate Distributions*, Vol. 2. New York: John Wiley & Sons.
Jones, M. C., & Faddy, M. J. (2003). A skew extension of the t-distribution, with applications. *Journal of the Royal Statistical Society. Series B*, *65*, Part 1, 159–174.
Kano, Y. (1995). An asymptotic expansion of the distribution of Hotellings T^2-statistic under general condition. *American Journal of Mathematical and Management Sciences*, *15*, 317–341.
Kariya, T., & Kim, P. (1997). Finite sample robustness of tests: an overview. In G. S. Maddala & C. R. Rao (Eds). *Handbook of Statistics, Vol. 15, Robust Inference* (pp. 645–660).
Konishi, S., Niki, N., & Gupta, A. K. (1988). Asymptotic expansions for the distribution of quadratic forms in normal variables. *Annals of the Institute of Statistical Mathematics*, *40*, 279–296.
Kuonen, D. (1999). Saddlepoint approximations for distributions of quadratic forms in normal variables. *Biometrika*, *86*, 929–935.
Lawley, D. N. (1959). Tests of significance in canonical analysis. *Biometrika*, *46*, 59–66.
Lu, Z. -H., & King, M. (2002). Improving the numerical technique for computing the accumulated distribution of a quadratic form in normal variables. *Econometric Reviews*, *21*, 149–165.
Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*, 519–530.
Mathai, A. M., & Provost, S. B. (1992). *Quadratic forms in random variables*. New York: Marcel Dekker.
Muirhead, R. J., & Waternaux, C. M. (1980). Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations. *Biometrika*, *67*, 31–43.
Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. New York: John Wiley & Sons.
Muirhead, R. J. (1985). Estimating a particular function of the multiple correlation coefficient. *Journal of the American Statistical Association*, *80*, 923–925.
Rao, C. R. (1973). *Linear Statistical Inference and Its Applications* (2nd Ed.). New York: Wiley.
Siotani, M., Hayakawa, T., & Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook*. Columbus, OH: American Sciences Press.
Solomon, H., & Stephens, M. A. (1977). Distribution of a sum of weighted chi-square variables. *Journal of the American Statistical Association*, *72*, 881–885.
Stuart, A., & Ord, J. K. (1987). *Kendall's Advanced Theory of Statistics*, Vol. 1 (5th Ed.). New York: Oxford University Press.

Sugiura, N. (1972). Locally best invariant test for sphericity and the limiting distribution. *Annals of Mathematical Statistics*, *43*, 1312–1316.
Tang, J., & Gupta, A. K. (1990). Testing independence of several groups of variables. *Journal of Statistical Computation and Simulation*, *37*, 27–35.
Xu, J., & Gupta, A. K. (2005). Confidence intervals for the mean value of response function in generalized linear models. *Statistica Sinica*, *15*, 1081–1096.
Yanagihara, H., Matsumoto, C., & Tonda, T. (2004). Asymptotic expansion of the null distribution of the modified normal likelihood ratio criterion for testing $\Sigma = \Sigma_0$ under nonnormality. *Hiroshima Mathematical Journal*, *34*, 81–100.