Smooth Test for Multivariate Normality
Yan Su\textsuperscript{a}, Ya-Ping Huang\textsuperscript{b}

School of Mathematics and Physics, North China Electric Power University, Baoding, China
\textsuperscript{a}suyanhd@163, \textsuperscript{b}huangyaping01@126.com

Keywords: Multivariate normality; Spherical harmonic; Smooth test; Power simulation.

Abstract. Based on the smooth test for uniformity on the surface of a unit sphere, a new test for multivariate normality is presented. The test statistic does not depend on the parameters in the multivariate normal distribution. We obtain the asymptotic null distribution of the test statistic. The test procedure for testing multivariate normality is given. Power simulation shows that the test based on all spherical harmonics of degree less than 3 offers good power against a wide variety of alternatives.

Introduction
The basic distribution in multivariate analysis is the multivariate normal distribution. Hence, testing for multivariate normality is often necessary in implementing multivariate analysis. Based on a characterization of the multivariate normal distribution, the Q-Q plots for testing multivariate normality were derived\cite{1}. Székely and Rizzo proposed a goodness-of-fit test for multivariate normality based on Euclidean distance between sample elements with estimated parameters\cite{2}.

The idea of smooth test is that the null hypothesis density is embedded in a parameter smooth alternative, such that when the vector of parameter $\eta = 0$, the alternative is the same as the hypothesized distribution. In this paper, The smooth test for multivariate normality is proposed, the test is based on spherical harmonics. The idea of the test is that the goodness-of-fit test for the multivariate normal distribution can be translated into the goodness-of-fit test for the uniform distribution on $\Omega_d$, the surface of a unit sphere. Therefore, the smooth test for uniformity on $\Omega_d$ can be used\cite{3}. Statistical analysis and power simulation indicated that smooth tests for $U(\Omega_d)$ based on spherical harmonics of degree at most 2 are generally powerful.

An advantage of the new test is that the estimation of the parameters of $N_d(\mu, \Sigma)$ can be avoided. Thus the test power can be increased.

The paper is organized as follows. In Section 2, we introduce some definitions and some lemmas. In Section 3, the smooth test statistic for multivariate normal distribution is presented. The asymptotic null distribution of the test statistic is obtained and the test procedure is proposed. Section 4 presents a Monte Carlo power study. Some discussions and future work are given in the last section. The proof of Theorem 1 is postponed to Appendix.

Definitions and some lemmas
Let $\Omega_d$ denote the surface of the unit sphere centered at the origin in $\mathbb{R}^d$, which we denote simply by $\Omega$ dropping the suffix. Let $U(\Omega_d)$ denote the uniform distribution on $\Omega_d$. The fundamental distribution on $\Omega_d$ is $U(\Omega_d)$. Let $U^{(d)}$ denote a random vector distributed uniformly on $\Omega_d$ in $\mathbb{R}^d$ and let $\|\|$ denotes the Euclidean norm.

Definition 1\cite{4} A $d \times 1$ random vector $X^{(d)}$ is said to have a spherical distribution if

$$X^{(d)} = RU^{(d)},$$

(1)

© 2015. The authors - Published by Atlantis Press
for some random variable $R \geq 0$ which is independent of random vector $U^{(d)}$. Here $=\text{signifies the two sides have the same distribution.}

**Definition 2**[4] Let $A$ be a $d \times k$ matrix of rank $k$ and let $\Sigma = AA^\top$. A $d \times 1$ random vector $X^{(d)}$ is said to have a multivariate normal distribution $N_d(\mu, \Sigma)$ with parameters $\mu$ and $\Sigma$ if

$$X^{(d)} = \mu + RAU^{(k)},$$

with $R^2 \chi^2_k$ (chi-square distribution with $k$ degrees of freedom), where the random variable $R \geq 0$ is independent of $U^{(k)}$.

**Lemma 1**[4] If $d \times 1$ random vector $X$ has a spherical distribution then $X / \|X\| \sim U(\Omega_d)$. Moreover, $\|X\|$ and $X / \|X\|$ are independent.

**Lemma 2**[5] Assume that $X_1^{(d)}, X_2^{(d)}, \ldots, X_n^{(d)}$ are i.i.d. $N_d(0, \Sigma)$. Define the random vectors

$$Y_i^{(d)} = \frac{1}{\sqrt{i(i+1)}}[X_1^{(d)} + \cdots + X_i^{(d)} - iX_{i+1}^{(d)}], \quad i = 1, \ldots, n-1,$$

Then $Y_1^{(d)}, Y_2^{(d)}, \ldots, Y_{n-1}^{(d)}$ are i.i.d. $N_d(0, \Sigma)$.

**Lemma 3**[5] Let $Y_1^{(d)}, Y_2^{(d)}, \ldots, Y_{n-1}^{(d)}$ be i.i.d. $N_d(0, \Sigma)$. Let

$$S_k = \sum_{i=1}^{k} Y_i^{(d)}Y_i^{(d)^\top}, \quad Z_k^{(d)} = S_k^{-1/2}Y_k^{(d)} = \left\|Z_k^{(d)}\right\|^2,

$$

for $k = d+1, \ldots, n-1$, where $S_k^{1/2}$ stands for the positive definite square root of $S_k$. Then

(a) $Z_{d+1,k}^{(d)}, \ldots, Z_{n-1}^{(d)}$ are mutually independent.

(b) $Z_k^{(d)}(k \geq d + 1)$ has a symmetric Pearson type II distribution.

(c) $\lambda_k \sim \text{Beta}(d/2, (k-d)/2)$ a beta distribution.

Let $t = (t_1, t_2, \ldots, t_d)^\top$ denotes a typical point in $R^d$. For $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)^\top$ a multi-index, define

$$t^\alpha = t_1^{\alpha_1}t_2^{\alpha_2} \cdots t_d^{\alpha_d}, \quad |\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_d, \quad D^\alpha = D_1^{\alpha_1}D_2^{\alpha_2} \cdots D_d^{\alpha_d},$$

where $D_j^{\alpha_j}$ denotes the $\alpha_j$th partial derivative with respect to the $j$th coordinate variable. The collection of all spherical harmonics of degree $m$ will be denoted by $H_m(\Omega)$.

Let $\sigma$ be the normalized surface-area measure on $\Omega_d$ (so that $\sigma(\Omega_d) = 1$). Let the inner product on $H^d_m(\Omega)$ be defined by

$$\int_{\Omega_d} p(u)q(u)d\sigma(u), \quad p, q \in H^d_m(\Omega).$$

**Lemma 4**[6] If $d > 2$ then the set $\{D^\alpha \|^{2-d}\alpha| = m \text{ and } \alpha_1 \leq 1\}$ is a vector space basis of $H^d_m(\Omega)$, where $D^\alpha$ is defined in (5).

A complete orthonormal basis(CONB) for $H_k(\Omega)$ can be obtained by Schmidt's orthogonalization.

**Lemma 5**[3] Let $N_{k,d} = \dim[H_k(\Omega)]$. Let

$$B_k = \{V_{k,j}(u) \in H_k(\Omega), j = 1, 2, \ldots, N_{k,d}\}$$

be a CONB for $H_k(\Omega)$. Let $B = \{B_k : k = 0, 1, \ldots, m\}$. Then $B$ is a set of orthonormal functions.

Let $\land = B \setminus B_0$ and let us denote $N = \#(\land)$, we have

$$N = d + \sum_{k=2}^{m} N_{k,d}$$

where $\#$ denotes cardinality. The elements of $\land$ are arranged with $k = 1, \ldots, m$. The set $\land$ can be written as $\land = \{h_i(u) : i = 1, \ldots, N\}$ with

$$h_i(u) = V_{1i}(u), \quad \cdots, h_N(u) = V_{mN_{k,d}}(u).$$
Let \( f(\cdot) \) be a density on \( \Omega_d \) and let \( a_d \) denote the surface-area of \( \Omega_d \). Let
\[
f_0(u) = \frac{1}{a_d}, u \in \Omega_d.
\]
Then \( f_0 \) is uniform on \( \Omega_d \). Consider the null hypothesis
\[
H_0 : f(u) = f_0(u).
\]
A smooth alternative probability density function can be defined by
\[
g_N(u, \eta) = C(\eta) \exp\left\{ \sum_{i=1}^{N} \eta_i h_i(u) \right\}, \quad \eta = (\eta_1, \eta_2, \ldots, \eta_N)'.
\]
(7)

**Lemma 6** [3] (Smooth test for \( U(\Omega_d) \)) Let \( U_1^{(d)}, \ldots, U_n^{(d)} \) be a random sample from \( g_N(u, \eta) \) defined in (7). Then
(a). The score statistic \( SR_N \) for testing \( H_0 : \eta = 0, H_1 : \eta \neq 0 \) is
\[
SR_N = \sum_{i=1}^{N} W_i^2, \quad W_i = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} h_j(U_i^{(d)}).
\]
(8)
(b). Under \( H_0 : \eta = 0 \), \( SR_N \) is asymptotically distributed as \( \chi^2_N \) random variable, where \( \chi^2_N \) represents chi-square distribution on \( N \) degrees of freedom.

**Remark 1** Let \( SR_n(B_1) \) and \( SR_n(B_2) \) denote \( SR_N \) in (8) constructed by \( B_1 \) and \( B_2 \), respectively. \( SR_n(B_1) \) can be used to detect the center of mass of \( U(\Omega_d) \), \( SR_n(B_2) \) can be used to detect the moment of inertia of \( U(\Omega_d) \) [3].

**Definition 3** [7] The random vector \( dX \) is said to have a Langevin (L) distribution on \( \Omega_d \) if the probability density function of \( dX \) is
\[
f_L(x, \kappa) = \frac{1}{a_d(\kappa)} \exp\{\kappa x\}, \kappa \geq 0, x \in \Omega_d,
\]
(9)
where \( a_d(\kappa) \) is a normalizing constant inserted so \( f_L(x, \kappa) \) integrates to one.

**Definition 4** [7] The random vector \( dX \) is said to have a Scheidegger-Watson (S-W) distribution on \( \Omega_d \) if the probability density function of \( dX \) is
\[
f_{SW}(x, \lambda) = \frac{1}{b_d(\lambda)} \exp\{\lambda (x')^2\}, x \in \Omega_d,
\]
(10)
where \( \lambda \in R \) and \( b_d(\lambda) \) is a normalizing constant inserted so \( f_{SW}(x, \lambda) \) integrates to one.

**Remark 2** The L distribution and S-W distribution are two particular cases of rotationally symmetric distributions on \( \Omega_d \). If \( \kappa > 0 \), The L distribution is unimodal. If \( 2\lambda > d - 3 \), the S-W distribution is symmetric bimodal and if \( 2\lambda < d - 3 \), the S-W distribution is symmetric annular. If \( \kappa \to 0 \) (or \( \lambda \to 0 \)), either distribution becomes uniform on \( \Omega_d \).

**Smooth test for multivariate normal distribution**

Let \( X_1^{(d)}, X_2^{(d)}, \ldots, X_n^{(d)} \) be an i.i.d. sample from a population with a distribution function (d.f.) \( F(x), x \in R^d \). We want to test
\[
H_0 : F(x) \text{ is the d.f. of a normal distribution } N_d(\mu, \Sigma),
\]
(11)
where \( \mu \) and \( \Sigma \) are unknown parameters. The goodness -of- fit test for the multivariate normal distribution is based on the goodness-of-fit test for \( U(\Omega_d) \).

**Theorem 1** Let \( X_1^{(d)}, X_2^{(d)}, \ldots, X_n^{(d)} \) be an i.i.d. sample from the \( N_d(\mu, \Sigma) \) distribution and \( Y_1^{(d)}, Y_2^{(d)}, \ldots, Y_m^{(d)} \) be defined by (3). Let \( Z_1^{(d)}, Z_2^{(d)}, \ldots, Z_{m-1}^{(d)} \) be defined by (4) and let
Let \( N \) be defined by (7). Let \( F_k(\cdot) \) be the distribution function (d.f.) of \( \lambda_k \) in (4). Let

\[
\varsigma_N = \sum_{i=1}^{N} \psi_i^2, \quad \tau_k = F_k(\lambda_k),
\]

and

\[
\varsigma_0 = \frac{12}{n-d-1} \left[ \sum_{i=d+1}^{n-1} (\tau_i - \frac{1}{2})^2 \right],
\]

\[
\psi = \varsigma_0 + \varsigma_N.
\]

Then \( \psi \to \chi^2_{N+1}, n \to \infty \), where \( \chi^2_{N+1} \) is the chi-square distribution with \( N+1 \) degrees of freedom.

The following test procedure can illustrate our proposed test for \( N(\mu, \Sigma) \):

1. Compute the values of \( Y_1^{(d)}, Y_2^{(d)}, \ldots, Y_n^{(d)} \) in (3) and \( Z_1^{(d)}, Z_2^{(d)}, \ldots, Z_{n-1}^{(d)} \) in (4), respectively.

2. Compute the values of \( U_1^{(d)}, U_2^{(d)}, \ldots, U_{n-1}^{(d)} \) in (12).

3. Compute the value of \( \varsigma_k \) in (13).

4. Compute the values \( \varsigma_0 \) in (14).

5. Compute the value of \( \psi \) in (15).

6. The multivariate normality is rejected for large value of \( \psi \).

**Remark 3** Theorem 1 indicates that if \( X_1^{(d)}, \ldots, X_n^{(d)} \) are i.i.d. \( N_d(\mu, \Sigma) \), then \( U_1^{(d)}, U_2^{(d)}, \ldots, U_{n-1}^{(d)} \) are i.i.d. \( U(\Omega_d) \). Thus, the critical values of \( \psi \) in (15) can be estimated by Monte Carlo simulation with \( \mu=0 \) and \( \Sigma=I_d \), where \( I_d \) is the \( d \times d \) unit matrix.

**Remark 4** Since \( Z_k^{(d)} \) in (4) is spherical distributed, it follows from lemma 1 that \( Z_k^{(d)} \) may be expressed as

\[
Z_k^{(d)} = \xi_k \cdot U_k^{(d)}, \quad \xi_k = \left\| Z_k^{(d)} \right\|, \quad k = d + 1, \ldots, n-1.
\]

\( \varsigma_0 \) and \( \varsigma_N \) correspond to \( \xi_k \) and \( U_k^{(d)} \) respectively. Thus the test statistic \( \psi \) can be used for testing multivariate normality.

**Power simulations**

The power of a statistical test is the probability of correctly rejecting a false null hypothesis.

Let \( X_1, X_2, \ldots, X_n \) be a random sample from a continuous distribution function on \( R^d \). The null hypothesis is that the sample follows the \( N_d(\mu, \Sigma) \) distribution. In this section, samples of sizes \( n = 20, 50, 80, 100 \), and 150 with replications of 5000 are generated. The critical values (percentiles) are estimated by Monte Carlo simulation of 20000 set of \( N_d(0, I_d) \) samples for each sample size \( n \). The significance level \( \alpha \) is 0.05.

Suppose \( \text{Exp}(\beta) \) refers to the exponential distribution, with the probability density function \( f(x) = \beta^{-1} e^{-x/\beta}, x > 0 \), zero otherwise, where \( \beta > 0 \). Let \( d = 4 \) and let the nonnegative random variable \( R \) be \( \text{Exp}(\beta) \) distribution. Let the alternatives be
\[
X^{(d)} = \mu + RA'\gamma^{(d)}, \mu = (0.01,1,50,1000), \quad A = \begin{pmatrix}
10 & 2 & 2 & 2 \\
2 & 10 & 2 & 2 \\
2 & 2 & 10 & 2 \\
2 & 2 & 2 & 10
\end{pmatrix},
\]

where \(\gamma^{(d)}\) is the rotationally symmetric distribution on \(\Omega_d\). The alternatives above represent a wide variety of distributions in \(R^d\).

Let \(f_L(x,\kappa)\) and \(f_{SW}(x,\lambda)\) be defined in (9) and (10), respectively. Let \(\psi(B_1),\psi(B_2)\) and \(\psi(B_1 \cup B_2)\) denote the test statistic \(\psi\) in (15) constructed by \(B_1, B_2\) and \(B_1 \cup B_2\) in Lemma 5, respectively.

**Conclusion**

The goodness-of-fit test for multivariate normality should be done before using many statistical analysis. Classical theory on Bayesian classifiers assume that the populations are normally distributed.

Our proposed statistic \(\psi\) in (15) for testing \(N_d(\mu,\Sigma)\) distribution is based on the statistic \(SR_N\) in (8) for testing \(U(\Omega_d)\) distribution. For \(n \leq 100\), the power simulation shows that the \(\psi(B_1 \cup B_2)\) test is superior to the \(\psi(B_1)\) and \(\psi(B_2)\) tests against the alternatives considered. When the sample size is large, the powers of the tests are very close or equal 100%. We conclude that \(\psi(B_1 \cup B_2)\) provides a powerful smooth test of multivariate normality. The power simulation revealed that the combination test of the center of mass and the moment of inertia should be applied[8]. Consequently, \(\psi(B_1 \cup B_2)\) can be used for testing multivariate normality against a wide range of alternatives.

The \(\psi(B_1)\) and \(\psi(B_2)\) are called components of the test statistic \(\psi\). The components are

---

From Table 1, Table 2 and Table 3, the following statements can be asserted:

(1) The test power of \(\psi\) increases with increasing the sample size \((n = 20 \rightarrow 150)\). This shows that the smooth tests for multivariate normality are consistent.

(2) For \(n \leq 100\), the test statistic \(\psi(B_1 \cup B_2)\) offers the highest power against the alternatives based on the unimodal type L distribution, the bimodal and annular type S-W distributions. For \(n = 150\), the \(\psi(B_1), \psi(B_2)\) and \(\psi(B_1 \cup B_2)\) have nearly the same power 100% against the alternatives considered.
asymptotically independent and they may indicate what alternative distribution would fit the data. A more detailed power simulation will be carried out to reveal the test power of the components against particular alternatives.

Appendix

Proof. Let \( \xi_k = \left\| Z_k^{(d)} \right\|, k = d + 1, \cdots, n-1 \). By lemma3, \( Z_k^{(d)}(k \geq d + 1) \) has a symmetric Pearson type II distribution which is a spherical distribution and \( Z_{d+1}^{(d)}, \cdots, Z_{n-1}^{(d)} \) are mutually independent.

By lemma 1, \( U_k^{(d)} \) and \( \xi_k \) are independent, where \( U_k^{(d)} \) is defined by (12). Thus, \( \hat{\lambda}_{d+1}, \cdots, \hat{\lambda}_{n-1} \) are i.i.d. \( \text{Beta}(d/2,(k-d)/2) \), \( U_{d+1}^{(d)}, \cdots, U_{n-1}^{(d)} \) are i.i.d. \( \sim U(\Omega_d) \), \( \{U_k^{(d)}\}_{k=d+1} \) and \( \{\xi_k\}_{k=d+1} \) are independent. Hence, \( \zeta_N \) in (13) and \( \zeta_0 \) in (14) are independent.

By the central limit theorem,

\[
\sqrt{2m} \left( \frac{1}{m} \sum_{i=d+1}^{n-1} \tau_i - \frac{1}{2} \right) \rightarrow N(0,1), \quad n \rightarrow \infty,
\]

where \( \tau_i \) is defined in (14) and \( m = n - d - 1 \). By (15) and lemma 6, we have

\[
\zeta_0 \rightarrow \chi^2_{1}, \quad \zeta_N \rightarrow \chi^2_{N}, \quad n \rightarrow \infty.
\]

Thus \( \psi \rightarrow \chi^2_{N+1}, n \rightarrow \infty \). This completes the proof.

References
