Biometrika (2013), **xx**, x, *pp*. 1–8

Printed in Great Britain

Multivariate-sign-based high-dimensional tests for sphericity

BY CHANGLIANG ZOU, LIUHUA PENG, LONG FENG AND ZHAOJUN WANG
School of Mathematical Sciences, Nankai University, Tianjin, China
chlzou@yahoo.com.cn, allenberify@yahoo.com.cn, flnankai@126.com, zjwang@nankai.edu.cn

SUMMARY

This article concerns tests for sphericity when data dimension is larger than the sample size. The existing multivariate-sign-based procedure (Hallin & Paindaveine, 2006) for sphericity is not robust against high dimensionality, producing tests with type I error rates much larger than nominal levels. This is mainly due to bias from estimating the location parameter. We develop a correction that makes the existing test statistic robust against high dimensionality. We show that the proposed test statistic is asymptotically normal under elliptical distributions. The proposed method allows dimensionality to increase as the square of sample size. Simulations show that it has good size and power for a wide range of settings.

Some key words: Asymptotic normality; Large p, small n; Spatial median; Spatial sign; Sphericity test.

1. INTRODUCTION

High-dimensional data have dimension p that increases to infinity as the number of observations $n \to \infty$. Traditional statistical methods may fail in this situation since they are often based the assumption that p remains constant. This challenge calls for new research on properties of traditional methods, see Chen et al. (2009) and Hjort et al. (2009), for instance, and new statistical approaches to deal with high-dimensional data. Some new methods were proposed by Chen & Qin (2010) for a two-sample test for means, Ledoit & Wolf (2002), Schott (2005) and Chen et al. (2010) for testing a specific covariance structure, and Tang & Leng (2010) and the references therein for variable selection.

Sphericity assumptions play a key role in a number of statistical problems. The need to test the sphericity come from domains of application such as geostatistics, paleomagnetic studies, animal navigation, astronomy, wind direction data and microarray analysis. See Tyler (1987), Baringhaus (1991), Marden & Gao (2002) or Sirkiä et al. (2009) for references. Given the interest in both high-dimensional data and testing sphericity, the asymptotic and finite-sample properties of sphericity tests in the high-dimensional setting are worthy of careful investigation.

Because of its importance for applications, testing sphericity has a long history and has generated a considerable body of literature which we only very briefly review. Let $X_1, \ldots, X_n$, denote a p-dimensional sample of size n. The distribution of a p-dimensional random vector X is called spherical if for some $\theta \in \mathbb{R}^p$, the distribution of $X - \theta$ is invariant under orthogonal transformations. For multinormal variables, sphericity is equivalent to the covariance matrix of X, Σ, being proportional to the identity matrix I_p. Thus, the approaches, such as the likelihood ratio test (Mauchly, 1940), based on the covariance matrix are quite popular. John (1971; 1972) considered the testing problem in the normal distribution case and showed that the statistic

$$Q_{\mathrm{J}} = \frac{np^2}{2} \mathrm{tr} \left\{ \frac{S}{\mathrm{tr}(S)} - \frac{1}{p} I_p \right\}^2 \qquad (1)$$

provides the locally most powerful invariant test for sphericity under the multivariate normality assumption, where S is the regular sample covariance matrix and I_p is the identity matrix. This test is valid only under multivariate normality. In the wider elliptical model one can use a modification of John's test statistic, $Q_{\mathrm{M}} = Q_{\mathrm{J}}/\widehat{\kappa}$, where $\widehat{\kappa}$ is the estimated kurtosis based on the standardized fourth moment of the marginal distribution (Muirhead & Waternaux, 1980).

Ledoit & Wolf (2002) evaluated Q_{J} when the dimension p increases at the same rate as n, so that $p/n \to c$ for a finite c. Chen et al. (2010) developed a high-dimensional test based on Q_{J}, and showed that their proposed test statistic is asymptotically normal by assuming that the data has the structure $X_i = \Gamma Z_i + \mu$, where Γ is a $p \times m$ matrix, $m \geq p$, and $Z_i = (Z_{i1}, \ldots, Z_{im})^T$ is a random vector such that

$$\begin{aligned} &E(Z_i) = 0,\ \mathrm{var}(Z_i) = I_m,\ E(Z_{il}^{4k}) = m_{4k} \in (0, \infty), \\ &E(Z_{ik_1}^{\alpha_1} Z_{ik_2}^{\alpha_2} \cdots Z_{ik_q}^{\alpha_q}) = E(Z_{ik_1}^{\alpha_1}) E(Z_{ik_2}^{\alpha_2}) \cdots E(Z_{ik_q}^{\alpha_q}), \end{aligned} \tag{2}$$

whenever $\sum_{k=1}^{q} \alpha_k \leq 4k$. Here $k, k_1, \ldots, k_q$ are positive integers. The data structure (2) generates a rich collection of X_i from Z_i with a given covariance Σ. For example, the distributions in the so-called independent component model lie in the family given by (2); see Example 2.6 in Oja (2010) and references therein. It is difficult, however, to justify this model. The condition that power transformations of different components of Z_i are uncorrelated is almost equivalent to saying that $Z_{i1}, \ldots, Z_{im}$ are independent and thus not easily met in practice. For instance, it can be verified that a random vector from the multivariate t distribution or mixtures of multivariate normal distributions does not satisfy (2). Moreover, the statistical performance of this test would be degraded when the non-normality is severe, especially for heavy-tailed distributions; see Section 3. This motivates us to consider using multivariate-sign-and/or-rank-based covariance matrices to construct robust tests for sphericity.

This approach has been adopted by Tyler (1987), Ghosh & Sengupta (2001), Marden & Gao (2002), Hallin & Paindaveine (2006), and Sirkiä et al. (2009), among others. Most of the tests proposed by these researchers are based on the signs and the ranks of the norms of the observations centered at θ, with test statistics that have structures similar to Q_{J}. These statistics are distribution-free under sphericity and elliptical distributional assumptions, or asymptotically so. Hallin & Paindaveine (2006) or Oja (2010, Chapter 9) gave a nice overviews of this topic. Among them, the test entirely based on multivariate signs, also called spatial-sign by some authors, is of particular interests due to its simplicity and effectiveness, and has been discussed in detail by Marden & Gao (2002), Hallin & Paindaveine (2006), and Sirkiä et al. (2009). In this paper, we focus on this type of test. The existing calibration method is not robust against high dimensionality in the sense that it would produce tests with type I error rates much larger than nominal levels. This is mainly due to biases in estimating the location parameter. In the next section, we develop a bias-correction to the existing test statistic that makes it robust against high dimensionality. We show that the proposed test statistic is asymptotically normal for elliptical distributions. Simulation comparisons show that our procedure has good size and power for a wide range of dimensions, sample sizes and distributions. Finite sample studies also show that the proposed method works reasonably well when the underlying distribution is not elliptical, especially for the observations from the data structure (2). All the proofs are given in the online Supplementary Material.

2. HIGH-DIMENSIONAL TEST FOR SPHERICITY USING MULTIVARIATE SIGNS

2·1. *Inference based on sign covariance matrix*

Let $X_1, \ldots, X_n$ be a random sample from a p-variate elliptical distribution with density function $\det(\Sigma_p)^{-1/2} g_p\{||\Sigma_p^{-1/2}(X - \theta_p)||\}$, where $||X|| = (X^T X)^{1/2}$ is the Euclidean length of the vector X, θ_p is the symmetry center and Σ_p is a positive definite symmetric $p \times p$ scatter matrix. The matrix Σ_p describing the covariances between the p variables, can be expressed as $\Sigma_p = \sigma_p \Lambda_p$, where $\sigma_p = \sigma(\Sigma_p)$ is a scale parameter and $\Lambda_p = \sigma_p^{-1}\Sigma_p$ is a shape matrix. The scale parameter is assumed to satisfy $\sigma(I_p) = 1$ and $\sigma(a\Sigma_p) = a\sigma(\Sigma_p)$ for all $a > 0$. We wish to test the null hypothesis $H_0 : \Sigma_p = \sigma I_p$. Under the assumption of ellipticity, finite second order moments need not exist and sphericity is equivalent to $\Lambda_p = I_p$. If one wishes to test the hypothesis $H_0 : \Lambda_p = V$, one can use the standardized observations $V^{-1/2}X_i$ instead of the original ones. In the following we assume that the shape matrix is standardized so that $\mathrm{tr}(\Lambda_p) = p$.

The multivariate sign function is defined as $U(X) = ||X||^{-1}XI(X \neq 0)$. The observed signs for the X_i's are $U_i = U(X_i - \theta_p)$. Accordingly, the sign covariance matrix is defined by $\Omega_{n,p} = n^{-1}\sum_{i=1}^{n} U_i U_i^T$ (Hallin & Paindaveine, 2006). Under the null hypothesis, we have $E(\Omega_{n,p}) = I_p/p$. The sign test statistic can be defined by mimicking John's test (1) with $\Omega_{n,p}$ (Hallin & Paindaveine, 2006; Sirkiä et al., 2009)

$$Q_{\mathrm{S}} = p\mathrm{tr}\left[\{\mathrm{tr}(\Omega_{n,p})\}^{-1}\Omega_{n,p} - p^{-1}I_p\right]^2 = p\mathrm{tr}\left(\Omega_{n,p} - p^{-1}I_p\right)^2.$$

It can be shown that when p is fixed, under the null hypothesis,

$$n(p+2)Q_{\mathrm{S}}/2 \to \chi^2_{(p+2)(p-1)/2} \tag{3}$$

in distribution as $n \to \infty$. See Hallin & Paindaveine (2006) for the proof.

In high-dimensional settings, p diverges to infinity as $n \to \infty$, so χ^2_p is asymptotically normal with mean p and variance $2p$, we may expect that

$$\{\mathrm{var}(Q_{\mathrm{S}})\}^{-1/2}\{Q_{\mathrm{S}} - E(Q_{\mathrm{S}})\} \to N(0,1) \tag{4}$$

in distribution as $n \to \infty$ and $p \to \infty$. In what follows, we will show that the above convergence in law is essentially correct under mild conditions. However, the main impact of high dimensionality on the validity of the sign-based test does not stem from the difference between two asymptotic calibrations of Q_{S}, (3) and (4). In the foregoing discussion, the true location parameter θ_p is used in the definition of the sign vector, but in practice θ_p usually must be replaced by an estimator $\widehat{\theta}_{n,p}$. Any root-n consistent estimator would be adequate, but in the literature the rotation-equivariant spatial median (Möttönen & Oja, 1995) which minimizes the criterion function $L(\theta) = \sum_{i=1}^{n} ||X_i - \theta||$ is usually recommended. Taking the gradient of the objective function, one sees that $\widehat{\theta}_{n,p}$ is the solution to the equation

$$\sum_{i=1}^{n} U(X_i - \theta) = 0. \tag{5}$$

When p is fixed, replacing θ_p with $\widehat{\theta}_{n,p}$ does not affect the asymptotic properties of Q_{S}. However, as we will show in the next section, this substitution would yield a bias-term which is not negligible when $n/p = O(1)$. Even worse, when $n/p = o(1)$, the test based on (3) or (4) would have asymptotic size 1 under H_0. We will propose a simple remedy to address this problem.

2·2. *A bias-correction sign-based procedure*

The test statistic Q_{S} can be rewritten as

$$Q_{\mathrm{S}} = \frac{p}{n} + \frac{n(n-1)}{n^2}\frac{p}{n(n-1)}\sum_{i \neq j}(U_i^T U_j)^2 - 1,$$

so, we consider the modified test statistic

$$Q'_{\mathrm{S}} = \frac{p}{n(n-1)}\sum_{i \neq j}(U_i^T U_j)^2 - 1.$$

Substituting the spatial median $\widehat{\theta}_{n,p}$ into U_i, the test statistic becomes

$$\tilde{Q} = \frac{p}{n(n-1)} \sum_{i \neq j} (\widehat{U}_i^T \widehat{U}_j)^2 - 1,$$

where $\widehat{U}_i = U(X_i - \widehat{\theta}_{n,p})$.

PROPOSITION 1. *The test statistic $\tilde{Q}$ is invariant under rotations.*

The value of $\tilde{Q}$ remains unchanged for $Z_i = aOX_i + c$ where a is a constant, c is a vector of constants and O is an orthogonal matrix. Thus, without loss of generality, we assume that $\theta_p = 0$ in what follows. It is easy to see that $E(Q'_{\mathrm{S}}) = 0$ under H_0. However, in high-dimensional settings, $E(\tilde{Q})$ is not negligible with respect to $\mathrm{var}^{1/2}(\tilde{Q})$. Before proceeding, we state a necessary assumption which is required throughout this paper. Let $R_i = ||X_i - \theta_p||$.

Assumption 1. The moments $E(R_i^{-k})$ for $k = 1, \ldots, 4$ exist for large enough p; $E(R_i^{-k})/\{E(R_i^{-1})\}^k \to d_k \in [1, \infty)$, $k = 2, 3$, and 4 as $p \to \infty$, where the d_k are constants.

This assumption ensures the validity of second-order expansions we use and the existence of our bias-correction term. The moments $E(R_i^{-k})$ may not exist for a fixed p. For example, for standard multivariate normal and t distributions, $E(R_i^{-2})$ is equal to $1/(p-2)$ and thus the second moment exists only when $p > 3$. In the Supplementary Material, we verify this assumption for three commonly used elliptical distributions, the multivariate normal, the multivariate t distribution, and mixtures of multivariate normal distributions. We also formulate this assumption using the g_p that fixes the distribution of the modulus R_i. The existence of $E(R_i^{-k})$ is guaranteed if $r^{p-1-k} g_p(r)$ is bounded for $r \in (0, \epsilon)$.

We define

$$\begin{aligned}\delta_{n,p} =& \frac{1}{n^2} \left(2 - \frac{2E(R_i^{-2})}{\{E(R_i^{-1})\}^2} + \left[\frac{E(R_i^{-2})}{\{E(R_i^{-1})\}^2} \right]^2 \right) \\ &+ \frac{1}{n^3} \left(\frac{8E(R_i^{-2})}{\{E(R_i^{-1})\}^2} - 6 \left[\frac{E(R_i^{-2})}{\{E(R_i^{-1})\}^2} \right]^2 + \frac{2E(R_i^{-2})E(R_i^{-3})}{\{E(R_i^{-1})\}^5} - \frac{2E(R_i^{-3})}{\{E(R_i^{-1})\}^3} \right),\end{aligned} \tag{6}$$

and $\tilde{\sigma}_0^2 = 4(p-1)/\{n(n-1)(p+2)\}$. From the proof of Theorem 1 in Appendix 1, we know that

$$E(\tilde{Q}) = p\delta_{n,p} + o(n^{-1}), \quad \mathrm{var}(\tilde{Q}) = \tilde{\sigma}_0^2 + o(n^{-2}).$$

Clearly, if $n < p$, using the normal calibration like (4) for $\tilde{Q}$ would result in a bias term which cannot be ignored, producing type I error rates much larger than nominal levels. Hence, the key for our proposal is to correct this bias through approximating $E(\tilde{Q})$. The following theorem establishes the asymptotic null distribution of $\tilde{Q}$.

THEOREM 1. *Under H_0 and Assumption 1, if $p = O(n^2)$, then $(\tilde{Q} - p\delta_{n,p})/\tilde{\sigma}_0 \to N(0,1)$ in distribution as $p \to \infty$, $n \to \infty$.*

The unknown quantities in $\delta_{n,p}$ are $E(R_i^{-2})/\{E(R_i^{-1})\}^2$ and $E(R_i^{-3})/\{E(R_i^{-1})\}^3$. A straightforward approach is to consider moment estimators. We let $\widehat{R}_i = ||X_i - \widehat{\theta}_{n,p}||$ and $\widehat{R}_{i*} = \widehat{R_i} + \widehat{\theta}_{n,p}^T \widehat{U}_i - 2^{-1} \widehat{R}_i^{-1} ||\widehat{\theta}_{n,p}||^2$. Then $\widehat{R}_{i*}$ can be seen as a second-order approximation of R_i; see Lemmas 1 and 2. Using $\widehat{R}_{i*}$ instead of $\widehat{R}_i$ would further reduce the bias in estimation of $E(R_i^{-2})/\{E(R_i^{-1})\}^2$. Then the test statistic $(\tilde{Q} - p\widehat{\delta}_{n,p})/\tilde{\sigma}_0$ converges to $N(0,1)$ under H_0 as long as $np(\widehat{\delta}_{n,p} - \delta_{n,p}) = o_p(1)$, where $\widehat{\delta}_{n,p}$ is the estimator of $\delta_{n,p}$ by using $n^{k-1} \sum_{i=1}^n \widehat{R}_{i*}^{-k} / (\sum_{i=1}^n \widehat{R}_{i*}^{-1})^k$ to replace $E(R_i^{-k})/\{E(R_i^{-1})\}^k$ in

(6). The proposed test with level of significance α rejects H_0 if $(\tilde{Q} - p\widehat{\delta}_{n,p})/\tilde{\sigma}_0 > z_\alpha$ where z_α is the $1-\alpha$ quantile of $N(0,1)$.

Although Theorem 1 allows the dimensionality to increase at the rate of the square of the sample size, in practice how large p is allowed to increase would depend mainly on the rate of ratio-consistency of $n\sum_{i=1}^n \widehat{R}_{i*}^{-2}/(\sum_{i=1}^n \widehat{R}_{i*}^{-1})^2$ to ensure $np(\widehat{\delta}_{n,p} - \delta_{n,p}) = o_p(1)$. Suppose that $n\sum_{i=1}^n \widehat{R}_{i*}^{-2}/(\sum_{i=1}^n \widehat{R}_{i*}^{-1})^2 = E(R_i^{-2})/\{E(R_i^{-1})\}^2\{1+O_p(a_n)\}$. It can be shown that under the null hypothesis and Assumption 1, $a_n = n^{-1/2}$. Also, $n^2\sum_{i=1}^n \widehat{R}_{i*}^{-3}/(\sum_{i=1}^n \widehat{R}_{i*}^{-1})^3 \to E(R_i^{-3})/\{E(R_i^{-1})\}^3$ in probability. Thus, without imposing any other conditions, the bias-corrected method is valid when $p = o(n^{3/2})$. In certain cases, a_n can be improved. For example, under the condition $\mathrm{var}(R_i^{-k})/\{E(R_i^{-k})\}^2 = o(p^{-1/2})$ for $k=1,2$, we can get $a_n = o(n^{-1} + n^{-1/2}p^{-1/4})$. In such cases, $p = O(n^2)$ can be allowed. The multinormal distribution clearly satisfies this condition because we can show that $\mathrm{var}(R_i^{-k})/\{E(R_i^{-k})\}^2 = O(p^{-1})$. Technical details can be found in the Supplemental Material.

Remark 1. When X_i comes from the normal distribution, $\lim_{p\to\infty} E(R_i^{-k})/\{E(R_i^{-1})\}^k = 1$ under the null hypothesis. In this case, $\delta_{n,p}$ can be simplified as $\delta_{n,p} \approx n^{-2} + 2n^{-3}$. We find that using this $\delta_{n,p}$ works almost as well as using $\widehat{\delta}_{n,p}$ in all the considered cases. Hence, it is recommended in practice when one wishes to reduce computational effort.

Next, we consider the asymptotic distribution of $\tilde{Q}$ under the alternative $H_1 : \Lambda_p = I_p + D_{n,p}$. Define

$$\tilde{\sigma}_1^2 = \tilde{\sigma}_0^2 + n^{-2}p^{-2}\left\{8p\mathrm{tr}(D_{n,p}^2) + 4\mathrm{tr}^2(D_{n,p}^2)\right\} + 8n^{-1}p^{-2}\left\{\mathrm{tr}(\Lambda_p^4) - p^{-1}\mathrm{tr}^2(\Lambda_p^2)\right\}.$$

THEOREM 2. *Suppose that* $n\mathrm{tr}(D_{n,p}^2)/p = O(1)$ *and* $p = O(n^2)$. *Under* H_1 *and Assumption 1,* $\{\tilde{Q} - \mathrm{tr}(D_{n,p}^2)/p - p\delta_{n,p}\}/\tilde{\sigma}_1 \to N(0,1)$ *in distribution as* $p\to\infty, n\to\infty$.

A direct application of Theorem 2 is the consistency of the proposed test.

COROLLARY 1. *Suppose the assumptions in Theorem 2 hold. If* $n\mathrm{tr}(D_{n,p}^2)/p \to \infty$ *, the test* $\tilde{\sigma}_0^{-1}(\tilde{Q} - p\delta_{n,p}) > z_\alpha$ *is consistent against* H_1 *as* $p\to\infty$ *and* $n\to\infty$.

Because $p^{-1}\mathrm{tr}(D_{n,p}^2)$ measures the departure from the null hypotheses for the sphericity hypothesis, this corollary ensures that as long as $p^{-1}\mathrm{tr}(D_{n,p}^2)$ is not shrinking faster than n^{-1}, the tests are asymptotically optimal; the consistency rates implied in Corollary 1 by $n\mathrm{tr}(D_{n,p}^2)/p\to\infty$ indeed in general are suboptimal. Neither John's test nor the sign-based test described here is asymptotically optimal because that their consistency rates are suboptimal. Under certain assumptions like multivariate normality, some rate-optimal tests can be constructed; see Onatski et al. (2011).

The following corollary provides the limiting efficiency comparison with Chen et al. (2010)'s test under multivariate normality. We consider the sequences of local alternatives $H_1 : \Lambda_p = I_p + D_{n,p}$ in which $C_1 \le n\mathrm{tr}(D_{n,p}^2)/p \le C_2$ with two positive constants C_1 and C_2.

COROLLARY 2. *Suppose the assumptions in Theorem 2 hold. Under multinormal distributions, the sign-based test* $\tilde{Q}$ *is asymptotically as efficient as Chen et al. (2010)'s test.*

When the dimension p is fixed, it can be expected that the proposed test, using only the direction of an observation from the origin, should be outperformed by the test constructed with original observations like that of Chen et al. (2010). However, as $p\to\infty$ as $n\to\infty$, the disadvantage diminishes. Theoretically comparing the proposed test with Chen et al. (2010)'s test under general multivariate distributions turns out to be difficult. This is because the asymptotic validity of Chen et al. (2010)'s test relies on model (2), while an elliptical assumption is required in Theorems 1 and 2. Thus, in Section 3, we compare these two methods using simulation.

3. NUMERICAL STUDIES

Three well-known multivariate elliptical distributions are considered: (I) the standard multivariate normal; (II) the standard multivariate t with four degrees of freedom, $t_{p,4}$; and (III) mixtures of two multivariate normal densities $\kappa f_p(\mu, I_p) + (1-\kappa) f_p(\mu, 9I_p)$, where $f_p(\cdot;\cdot)$ is the p-variate multivariate normal density. The value κ is chosen to be 0.8. The model (2) is also included, allowing us to have a more broader picture of the robustness and efficiency of the proposed method. Following Chen et al. (2010), we choose $\Gamma = I_p$ and for each Z_i, p independent identically distributed random variables Z_{ij}'s are generated. Two distributions for the Z_{ij} are considered: (IV) the standardized Gamma(4, 0.5) distribution considered by Chen et al. (2010); (V) the standardized t distribution with four degrees of freedom, t_4. The random vectors generated from scenarios (IV) and (V) are not elliptically distributed, while neither scenario (II) nor (III) corresponds to model (2); see the proof in the Supplementary Material. Only scenario (I) satisfies both the elliptical assumption and the form of model (2).

The settings of the combinations of p and n in Chen et al. (2009) are adopted here. The sample size n is chosen as 20, 40, 60 and 80. Dimensions $p = 38, 55, 89, 181, 331$ and 642 are considered for each sample size. Set $A = \mathrm{diag}(2^{1/2} 1_{[vp]}, 1_{p-[vp]})$, where $[x]$ denotes the integer truncation of x. To evaluate the size and power of the sphericity test, we generate multivariate random vectors from scenarios (I)-(V) say, Y_is, and then obtain the observations $X_i = AY_i$. Three levels of v were considered: 0, 0.125 and 0.25. We compare the proposed test, called the bias-corrected sign test hereafter, with three testing methods for sphericity: the sign test (3) studied by Hallin & Paindaveine (2006); the test proposed by Ledoit & Wolf (2002); and the sphericity test proposed by Chen et al. (2010).

Tables 1 and 2 report empirical sizes and power under scenarios (I)-(III) with $n = 40, 80$ and $p = 55, 181, 642$. The complete simulation results can be found in the supplementary Material. For each experiment we run 2,500 replications. Ledoit & Wolf's (2002) test is not included in Table 2 because it is applicable only for normal distribution and it encounters serious size distortion. Table 1 shows the sign-based test without the bias-correction has type I error rates much larger than nominal levels, especially for large p and small n. This is consistent with the asymptotic analysis in Section 2. For this reason we do not report its results in Table 2.

Table 1. *Empirical sizes and power (%) of four tests at 5% significance for multivariate normal random vectors (scenario I). T_{BCS}: the proposed bias-corrected sign test; T_{HP}: the sign test (3) studied by Hallin & Paindaveine (2006); T_{LW}: Ledoit & Wolf's (2002) test; T_{CZZ}: Chen et al. (2010)'s test.*

	v	T_{LW}	T_{CZZ}	T_{HP}	T_{BCS}	T_{LW}	T_{CZZ}	T_{HP}	T_{BCS}	T_{LW}	T_{CZZ}	T_{HP}	T_{BCS}
		$p = 55$				$p = 181$				$p = 642$			
	0.000	4.8	5.6	17	4.9	4.8	5.8	74	4.9	5.8	6.2	100	5.1
$n = 40$	0.125	44	44	65	41	50	50	98	47	50	52	100	49
	0.250	68	69	84	64	72	72	99	68	74	74	100	72
	0.000	4.6	5.5	9.0	4.7	4.9	6.0	31	5.2	5.2	5.2	100	5.1
$n = 80$	0.125	87	87	90	84	93	92	99	91	95	95	100	94
	0.250	99	99	99	99	100	100	100	100	100	100	100	100

Under scenario (I), the empirical sizes of the tests converge to the nominal level as p and n increase together. Ledoit & Wolf's (2002) test performs best in both terms of size and power, as we could expect because it is based on normality. The bias-corrected sign test has similar empirical sizes. The power of the proposed test is largely dependent on the sample size and levels of v as they determine $\mathrm{tr}(D_{n,p}^2)$. Chen et al. (2010)'s test outperforms the bias-corrected sign test in terms of power in most cases, but when n and p increase the advantage tends to vanish. When $n = 80$ and $p \geq 181$, the two tests are largely comparable and the difference is at least partly due to the unequal size of the test. This can be understood from Corollary 2.

Table 2. *Empirical size and power (%) at 5% significance for multivariate t random vectors (scenario II) and mixtures of multivariate normal random vectors (scenario III)*

		Scenario (II)				Scenario (III)			
		$n = 40$		$n = 80$		$n = 40$		$n = 80$	
	v	T_{CZZ}	T_{BCS}	T_{CZZ}	T_{BCS}	T_{CZZ}	T_{BCS}	T_{CZZ}	T_{BCS}
	0.000	12	4.7	14	5.1	16	5.4	16	5.4
$p = 55$	0.125	34	39	61	84	37	40	57	82
	0.250	46	65	76	98	44	64	75	98
	0.000	12	5.9	14	4.4	15	5.0	16	4.5
$p = 181$	0.125	36	47	63	92	38	46	60	92
	0.250	47	69	79	99	44	69	79	100
	0.000	12	6.6	13	5.7	16	6.2	16	5.4
$p = 642$	0.125	35	50	61	94	35	50	62	94
	0.250	46	71	77	100	45	72	79	100

Table 2 gives simulation values with the other two elliptical distributions. The proposed test can achieve the nominal size, but Chen et al. (2010)'s test has considerable bias in size. Even worse, the empirical sizes of Chen et al. (2010)'s test hardly improve when n and/or p increase. The bias-corrected sign test is more efficient under H_1 in both scenarios in the sense that even when the observed size is much smaller than that of Chen et al. (2010)'s test, the empirical power increases much faster with v increases. When $n = 80$, the proposed test performs uniformly much better than Chen et al. (2010)'s test, and the difference is quite remarkable. Certainly, this is not surprising as neither $t_{p,4}$ nor mixture of multivariate normal distributions belongs to model (2) on which the validity of Chen et al.'s (2010) test depends much.

Empirical sizes and power under scenarios (IV) and (V) are given in the Supplementary Material. Although our test is not asymptotically justified under model (2), it is quite robust to the two distributions that belong to that model. Its sizes are close to nominal and even closer than those of Chen et al. (2010)'s test. With respect to the performance under alternatives, our test has quite good power and generally performs similarly to Chen et al. (2010)'s test, although its sizes are usually smaller. These results suggest that the proposed test is quite robust and efficient in testing sphericity, especially for heavy-tailed or skewed distributions.

4. Concluding remarks

The bias-correction procedure takes advantage of the relatively simple form of multivariate-sign-based tests for sphericity. However we believe that this procedure can be extended to more general elliptical distributions with $\Sigma_p = \mathrm{diag}\{\sigma_{11}, \ldots, \sigma_{pp}\}$ with unknown σ_{ii}'s. Moreover, Theorem 1 is established under $p = O(n^2)$. The issue preventing p from growing faster than n^2 is that a higher-order expansion is required for bias-correction.

Hallin & Paindaveine (2006) proposed a family of signed-rank test statistics based on the sign vectors U_i and ranks of the moduli R_i. Their tests appear to be asymptotically optimal at given target densities. Deriving similar bias-corrected procedures for those tests is hard due to their complicated construction and deserves some future research. Furthermore, tests based on symmetrised- spatial-signs and spatial-ranks derived in Sirkiä et al. (2009) also warrant future study in a high-dimensional setting.

Supplementary Material

The Supplementary Material contains the proofs of Theorems 1 and 2, the calculation of $E(\Delta_i)$, additional simulation results and some other technical details.

Acknowledgment

The authors would like to thank the Editor, Associate Editor and two anonymous referees for their many helpful comments that have resulted in significant improvements in the article. This research was supported by the NNSF of China Grants No.11001138, 11071128, 11131002, 11101306, 70931004, the RFDP of China Grant No. 20110031110002, Foundation for the Author of National Excellent Doctoral Dissertation of PR China H0512101 and New Century Excellent Talents in University NCET-12-0276.

REFERENCES

BARINGHAUS, L. (1991). Testing for spherical symmetry of a multivariate distribution. *Ann. Statist.* **19**, 899–917.

CHEN, S. X., PENG, L. & QIN, Y. L. (2009). Effects of data dimension on empirical likelihood. *Biometrika* **96**, 1–12.

CHEN, S. X. & QIN, Y. L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Ann. Statist.* **38**, 808–835.

CHEN, S. X., ZHANG, L. X. & ZHONG, P. S. (2010). Tests for high-dimensional covariance matrices. *J. Am. Statist. Assoc.* **105**, 801–815.

GHOSH, S. K. & SENGUPTA, D. (2001). Testing for proportionality of multivariate dispersion structures using interdirections. *J. Nonparametr. Statist.* **13**, 331–349.

HALLIN, M. & PAINDAVEINE, D. (2006). Semiparametrically efficient rank-based inference for shape. I: Optimal rank-based tests for sphericity. *Ann. Statist.* **34**, 2707–2756.

HJORT, H. L., MCKEAGUE, I. W. & VAN KEILEGOM, I. (2009). Extending the scope of empirical likelihood. *Ann. Statist.* **37**, 1079–1115.

LEDOIT, O. & WOLF, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Ann. Statist.* **30**, 1081–1102.

JOHN, S. (1971). Some optimal multivariate tests. *Biometrika* **59**, 123–127.

JOHN, S. (1972). The distribution of a statistic used for testing sphericity of normal distributions. *Biometrika* **59**, 169–173.

MARDEN, J. & GAO, Y. (2002). Rank-based procedures for structural hypotheses on covariance matrices. *Sankhyā Ser. A* **64**, 653–677.

MAUCHLY, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *Ann. Math. Statist.* **11**, 204–209.

MÖTTÖNEN J. & OJA, H. (1995). Multivariate spatial sign and rank methods. *J. Nonparametr. Statist* **5**, 201–213.

MUIRHEAD, R. J. & WATERNAUX, C. M. (1980). Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations. *Biometrika* **67**, 31–43.

OJA, H. (2010). *Multivariate Nonparametric Methods with R.* Springer, New York.

ONATSKI, A., MOREIRA, M. J. & HALLIN, M. (2013). Asymptotic power of sphericity tests for high-dimendional data. *Ann. Statist.* **41**, 1204–1231.

SCHOTT, J. R. (2005). Testing for complete independence in high dimensions. *Biometrika* **92**, 951–956.

SIRKIÄ, S., TASKINEN, S., OJA, H. & TYLER, D. E. (2009). Tests and estimates of shape based on spatial signs and ranks. *J. Nonparametr. Statist.* **21**, 155–176.

TANG, C. Y. & LENG, C. (2010). Penalized high-dimensional empirical likelihood. *Biometrika* **97**, 905–920.

TYLER, D. E. (1987). Statistical analysis for the angular central Gaussian distribution on the sphere. *Biometrika* **74**, 579–589.