Statistics and Probability Letters 137 (2018) 105–112

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Remarks for the singular multivariate skew-normal distribution and its quadratic forms

Baokun Li [a], Weizhong Tian [b,c,*], Tonghui Wang [d]

[a] School of Statistics, Southwestern University of Finance and Economics, China
[b] Department of Mathematical Sciences, Eastern New Mexico University, USA
[c] School of Science, Xi'an University of Technology, China
[d] Department of Mathematical Sciences, New Mexico State University, USA

ARTICLE INFO

Article history:
Received 29 October 2017
Received in revised form 7 January 2018
Accepted 10 January 2018
Available online 2 February 2018

Keywords:
Multivariate skew normal distribution
Independent sub vectors
Moment generating function
Noncentral skew chi square distribution

ABSTRACT

Under the singular multivariate skew normal (SMSN) setting, we showed that, in this paper, the necessary and sufficient conditions for independence of two sub-vectors given in Young et al. (2017) are equivalent to the results in Wang et al. (2009). In addition, the distribution of quadratic form of SMSN random vector is derived, with this new definition of the noncentral skew chi-square distribution. Several examples are given to illustrate our main results.

1. Introduction

In many real world problems, the assumptions of normality are violated as the data possess some level of skewness. The class of skew normal distributions is an extension of the normal distribution, allowing for the presence of skewness was introduced by Azzalini (1985) and Azzalini and Dalla (1996).

The random variable X is said to have the skew normal distribution with the location parameter $\mu \in \Re$, scale parameter $\sigma > 0$, and skewness parameter $\alpha \in \Re$, denoted by $X \sim SN(\mu, \sigma^2, \alpha)$, if its probability density function (pdf) is of the form

$$f_X(x) = 2\phi(x, \mu, \sigma^2)\Phi\left\{\alpha\left(\frac{x-\mu}{\sigma}\right)\right\}, \tag{1.1}$$

where $\phi(\cdot)$ is the normal pdf and $\Phi(\cdot)$ is cumulative distribution function of standard normal distribution respectively. The class of multivariate skew normal distributions has been studied by Azzalini and Capitanio (1999). The random vector $\mathbf{x} \in \Re^n$ is said to have a **multivariate skew normal distribution** with location parameter $\boldsymbol{\mu} \in \Re^n$, positive definite matrix $\boldsymbol{\Sigma} \in M_{n\times n}$, and skewness vector $\boldsymbol{\alpha} \in \Re^n$, and denoted by $\mathbf{x} \sim SN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$, if its pdf is given by

$$f(\mathbf{x}) = 2\phi_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})\Phi\left(\boldsymbol{\alpha}'(\mathbf{x}-\boldsymbol{\mu})\right), \tag{1.2}$$

where $\phi_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the n-dimensional normal density with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Also Wang et al. (2009) proposed a new family of multivariate skew normal distribution and a new version of Cochran's theorem based on it. Let $\mathbf{z} \sim SN_k(\mathbf{0}, \mathbf{I}_k, \boldsymbol{\alpha})$, where $\mathbf{I}_k$ is the identity matrix order k. The distribution of $\mathbf{y} = \boldsymbol{\mu} + \mathbf{B}'\mathbf{z}$ is called multivariate skew

* Corresponding author at: Department of Mathematical Sciences, Eastern New Mexico University, USA.
E-mail addresses: bali@swuef.edu.cn (B. Li), weizhong.tian@enmu.edu (W. Tian), twang@nmsu.edu (T. Wang).

https://doi.org/10.1016/j.spl.2018.01.008
0167-7152/

normal with the location parameter $\boldsymbol{\mu} \in \Re^n$, the scale parameter $\mathbf{B} \in \Re^{k\times n}$, and the shape parameter $\boldsymbol{\alpha} \in \Re^k$, denoted by $\mathbf{y} \sim \mathcal{SN}_n(\boldsymbol{\mu}, \mathbf{B}, \boldsymbol{\alpha})$. The moment generating function of $\mathbf{y}$ is given by

$$M_{\mathbf{y}}(\mathbf{t}) = 2\exp\left\{\mathbf{t}'\boldsymbol{\mu} + \frac{\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}{2}\right\}\Phi\left\{\frac{\boldsymbol{\alpha}'\mathbf{B}\mathbf{t}}{(1+\boldsymbol{\alpha}'\boldsymbol{\alpha})^{1/2}}\right\} \tag{1.3}$$

and the density function of $\mathbf{y}$, if it exists, is given by,

$$f(\mathbf{y};\boldsymbol{\mu},\mathbf{B},\boldsymbol{\alpha}) = 2\phi_k(\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma})\,\Phi\left\{\frac{\boldsymbol{\alpha}'\mathbf{B}\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})}{\left[1+\boldsymbol{\alpha}'(I_k-\mathbf{B}\boldsymbol{\Sigma}^{-1}\mathbf{B}')\boldsymbol{\alpha}\right]^{\frac{1}{2}}}\right\}, \tag{1.4}$$

where $\boldsymbol{\Sigma} = \mathbf{B}'\mathbf{B}$.

This paper is organized as follows. Related results of Young et al. (2016, 2017) and other multivariate skew-normal distributions are listed in Section 2. The equivalence between the results of Wang et al. (2009) and results of Young et al. (2017) are discussed in Section 3, especially for the necessary and sufficient conditions for independence. In Section 4 distributions of quadratic forms for singular multivariate skew-normal variables are derived following a new definition of noncentral skew chi square distribution. Several examples are given for the illustration of our main results.

2. Singular multivariate skew-normal distributions

In this section, we would like to adopt the notations of Young et al. (2016, 2017). Let $\Re^{m\times n}$ be the vector space of all $m \times n$ matrices over the real field $\Re$. Let $\mathbf{A}^+$ be the Moore–Penrose inverse of the nonnegative definite $\mathbf{A} \in \Re^{n\times n}$ with rank $r < n$, let $\mathbf{A}^{-\frac{1}{2}}$ be the symmetric matrix such that $\mathbf{A}^{-\frac{1}{2}}\mathbf{A}^{-\frac{1}{2}} = \mathbf{A}^+$, and let $\sigma_i(\mathbf{A})$ be the ith largest eigenvalue of $\mathbf{A}$ and $\det_r(\mathbf{A}) = \prod_{i=1}^r \sigma_i(\mathbf{A})$.

The new version of multivariate skew normal distribution, different from the one given in Azzalini and Dalla (1996), was given by Vernic (2006). A random vector $\mathbf{y} \in \Re^n$ is said to have a multivariate skew normal distribution with skewness parameter $\boldsymbol{\gamma} \in \Re^n$, written $\mathbf{y} \sim MSN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \delta, \boldsymbol{\gamma})$, if its density function is

$$f(\mathbf{y}) = \frac{1}{\Phi(\delta)}\phi_n(\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma})\Phi\left[\frac{\delta+\boldsymbol{\gamma}'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})}{\sqrt{1-\boldsymbol{\gamma}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}}}\right], \tag{2.1}$$

where $\boldsymbol{\mu} \in \Re^n$, $\boldsymbol{\Sigma} \in \Re^{n\times n}$ such that $\boldsymbol{\gamma}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma} < 1$, and $\delta \in \Re$. Similar to the approach of the **singular multivariate normal distribution** given by Van Perlo-ten Kleij (2004), Young et al. (2016) defined the singular multivariate skew normal (SMSN) distribution as follows.

Definition 2.1. A random vector $\mathbf{y} \in \Re^n$ is said to have SMSN distribution with skewness parameter $\boldsymbol{\gamma} \in \Re^n$, denoted as $\mathbf{y} \sim SMSN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \delta, \boldsymbol{\gamma})$, if its density function is

$$f(\mathbf{y}) = \frac{1}{\Phi(\delta)}\phi_n^s(\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma})\Phi\left[\frac{\delta+\boldsymbol{\gamma}'\boldsymbol{\Sigma}^{+}(\mathbf{y}-\boldsymbol{\mu})}{\sqrt{1-\boldsymbol{\gamma}'\boldsymbol{\Sigma}^{+}\boldsymbol{\gamma}}}\right], \tag{2.2}$$

where $\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\gamma}$ are n-vectors, $\boldsymbol{\Sigma} \in \Re^{n\times n}$ is a nonnegative definite matrix with $rank(\boldsymbol{\Sigma}) = r < n$, $\boldsymbol{\gamma}$ in the column space of $\boldsymbol{\Sigma}$, say $\mathcal{C}(\boldsymbol{\Sigma})$, such that $\boldsymbol{\gamma}'\boldsymbol{\Sigma}^+\boldsymbol{\gamma} < 1$, $\delta \in \Re$ and

$$\phi_n^s(\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma}) = (2\pi)^{-n/2}\det_r(\boldsymbol{\Sigma})^{-1/2}\exp\left\{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^+(\mathbf{y}-\boldsymbol{\mu})\right\}.$$

Remark 2.1. In (1.3), let $\mathbf{B} = \boldsymbol{\Sigma}^{1/2}$ and $\boldsymbol{\gamma} = \frac{\boldsymbol{\Sigma}^{1/2}\boldsymbol{\alpha}}{\sqrt{1+\boldsymbol{\alpha}'\boldsymbol{\alpha}}}$. Then the moment generating function of $\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\mathbf{z}$ is

$$M_{\mathbf{y}}(\mathbf{t}) = 2\exp\left\{\mathbf{t}'\boldsymbol{\mu} + \frac{\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}{2}\right\}\Phi(\boldsymbol{\gamma}'\mathbf{t}), \tag{2.3}$$

which is the moment generating function of $\mathrm{SMSN}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$ given in Young et al. (2016). Thus if we use the moment generating function to define the SMSN distribution, then $SMSN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$ is a special case in the family of multivariate skew normal distribution defined in Wang et al. (2009) with $\mathbf{B} = \boldsymbol{\Sigma}^{1/2}$. More details will be given in next section.

3. Independence between two sub-vectors

First we would like to obtain the density function of $\mathbf{y} \sim \mathcal{SN}_n(\boldsymbol{\mu}, \mathbf{B}, \boldsymbol{\alpha})$ in (1.3) for any $\mathbf{B} \in \Re^{k\times n}$ using moment generating function of $\mathbf{x} \sim SMSN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma})$ for the case of $\delta = 0$.

Theorem 3.1. *Let $\mathbf{y} \sim \mathcal{SN}_n(\boldsymbol{\mu}, \mathbf{B}, \boldsymbol{\alpha})$ with the moment generating function given in (1.3), then the density function of $\mathbf{y}$ is given by (2.2) with $\delta = 0$, $\boldsymbol{\gamma} = \mathbf{B}'\boldsymbol{\alpha}/\sqrt{1+\boldsymbol{\alpha}'\boldsymbol{\alpha}}$ and $\boldsymbol{\Sigma} = \mathbf{B}'\mathbf{B}$.*