Available at WWW.MATHEMATICSWEB.ORG
POWERED BY SCIENCE DIRECT
Journal of Multivariate Analysis 84 (2003) 284–298
Journal of Multivariate Analysis
http://www.elsevier.com/locate/jmva

Conditional tests for elliptical symmetry

Li-Xing Zhu[a,b,*] and Georg Neuhaus[c]

[a] Department of Statistics and Actuarial Sciences, The University of Hong Kong, Pokfulam Road, Hong Kong
[b] Academy of Mathematics and System Science, China
[c] Center for Statistics and Stochastics, Hamburg University, Germany

Received 9 March 2000

Abstract

In this paper, we suggest the conditional test procedures for testing elliptical symmetry of multivariate distribution. The conditional tests are exactly valid if the symmetric center and the shape matrix are given and are asymptotically valid if they are unknowns to be estimated. The equivalence, in the large sample sense, between the conditional tests and their unconditional counterparts is established. The power behavior of the tests under global as well as local alternatives is investigated theoretically. A small simulation study is performed.

AMS 1991 subject classifications: primary 62G10; secondary 62G20; 62G09

Keywords: Conditional test; Elliptical symmetry; Empirical process; Validity of test

1. Introduction

Let X be a d-dimensional random vector. The distribution of X is said to be elliptically symmetric with a center $\mu \in R^d$ and a matrix A if for all orthogonal $d \times d$ matrix H the distributions of $HA(X - \mu)$ are identical. Throughout this paper, we assume that the covariance matrix Σ of X is positive definite. In this case, A is equal to $\Sigma^{-1/2}$. We call Σ the shape matrix.

In practical use, the elliptically symmetric distribution (the elliptical distribution for short) has received considerable attention. The elliptical distribution possesses

*Corresponding author. Department of Statistics and Actuarial Sciences, The University of Hong Kong, Pokfulam Road, Hong Kong.
E-mail address: lzhu@hku.hk (L.-X. Zhu).

0047-259X/03/$ - see front matter
doi:10.1016/S0047-259X(02)00036-2

many nice properties which are analogous to those of multivariate normal distribution. Hence, if one has known that the variable is elliptically symmetrically distributed, some tools for classical multivariate analysis, as Friedman [10] pointed out, may be still applicable for analyzing the data. Additionally, the dimensional reduction techniques have quickly been developed in recent years for overcoming the curse of dimensionality in data analysis. One of them is sliced inverse regression (SIR) (see [13]). The most important subclass of distributions satisfying the designed condition of SIR is just the one of the elliptical distributions. Consequently, testing elliptical symmetry is important and relevant in multivariate analysis. The hypothesis to be tested is that

$$\mathrm{H}_0: \text{ The distribution of } X \text{ is elliptically symmetric.} \tag{1.1}$$

There are many references concerning this issue in the literature. For example, among others, [1–3,5,7,9,11,12,16,18]. Most of them are for spherical symmetry, a special case of the elliptical symmetry, where the symmetric center and the shape matrix of the random vector X are both known. Li et al. [14] proposed some Q–Q plots to test for the spherical and elliptical symmetry.

The proposed test statistics often have intractable sampling and limiting null distributions. Hence how to determine critical values is a crucial issue in this setup. Zhu et al. [17] suggested a bootstrap test for the spherical symmetry and proved that the proposed bootstrap test is asymptotically valid. As an alternative, we in the present paper develop conditional test procedures. The tests are easy to implement and have some nicer properties than the bootstrap tests. There are the following three points: (a) the conditional tests are exactly valid for the spherical symmetry; (b) they are distribution-free under the null hypothesis; (c) the distributions of the conditional tests approximate to the null distributions of their unconditional counterparts.

The paper is organized in the following way. Section 2 contains some tests which handle separately the cases with the known and unknown center and shape matrix and the asymptotic behavior of the test statistics under null hypothesis and alternatives. The conditional test procedures are defined in Section 3 and the exact and asymptotic validities of the tests are also presented in this section. A small simulation is contained in Section 4. All proofs of the theorems are postponed to Section 5.

2. Test statistic and asymptotics

As a slight extension of Ghosh and Ruymgaart's [11] test, we define test statistic as follows:

$$\int_{S^d}\int_I (\sqrt{n}P_n\{\sin(ta'\hat{A}(X-\hat{\mu}))\})^2 w(t)\,dt\,dv(a), \tag{2.1}$$

where P_n is the empirical measure based on the sample points $\{X_l, \ldots, X_n\}$ which are iid copies of X, $P_n(f)$ stands for $(1/n)\sum_{j=1}^n f(X_j)$ for each function $f(\cdot)$, $\hat{A} = \Sigma^{-1/2}$

or $=\hat{\Sigma}^{-1/2}$, the sample covariance and $\hat{\mu}=\mu$ or $=\bar{X}$ the sample mean, respectively, in accordance to the parameters being known or unknown, $w(\cdot)$ is a weight function with a compact support, $a\in S^d=\{a:||a||=1, a\in R^d\}$, v is the uniform distribution on S^d and I is a working region. In this paper, I is compact subset of the real line R. The null hypothesis H_0 is rejected for the large values of the test statistic.

In order to study the asymptotic properties of the test statistic, we define the empirical process by

$$V_n=\{V_n(\mathbf{X}_n,\hat{\mu},\hat{A},t,a)=\sqrt{n}P_n\{\sin(ta'\hat{A}(X-\hat{\mu}))\}:(t,a)\in I\times S^d\}, \tag{2.2}$$

and the test statistic in (2.1) can be rewritten as

$$E_n=\int_{S^d}\int_I\{V_n(\mathbf{X}_n,\hat{\mu},\hat{A},t,a)\}^2\,dw(t)\,dv(a). \tag{2.3}$$

The limit behavior, under the null hypothesis, of the empirical processes defined above is presented in the following theorem and corollary. For simplicity, call a Gaussian process with index set $I\times S^d$ is continuous if its sample paths are bounded and uniformly continuous with respect to $(t,a)\in I\times S^d$.

Theorem 2.1. *Assume that $P\{X=\mu\}=0$ and $E||X-\mu||^4<\infty$. Then under* H_0

(1) *If the center μ is given and then $\hat{\mu}=\mu$, the process V_n converges in distribution to a centered continuous Gaussian process $V_1=\{V_1(t,a):(t,a)\in I\times S^d\}$ with the covariance kernel*:

$$E\{\sin(ta'A(X-\mu))\sin(sb'A(X-\mu))\}:\textit{for}\ (t,a),(s,b)\in I\times S^d. \tag{2.4}$$

(2) *Let*

$$k(t,a,x)=\sin(ta'A(x-\mu))-ta'A(x-\mu)E(\cos(ta'A(X-\mu))). \tag{2.5}$$

If the center μ is an unknown parameter and then $\hat{\mu}=\bar{X}$, the process V_n converges in distribution to a centered continuous Gaussian process $V_2=\{V_2(t,a):(t,a)\in I\times S^d\}$ with the covariance kernel: for $(t,a),(s,b)\in I\times S^d$,

$$E\{k(t,a,x)k(s,b,x)\}. \tag{2.6}$$

The convergence of the test statistic is a direct consequence of Theorem 2.1.

Corollary 2.2. *The test statistic E_n associated with known and unknown center converges in distribution to the quadratic functionals $\int_{S^d}\int_I V_1^2(t,a)\,dw(t)\,dv(a)$ and $\int_{S^d}\int_I V_2^2(t,a)\,dw(t)\,dv(a)$, respectively.*

We now investigate the behavior of the test under alternatives. For convenience, let $\sin^{(i)}(c)$ be ith derivative of $\sin(\cdot)$ at point c. If there is a direction $a\in S^d$ such that

$E[\sin(ta'A(X-\mu))] \neq 0$ for some $t \in I$, it is easily derived that from the continuity of function $E[\sin(ta'A(X-\mu))]$ w.r.t. (t,a), the test statistic E_n converges in distribution to infinity as in this case the process V_n converges in distribution to infinity. This means that the tests are consistent against global alternatives. The rest of this section focuses on the investigation with local alternatives.

Suppose that the iid d-variate vectors $X_i = X_{in}$ have the expression $Z_i + Y_i/n^\alpha$, $i = 1, \ldots, n$ for some $\alpha > 0$. The center $\mu = \mu_n = E(Z) + E(Y)/n^\alpha$. When Z_i is independent of Y_i, the distribution of X_{in} is a convolution of two distributions, and one of them converges to the degenerate distribution at zero with the rate n^α in certain sense.

Theorem 2.3. *Assume that the following conditions hold*:

(1) *Both distributions of Z and Y are continuous. In addition, Z is elliptically symmetric with the center $E(Z)$ and the shape matrix Σ.*
(2) *There is an integer l being the smallest one such that*

$$\sup_{(t,a)\in I\times S^d} |B_l(t,a)| =: \sup_{(t,a)\in I\times S^d} |E((ta'A(Y-E(Y))^l \sin^{(l)}(ta'A(Z-E(Z))))| \neq 0,$$

$$E(||Y||^{2l}) < \infty, \quad \textit{and} \quad E(||Y||^{2(l-1)}||Z||^2) < \infty. \tag{2.7}$$

Then when $\alpha = 1/(2l)$, if $\hat{\mu} = \mu$

$$E_n \Rightarrow \int_{S^d}\int_I (V_1(t,a) + 1/l!B_l(t,a))^2 \, dw(t)v(a), \tag{2.8}$$

and if $\hat{\mu} = \bar{X}$,

$$E_n \Rightarrow \int_{S^d}\int_I (V_2(t,a) + 1/l!B_l(t,a))^2 \, dw(t)v(a), \tag{2.9}$$

where "$\Rightarrow$" stands for the convergence in distribution, V_1 and V_2 are the Gaussian processes defined in Theorem 2.1.

Remark 2.1. Comparing with the limiting variables under the null hypothesis in Corollary 2.2 and the ones under the alternative in Theorem 2.3, we see that the test can detect the local alternatives distinct $O(n^{-1/(2l)})$ from the null. In some cases, this rate can reach a parametric one, that is, $l = 1$. For example, suppose that Z has the uniform distribution on S^d and $Y = (Z_1^2 - 1, \ldots, Z_d^2 - 1)$, we can see easily that, via a little elementary calculation, $\sup_{(t,a)\in I\times S^d}|E(ta'AY\cos(ta'AZ))| \neq 0$. Hence, $l = 1$. On the other hand, when Z and Y are independent, $l \geqslant 3$, namely, the test can detect, at most, alternative distinct $O(n^{-1/6})$ from the null. In fact, it is clear that for $l = 1, 2$

$$\sup_{(t,a)\in I\times S^d} |E((ta'AY)^l \sin^{(l)}(ta'AZ))| = 0.$$

We also note that the test is Cramer–von Mises type and then is omnibus because of the absolute and square values in the test statistic, and that therefore the test is asymptotically unbiased for all shapes of the function B_l.

3. Conditional test procedures

The basic idea of our method is quite simple and the test procedures are easy to implement. Consider the known center-shape matrix case first. The conditional test procedures are based on the following property of elliptical distribution that X is elliptically symmetric if and only if

$$A(X-\mu) = v||A(X-\mu)||, \tag{3.1}$$

where $v = A(X-\mu)/||A(X-\mu)||$ is uniformly distributed on S^d and independent of $||A(X-\mu)||$ (see, e.g. [6] or [3]). Hence for any u being uniformly distributed on S^d, $u||A(X-\mu)||$ has the same distribution as $A(X-\mu)$. This leads up to that a statistic, say $T_n(A(X_1-\mu),\ldots,A(X_n-\mu))$ based on X_i's, has the same distribution as that of $T_n(u_1||A(X_1-\mu)||,\ldots,u_n||A(X_n-\mu)||)$ with u_i having the uniform distribution on S^d, and, given $||A(X_i-\mu)||$'s, the conditional distributions of $T_n(A(X_1-\mu),\ldots,A(X_n-\mu))$ and $T_n(u_1||A(X_1-\mu)||,\ldots,u_n||A(X_n-\mu)||)$ are also identical. Note that the unconditional distribution is the expectation of the conditional distribution over $||A(X_i-\mu)||$'s. Intuitively, from the LLN, the conditional distribution of the test statistic based on the generated data $u_i||A(X_i-\mu)||$'s may converge to the unconditional distribution which is also the unconditional distribution of the test statistic based on the original data as described above. Hence our approximation can be consistent. We shall prove this later. Furthermore, comparing with bootstrap methods, our procedure creates new data having elliptical distributions (each data point, when given $||A(X_i-\mu)||$, has a uniform distribution which is itself elliptical on $\{a : a \in R^d, ||a|| = ||A(X_i-\mu)||\}$, the conditional hypothetical distributions of the original data, while bootstrap methods generally create the new data whose distribution is the empirical distribution or its variants when given data points which relies on the underlying distribution. Consequently, the conditional distribution and then the critical values determined by our procedures may not change with the underlying distribution significantly while the ones determined by the bootstrap methods may do. Actually, we can have the exact validity of the test. It will be presented in Theorem 3.1 below. For other cases in which the A and μ are the unknowns to be estimated, the distributions of the tests mentioned above may be identical asymptotically. From these observations, we can generate u_i and then, for given data X_i's, approximate the conditional distribution of T_n given $||A(X_i-\mu)||$'s by the Monte Carlo method. We shall later verify theoretically the above observations, namely, prove the conditional distribution based on the generated data given $||A(X_i-\mu)||$'s to be a consistent approximation for almost all sequences $\{X_1,\ldots,X_n,\ldots\}$ to the null distribution of T_n.

Furthermore, from the above description, we know that property (3.1) is the key for constructing the conditional test procedures. It also means that testing elliptical symmetry cannot be simply reduced to testing uniformity because for other distribution v is not independent of $||A(X-\mu)||$, therefore, we cannot use the above Monte Carlo method to obtain a conditional distribution which is a good approximation of the unconditional counterpart.

A simple algorithm of the conditional test procedure can be performed to approximate critical values as follows. Assume that both μ and Σ are given. It is worthwhile to mention that the algorithm can be applied to any test statistic T_n when the center and the shape matrix are given. But when the center is unknown, we need some modification for the algorithm. Therefore, we in the following describe the algorithm for the specific test statistic defined in the paper so that it also can be used for unknown center case with some modification.

Step 1: Generate by computer iid random vectors, say u_i, of size n with uniform distribution on S^d, let $\mathbf{U}_n = (u_1, \dots, u_n)$. The new data are $u_i||A(X_i - \mu)||$.

Step 2: Accordingly as the empirical process defined in (2.2), we define a conditional empirical process. For fixed $\mathbf{X}_n = (X_1, \dots, X_n)$, let

$$V_{n1}(\mathbf{U}_n) = \{V_{n1}(\mathbf{X}_n, \mathbf{U}_n, t, a) = \sqrt{n}P_n\{\sin(ta'u||A(X - \mu)||)\} : (t, a) \in I \times S^d\}, \tag{3.2}$$

and calculate the value of the statistic

$$E_{n1}(\mathbf{U}_n) = \int_{S^d} \int_I \{V_{n1}(\mathbf{X}_n, \mathbf{U}_n, t, a)\}^2 \, dw(t) \, dv(a). \tag{3.3}$$

Step 3: Repeat steps 1 and 2 m times to get m values $E_{n1}(\mathbf{U}_n^{(j)})$, $j = 1, \dots, m$.

Step 4: Define $E_{n1}(U_n^{(0)})$ as the value of E_n. Estimate the p-value by $p = k/(m+1)$ where k is the number that $E_{n1}(U_n^{(j)})$ $j = 0, 1, \dots, m$ are greater than or equal to $E_{n1}(U_n^{(0)})$.

The following theorem states the exact validity of the test $E_{n1}(\mathbf{U}_n)$.

Theorem 3.1. *Assume that $X_1, \dots, X_n$ are iid d-variate vectors which are elliptically symmetric with known center and shape matrix μ and Σ. Let $\mathbf{U}_n^{(1)}, \dots, \mathbf{U}_n^{(m)}$ be the independent copies of $\mathbf{U}_n$. Then for any $0 < \alpha < 1$,*

$$G_{n,m}^{(1)}(\alpha) = P\{p \leqslant \alpha\} \leqslant \frac{[m\alpha] + 1}{m + 1}, \tag{3.4}$$

where $[z]$ stands for the largest integer part of z.

The above algorithm can be extended directly to the case where the center is known but the shape matrix is unknown. For the unknown center case, the situation is not so simple. In order to ensure the equivalence between the conditional empirical process below and its unconditional counterpart, we shall use the following fact to construct conditional empirical process. It can be derived that uniformly on

$t \in I$ and $a \in S^d$

$$\begin{aligned}
&\sqrt{n}P_n(\sin(ta'\hat{A}(X-\bar{X}))) \\
&\quad = \sqrt{n}P_n(\sin(ta'\hat{A}(X-\mu))\cos(ta'\hat{A}P_n(X-\mu))) \\
&\qquad - \sqrt{n}(P_n((ta'\hat{A}(X-\mu))\sin(ta'\hat{A}P_n(X-\mu))) \\
&\quad = \sqrt{n}P_n(\sin(ta'A(X-\mu)) \\
&\qquad - \sqrt{n}\sin(ta'AP_n(X-\mu)))(P_n(ta'A\cos(t'(X-\mu))+o_p(1).
\end{aligned}$$

We then define a conditional empirical process in Step 2 of the algorithm as

$$V_{n2}(\mathbf{U}_n) = \{V_{n2}(\mathbf{X}_n, \mathbf{U}_n, \hat{\mu}, \hat{A}, t, a) : (t,a) \in I \times S^d\}, \tag{3.5}$$

where

$$\begin{aligned}
&V_{n2}(\mathbf{X}_n, \mathbf{U}_n, \hat{\mu}, \hat{A}, t, a) \\
&\quad = \sqrt{n}P_n\{\sin(ta'u||\hat{A}(X-\hat{\mu})||)\} \\
&\qquad - \sqrt{n}P_n\{\cos(ta'u||\hat{A}(X-\hat{\mu})||)\sin(ta'P_n(u||\hat{A}(X-\hat{\mu})||))\}.
\end{aligned} \tag{3.6}$$

The associated conditional statistic is defined as

$$E_{n2}(\mathbf{U}_n) = \int_{S^d}\int_I \{V_{n2}(\mathbf{X}_n, \mathbf{U}_n, \hat{\mu}, \hat{A}, t, a)\}^2 \, dw(t)\, dv(a). \tag{3.7}$$

We in the following theorem present the asymptotic equivalence between the conditional empirical processes $V_{n1}(\mathbf{U}_n)$ and $V_{n2}(\mathbf{U}_n)$ and their unconditional counterparts. The asymptotic validity of $E_{n1}(\mathbf{U}_n)$ and $E_{n2}(\mathbf{U}_n)$ is a direct consequence.

Theorem 3.2. *Assume, in addition to the conditions of Theorem* 3.1, *that* $P\{X=\mu\}=0$. *Then the conditional empirical processes* $V_{n1}(\mathbf{U}_n)$ *and* $V_{n2}(\mathbf{U}_n)$ *given* $\mathbf{X}_n$ *in* (3.2) *and* (3.5) *converge, for almost all sequences* $\{X_1, \dots, X_n, \dots\}$, *in distribution to the Gaussian process* V_1 *and* V_2 *defined in Theorem* 2.1, *respectively, which are the limits of the unconditional counterparts* V_n *with known and unknown centers. This leads up to that the conditional statistics* $E_{n1}(\mathbf{U}_n)$ *and* $E_{n2}(\mathbf{U}_n)$ *given* $\mathbf{X}_n$ *in* (3.3) *and* (3.7) *have almost surely the same limits as those of the statistics* E_n *with known and unknown centers respectively,* $E = \int (V(a,t))^2\, dw(t)\, dv(a)$ *and* $E_1 = \int (V_1(a,t))^2\, dw(t)\, dv(a)$.

Remark 3.1. The optimal choice of the working region I and the weight function $w(\cdot)$ is beyond the scope of this paper. In our simulations, the working region was $[-2,2]$ and $w(\cdot)$ was a constant, the uniform distribution density. It is worth mentioning that in some cases, the choice of working regions is not very important. We now show an example in which the fact that the imaginary part of the characteristic function equals zero in a compact subset of R^d such as $[-2,2] \times S^d$ is equivalent to that the imaginary part is zero in whole space R^d. Suppose that the moment generating function of a multivariate vector, X say, exists in a sphere

$[-b, b] \times S^d$, $b > 0$. Then the moment generating function of $a'X$, the linear projector of X on R^1, exists in an interval $[-b_1, b_1]$ for any $a \in S^d$, where b_1 does not depend on a. If the imaginary part of the characteristic function of X equals zero in a sphere $[-b_2, b_2] \times S^d$, so does the one of the characteristic function of $a'X$ in an interval $[-b_3, b_3]$. It is easy to see that all moments of $a'X$ with odd orders equal zero. This means that the characteristic function of $a'X$ is real, and then $a'X$ is symmetric about the origin for any a. This conclusion implies, in turn, that imaginary part of the characteristic function of X is zero in R^d. Consequently, the choice of working region is not very important in such a case.

Remark 3.2. It is noted that under the local alternative the distribution of the conditional test statistic is convergent to that under the null hypothesis. We omit the detail here, the reader can refer the technical report of the University of Hong Kong [19].

Remark 3.3. Romano [16] proposed a general method of the randomization tests. From the idea of permutation test proposed by Hoeffding (1952), the randomization tests are constructed in terms of the invariance of the distribution for a class $\mathbf{G}_n$ of transformations, [16, p. 151]. The spherically symmetric distribution has such an invariance property. Similar argument is used in [4,7,8]. For testing spherical symmetry, our test procedure is similar to Romano's and Diks and Tong's.

4. A simulation study

In this section, a small simulation study was performed. In the simulation results reported in Tables 1 and 2, the sample size $n = 20, 50$. The dimension of random vector X, $d = 2, 4, 6$. We consider that (1) both μ and Σ are known (testing for spherical symmetry); (2) μ is known, Σ needs to be estimated; and (3) both μ and Σ are unknown. The test statistic E_n was rewritten as E_{ni}, $i = 1, 2, 3$ in accordance to these three cases, respectively. For power study, we consider the vector $X = Z + b \cdot Y$ for $b = 0.00, 0.25, 0.5, 0.75, 1.00$, and 1.25, where Z has the normal distribution $N(\mu, \Sigma)$ and Y is the random vector with the independent χ_1^2 components. The hypothetical distribution was normal $N(\mu, \Sigma)$. That is, $b = 0.00$ corresponds to the null hypothesis H_0. When $b \neq 0.00$, the distribution will not be elliptically symmetric. In simulation, we generated data from $N(0, I_3)$ and, accordingly as different setup, regarded the symmetric center and the shape matrix as the given ones or unknown parameters separately.

In order to get a critical value when given the data $\{(Y_1, Z_1), \ldots, (Y_n, Z_n)\}$, we generated 1000 $\mathbf{U}_n$ pseudo-random vectors of $n = 20$ and 50 by Monte Carlo method. The basic experiment was replicated 1000 times for each combination of the sample sizes and the underlying distributions of the vectors. The nominal level was 0.05. The proportion of times that the values of the statistics exceeded the critical values were recorded as the empirical power of the tests.

Table 1
Power of the tests with $n = 20$

	b	0.00	0.25	0.50	0.75	1.00	1.25
$d = 2$	E_{n1}	0.048	0.130	0.292	0.579	0.785	0.884
	E_{n2}	0.046	0.201	0.332	0.584	0.776	0.870
	E_{n3}	0.043	0.223	0.391	0.578	0.630	0.663
$d = 4$	E_{n1}	0.052	0.120	0.278	0.573	0.697	0.872
	E_{n2}	0.045	0.181	0.315	0.577	0.679	0.861
	E_{n3}	0.040	0.238	0.402	0.579	0.633	0.643
$d = 6$	E_{n1}	0.056	0.142	0.300	0.574	0.685	0.865
	E_{n2}	0.053	0.195	0.345	0.581	0.667	0.860
	E_{n3}	0.038	0.251	0.407	0.576	0.616	0.654

Table 2
Power of the tests with $n = 50$

	b	0.00	0.25	0.50	0.75	1.00	1.25
$d = 2$	E_{n1}	0.048	0.250	0.432	0.649	0.880	0.954
	E_{n2}	0.046	0.261	0.392	0.664	0.874	0.950
	E_{n3}	0.046	0.283	0.455	0.648	0.797	0.853
$d = 4$	E_{n1}	0.052	0.220	0.378	0.636	0.892	0.952
	E_{n2}	0.045	0.281	0.385	0.637	0.881	0.957
	E_{n3}	0.046	0.288	0.462	0.635	0.831	0.850
$d = 6$	E_{n1}	0.056	0.242	0.380	0.638	0.883	0.965
	E_{n2}	0.053	0.295	0.395	0.640	0.866	0.960
	E_{n3}	0.043	0.311	0.487	0.641	0.818	0.846

Looking at Table 1 with $n = 20$, we see that, under the null hypothesis, that is $b = 0.00$, the size of the tests E_{n1} and E_{n2} are close to the nominal one and E_{n3} is somewhat conservative. But it gets better with increase of the sample size. Under the alternatives, namely $b \neq 0.00$, when b is small, it seems that the test with estimated center and shape matrix would be more sensitive, see the case with $b = 0.25, 0.50$. With the increase of b, the power performance is reverse. E_{n1} is the best while E_{n3} gets worse. The situation with $n = 50$, looking at Table 2, is similar except for higher power than that with $n = 20$. Furthermore, the power performance of the tests are less affected by the dimension of the variable.

Acknowledgments

Research supports from NNSF of China, a UGC Grant of Hong Kong, HKSAR, China and the contract "Projet d'Actions de Recherche Concertees" (PARC

No. 93/98-164) of the Belgian Government are gratefully acknowledged. We thank the Editor and the two referees for their comments which greatly improved the presentation of the paper and Prof. Dr. Norbert Henze for the helpful discussion.

Appendix A. Proofs of theorems

A.1. Proof of theorems in Section 2

Proof of Theorem 2.1. Ghosh and Ruymgaart [11] has proved that, when the center and the shape matrix are given, the process V_n converges in distribution to V_1 with the covariance kernel in (2.4). When the shape matrix is replaced by sample covariance matrix $\hat{\Sigma}$, applying the triangle identity; we have

$$\begin{aligned}\sqrt{n}P_n(\sin(ta'\hat{A}(X-\mu))) &= \sqrt{n}P_n(\sin(ta'A(X-\mu)))\cos(ta'(\hat{A}-A)(X-\mu)))\\ &\quad + \sqrt{n}(P_n(\cos(ta'A(X-\mu)))\sin(ta'(\hat{A}-A)(X-\mu)))\\ &=: I_{n1}(t,a)+I_{n2}(t,a).\end{aligned}$$

It is well-known that by the conditions $\sqrt{n}(\hat{A}A^{-1}-I_d)=O_p(1), \max_{1\leqslant j\leqslant n}||X_j-\mu||/n^{1/4}\to 0$, *a.s.* and $E(A(X-\mu)\cos(ta'A(X-\mu)))=0$ which is implied by the spherical symmetry of $A(X-\mu)$, we then easily derive that, uniformly over $(t,a)\in I\times S^d$,

$$\begin{aligned}I_{n1}(t,a) &= \sqrt{n}P_n(\sin(ta'A(X-\mu)))+O_p(1/\sqrt{n}),\\ I_{n2}(t,a) &= ta'\sqrt{n}(\hat{A}A^{-1}-I_d)(P_n(A(X-\mu)\cos(ta'A(X-\mu)))=o_p(1).\end{aligned}$$

This implies that V_n with the sample covariance matrix $\hat{\Sigma}$ is asymptotically equivalent to that with Σ. Conclusion (1) is proved. For conclusion (2), the argument is analogous since we can derive that

$$\begin{aligned}\sqrt{n}P_n(\sin(ta'\hat{A}(X-\hat{\mu}))) &= \sqrt{n}P_n(\sin(ta'\hat{A}(X-\mu)))\cos(ta'\hat{A}(\hat{\mu}-\mu)))\\ &\quad - \sqrt{n}(P_n(\cos(ta'\hat{A}(X-\mu)))\sin(ta'\hat{A}(\hat{\mu}-\mu)))\\ &= \sqrt{n}P_n(\sin(ta'\hat{A}(X-\mu)))\\ &\quad - \sqrt{n}ta'\hat{A}(\hat{\mu}-\mu))E(\cos(ta'\hat{A}(X-\mu)))+o_p(1).\end{aligned}$$

The proof of Theorem 2.1 is completed. □

Proof of Theorem 2.3. Consider the case of $\hat{\mu}=\mu$ first. Assume no loss of generality that the center $\mu=0$ and the covariance matrix of X_{in} is $\Sigma_n=(A_n)^{-2}$. Note that Σ_n converges to the covariance matrix of the variable Z, Σ say. Applying the Taylor expansion to *sine* function, for any $(t, a) \in I \times S^d$,

$$\begin{aligned}
&\sqrt{n}P_n\{\sin(ta'A_n(Z + Y/n^{1/(2l)}))\} \\
&\quad = \sqrt{n}P_n\{\sin(ta'A_nZ)\} + \sum_{i=1}^{l-1} (1/i!)n^{-i/(2l)}\sqrt{n}P_n\{(ta'A_nY)^i \sin^{(i)}(ta'A_nZ)\} \\
&\qquad +(1/l!)n^{-1}\sum_{j=1}^{n}\{(ta'A_nY_j)^l \sin^{(l)}(ta'A_n(Z_j + (t'Y_j)^*/n^{1/(2l)})) - \sin^l(ta'A_nZ_j))\} \\
&\qquad + (1/l!)P_n\{(ta'A_nY)^l \sin^{(l)}(ta'A_nZ)\},
\end{aligned} \tag{A.1}$$

where $(ta'A_nY_j)^*$ is a value between 0 and $ta'A_nY_j$. We need to show that the second and third summands in RHS of (A.1) tend to zero in probability as $n \to \infty$, and the fourth summand converges in probability to $E\{(ta'AY)^l)\sin^{(l)}(ta'AZ)\}$. The convergence of the fourth term is obvious. Noticing that $E\{(ta'AY)^i)\sin^{(i)}(ta'AZ)\} = 0$ for $1 \leqslant i \leqslant l-1$, and similar argument used in the proof of Theorem 2.1 can be applied. The proof for V_n with the known center and then for E_n is finished.

For V_n with the estimated covariance matrix, we note that $\max_{1\leqslant j\leqslant n} ||Y_j||/n^{1/(2l)} \to 0$, *a.s.*, $\sqrt{n}(\hat{A}_n - A_n) = O_p(1)$ and $A_n - A = o(1)$. Furthermore,

$$\begin{aligned}
&\sup_{(t,a)\in I\times S^d} |P_n(\sin(ta'\hat{A}_n(Z - E(Z)) + (Y - E(Y)/n^{1/(2l)})) - \sin(ta'\hat{A}_n(Z - E(Z))))| \\
&\quad \leqslant cP_n||\hat{A}_n(Y - E(Y))||/n^{1/(2l)} = O(n^{-1/(2l)}) \text{ a.s.,}
\end{aligned}$$

and

$$\begin{aligned}
&\sup_{(t,a)\in I\times S^d} |1 - \cos(ta'P_n(\hat{A}_n(Z - E(Z)) + \hat{A}_n(Y - E(Y)/n^{1/(2l)})))| \\
&\quad \leqslant c(||P_n\hat{A}_n(Z - E(Z))||^2 + ||P_n\hat{A}_n(Y - E(Y))||^2/n^{1/l}) = O_p(n^{-1}).
\end{aligned}$$

Similar argument used in the proof of Theorem 2.1 can be applied again. Omit the details. From the convergence of V_n we immediately derive the convergence of E_n in (2.8).

For the case of $\hat{\mu} = \bar{X}$, we further note that

$$\begin{aligned}
&\sup_{(t,a)\in I\times S^d} |\sqrt{n}(\sin(ta'P_n\hat{A}_n((Z - E(Z)) + (Y - E(Y)/n^{1/(2l)}))) \\
&\quad - \sin(ta'P_n\hat{A}_n(Z - E(Z))))| \\
&\qquad \leqslant c\sqrt{n}||\hat{A}_n(P_nY - E(Y))||/n^{1/(2l)} = O_p(n^{-1/(2l)}).
\end{aligned}$$

Based on the above inequalities and the triangle identity, it is easy to see that

$$\begin{aligned}
&\sqrt{n}P_n(\sin(ta'\hat{A}_n(Z + Y/n^{1/(2l)} - (\bar{Z} + \bar{Y}/n^{1/(2l)}))))\\
&\quad = \sqrt{n}P_n(\sin(ta'A_n(Z + Y/n^{1/(2l)} - (E(Z) + E(Y)/n^{1/(2l)}))))\\
&\qquad - \sqrt{n}P_n(\cos(ta'A_n(Z - E(Z)))\sin(ta'P_nA_n(Z - E(Z))) + O_p(n^{-1/(2l)})\\
&\quad = \sqrt{n}P_n(\sin(ta'A(Z - E(Z)))) + (1/l!)E\{(ta'A(Y - E(Y)))^l \sin^{(l)}(ta'A(Z - E(Z)))\}\\
&\qquad - \sqrt{n}\sin(ta'P_nA(Z - E(Z)))E(\cos(ta'A(Z - E(Z))) + o_p(1)\\
&\quad \Rightarrow V_2(t,a) + (1/l!)B_l(t,a).
\end{aligned} \tag{A.2}$$

It implies the convergence of E_n in (2.9). The proof is completed. □

A.2. Proof of theorems in Section 3

Proof of Theorem 3.1. As μ and Σ are known, we assume without loss of generality that $\mu = 0$ and $\Sigma = I_d$, the identity matrix. Hence $A(X-\mu) = X$ write $||\mathbf{X}_n|| = (||X_1||, \dots, ||X_n||)$ and $\mathbf{U}_n^0 = (v_1, \dots, v_n)$. Note that $X = v||X||$ and $\mathbf{U}_n^0$ has the same distribution as that of $\mathbf{U}_n^{(j)}$'s. Then given $||\mathbf{X}_n||, E_{n1}$ can be written as $E_{n1}(\mathbf{U}_n^0)$ and $E_{n1}(\mathbf{U}_n^{(j)})$ $j = 0, 1, \dots, m$ are $m+1$ iid variables, which implies that

$$P\{p \leqslant \alpha | ||\mathbf{X}_n||\} \leqslant \frac{[m\alpha] + 1}{m+1}.$$

The proof is concluded from integrating out over $||\mathbf{X}_n||$. □

Proof of Theorem 3.2. We only need to show the convergence of the processes, which implies the convergence of the test statistics. First we show that $\{V_{n1}(\mathbf{U}_n, \mathbf{X}_n, t, a) : (t,a) \in I \times S^d\}$ given $\mathbf{X}_n$ converges almost surely to the process $\{V_1(t) : (t,a) \in I \times S^d\}$ which is the limit of V_n with the known center. The argument of the proof will be applicable for showing the convergence of the process $V_{n2}(\mathbf{U}_n)$.

For the simplicity of notation, write X_j for $A(X_j - \mu)$. Define sets

$$D_1 = \left\{ \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^n ||X_j||^2 = E||X||^2 \right\},$$

$$D_2 = \left\{ \lim_{n\to\infty} \sup_{(t,a),(s,b)} \left| \frac{1}{n} \sum_{j=1}^n (\sin(ta'X_j)\sin(sb'X_j)) - E((\sin(ta'X)\sin(sb'X)) \right| = 0 \right\}$$

and $D = D_1 \cap D_2$. By the Lipschitz continuity of the *sine* function and Glivenko–Cantelli theorem for the general class of functions (e.g. [15, Theorem II 24, p. 25]), it is clear that D is a subset of sample space with probability measure one.

We assume without further mentioning that the given $\{X_1, \dots, X_n, \dots\} \in D$ in the following.

For the convergence of the empirical process defined in the theorem, all we need to do is to prove *Fidis convergence* and *Uniform tightness*. The proof of the *fidis convergence* is standard, so we only describe an outline. For any integer $k, (t_1, a_1) \ldots (t_k, a_k) \in I \times S^d$. Let

$$V^{(k)} = (\mathrm{cov}(\sin(t_i a_i' x), \sin(t_l a_l' x)))_{1 \leqslant i,l \leqslant k}$$

it needs to show that

$$V_{n1}^{(k)} = \{V_{n1}(\mathbf{U}_n, \mathbf{X}_n, t_i, a_i) : i = 1, \ldots, k\} \Rightarrow N(0, V^{(k)}).$$

It suffices to show that for any unit k-dimensional vector γ

$$\gamma' V_{n1}^{(k)} \Rightarrow N(0, \gamma' V^{(k)} \gamma). \tag{A.3}$$

Note that the variance of LHS in (A.3), as follows, converges in probability to $\gamma' V^{(k)} \gamma$

$$\gamma' (\widehat{\mathrm{Cov}}_{i,l})_{1 \leqslant i,l \leqslant k} \gamma,$$

with $\widehat{\mathrm{Cov}}_{i,l} = \frac{1}{n} \sum_{j=1}^{n} E(\sin(t_i a_i' u ||X_j||) \sin(t_l a_l' u ||X_j||))$ where the expectation is taken over u. Hence if $\gamma' V^{(k)} \gamma = 0$, (A.3) is trivial. Assume $\gamma' V^{(k)} \gamma > 0$. Invoking the boundedness of the *sine* function and the Lindeberg condition,

$$\gamma' V_{n1}^{(k)} / \sqrt{\gamma' V^{(k)} \gamma} \to N(0, 1).$$

That is (A.3) holds, the *fidis convergence* is then proved.

We now turn to prove *Uniform tightness* of the process. All we need to do is to show that for any $\eta > 0$ and $\varepsilon > 0$, there exists an $\delta > 0$ for which

$$\limsup_{n \to \infty} P\left\{ \sup_{[\delta]} |V_{n1}(\mathbf{U}_n, \mathbf{X}_n, t, a) - V_{n1}(\mathbf{U}_n, \mathbf{X}_n, s, b)| > 2\eta ||| \mathbf{X}_n || \right\} < \varepsilon, \tag{A.4}$$

where $[\delta] = \{((t, a), (s, b)) : ||ta - sb|| \leqslant \delta\}$. Since the limiting properties are investigated with $n \to \infty, n$ is always considered to be large enough below which simplifies some arguments of the proof.

It is easy to show that if a d-variate vector u is uniformly distributed on S^d, then u can be expressed as $e \cdot u^*$ where $e = \pm 1$ with probability one half, u^* has the same distribution as u and e and u^* are independent. From which, the LHS of (A.4) can be written as

$$\begin{aligned} &P\left\{ \sup_{[\sigma]} \sqrt{n} |P_n(\sin(ta' e \cdot u^* ||X||) - \sin(sb' e \cdot u^* ||X||))| > \eta ||| \mathbf{X}_n || \right\} \\ &\quad = P\left\{ \sup_{[\sigma]} \sqrt{n} |P_n{}^{\circ}(\sin(ta' u^* ||X||) - \sin(sb' u^* ||X||))| > \eta ||| \mathbf{X}_n || \right\}, \end{aligned} \tag{A.5}$$

where $P_n{}^{\circ}$ is the signed measure that places mass e_i/n at $u_i ||X_i||$, which is analogous to that in [15, p. 14].

We now consider conditional probability given $\mathbf{U}_n^* = (u_1^*, \ldots, u_n^*)$ and $||\mathbf{X}_n||$. Combining (A.5) with that $|\sin(ta'u^*||X||) - \sin(sb'u^*||X||)| \leqslant ||ta - sb||||X||$, the Hoeffding inequality implies that

$$P\{\sqrt{n}|(P_n{}^\circ(\sin(ta'u^*||X||) - \sin(sb'u^*||X||) > \eta c||ta - sb||\ |||\mathbf{X}_n||, \mathbf{U}_n^*\}$$
$$\leqslant 2\exp(-\eta^2/32).$$

In order to apply the chaining lemma (e.g. [15, p. 144]), we need to check, together with the above inequality, the covering integral

$$J_2(\delta, ||\cdot||, I \times S^d) = \int_0^\delta \{2\log\{(N_2(u, ||\cdot||, I \times S^d))^2/u\}\}^{1/2}\, du \tag{A.6}$$

is finite for small $\delta > 0$, where $||\cdot||$ is the Euclidean norm in R^d and the covering number $N_2(u, ||\cdot||, I \times S^d)$ is the smallest l for which there exist l points $t_1, \ldots, t_l$ with $\min_{1 \leqslant i \leqslant l}||ta - t_i a_i|| \leqslant u$ for every $(t, a) \in I \times S^d$. It is clear that

$$N_2(u/c, ||\cdot||, I \times S^d) \leqslant cu^{-d}.$$

Consequently, for small $\delta > 0$,

$$J_2(\delta, ||\cdot||, I \times S^d) \leqslant c\int_0^\delta (\log(1/u))^{1/2}\, du \leqslant c\delta\log(1/\delta) \leqslant c\delta^{1/2}.$$

(A.6) holds. Applying now the chaining lemma, there exists a countable dense subset $[\delta]^*$ of $[\delta]$ such that

$$P\left\{\sup_{[\delta]^*}\sqrt{n}|(P_n{}^\circ(\sin(ta'u^*||X||) - \sin(sb'u^*||X||))| > 26cJ_2(\delta, ||\cdot||, I \times S^d)|||\mathbf{X}_n||, \mathbf{U}_n^*\right\}$$
$$\leqslant 2c\delta.$$

The countable dense subset $[\delta]^*$ can be replaced by $[\delta]$ itself because $\sqrt{n}P_n{}^\circ\{\sin(ta'u||X||) - \sin(sb'u||X||)\}$ is a continuous function with respect to ta and sb for each fixed $||\mathbf{X}_n||$. Hence, choosing properly small δ, and integrating out over $\mathbf{U}_n^*$, the uniform tightness in (A.4) is proved. Therefore, the convergence of the process is proved. Then the convergence of $E_{n1}(\mathbf{U}_n)$ follows. The convergence of the process $V_{n2}(\mathbf{U}_n)$ can be proved by following the above argument and noticing $\hat{A} - A = O_p(1/\sqrt{n})$ and $\hat{\mu} - \mu = O_p(1/\sqrt{n})$. The limit of $V_{n2}(\mathbf{U}_n)$ is V_2, the limit of its unconditional counterparts. The asymptotic validity of E_{n2} then follows. The proof of Theorem 3.2 is finished. □

References

[1] L. Baringhaus, Testing for spherical symmetry of a multivariate distribution, Ann. Statist. 19 (1991) 899–917.

[2] L. Baringhaus, N. Henze, Limit distributions for measures of skewness and kurtosis based on projections, J. Multivariate Anal. 38 (1991) 51–69.

[3] R. Beran, Testing for ellipsoidal symmetry of a multivariate density, Ann. Statist. 7 (1979) 150–162.
[4] J. Besag, P.J. Diggle, Simple Monte Carlo tests for spatial pattern, Appl. Statist. 26 (1977) 327–333.
[5] D.K. Blough, Multivariate symmetry via projection pursuit, Ann. Inst. Statist. Math. 41 (1989) 461–475.
[6] A.P. Dempster, Elements of Continuous Multivariate Analysis, Addison-Wesley, Reading, MA, 1969.
[7] C. Diks, H. Tong, A test for symmetries of multivariate probability distributions, Biometrika 86 (1999) 605–614.
[8] S. Engen, M. Lillegård, Stochastic simulations conditioned on sufficient statistics, Biometrika 84 (1997) 235–240.
[9] K.T. Fang, L.X. Zhu, P.M. Bentler, A necessary test of goodness of fit for sphericity, J. Multivariate Anal. 44 (1993) 34–55.
[10] J.H. Friedman, Exploratory projection pursuit, J. Amer. Statist. Assoc. 82 (1987) 249–266.
[11] S. Ghosh, F.H. Ruymgaart, Applications of empirical characteristic functions in some multivariate problem, Canad. J. Statist. 20 (1992) 429–440.
[12] T. Kariya. M.L. Eaton, Robust tests for spherical symmetry, Ann. Statist. 1 (1977) 206–215.
[13] K.C. Li, Sliced inverse regression for dimension reduction (with discussions), J. Amer. Statist. Assoc. 85 (1991) 316–342.
[14] R.Z. Li, K.T. Fang, L.X. Zhu, Some Q-Q probability plots to test spherical and elliptical symmetry, J. Comput. Graph. Statist. 6 (1997) 435–450.
[15] D. Pollard, Convergence of Stochastic Processes, Springer, New York, 1984.
[16] J.P. Romano, Bootstrap and randomization tests of some nonparametric hypotheses, Ann. Statist. 17 (1989) 141–159.
[17] L.X. Zhu, K.T. Fang, I.M. Bhatti, P.M. Bentler, Testing sphericity of a high-dimensional distribution based on bootstrap approximation, Pakistan J. Statist. 14 (1995) 49–65.
[18] L.X. Zhu, G. Neuhaus, Nonparametric Monte Carlo tests for multivariate distribution, Biometrika 67 (2000) 129–153.
[19] L.X. Zhu, G. Neuhaus, Conditional test for elliptical symmetry, Technical report, Department of Statistics and Actuarial Science, The University of Hong Kong, 2001.

Further reading

C.R. Heathcote, S.T. Rachev, B. Cheng, Testing multivariate symmetry, J. Multivariate Anal. 54 (1995) 91–112.
G. Neuhaus, L.X. Zhu, Permutation tests for reflected symmetry, J. Multivariate Anal. 67 (1998) 129–153.
J.P. Royston, Some techniques for assessing multivariate normality based on the Shapiro-Wilk W, Appl. Statist. 32 (1983) 121–133.
L.X. Zhu, Convergence rates of empirical processes indexed by classes of functions and their application, J. Systems Sci. Math. Sci. 13 (1993) 33–41 (in Chinese).