EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 3, No. 2, 2010, 128-140
ISSN 1307-5543 – www.ejpam.com

A Matrix Variate Skew Distribution

Deniz Akdemir*[1], Arjun K. Gupta[2]

[1] *Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, Ohio 43403, USA (Now at the Ohio Northern University, Ada, Ohio, 45810, USA)*
[2] *Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, Ohio 43403, USA*

Abstract. Typical multivariate analysis assumes independence among the individual observations as well as elliptical symmetry of distributions. In many situations these assumptions may be too restrictive. This paper studies a class of flexible matrix variate distribution models that can represent both skewed and symmetric distributions which can also account for dependence among individual observations. We derive the moment generating function and study linear and quadratic forms of interest that help understand the properties of these models.

2000 Mathematics Subject Classifications: Primary 62H10, Secondary 62H05.

Key Words and Phrases: Skew Normal Distribution, Normal Distribution, Multivariate Distribution, Matrix Variate Skew Normal Distribution.

1. Introduction

1.1. Distribution of Random Sample

In order to make inferences about the population parameters of a k dimensional parametric distribution, we work with a random sample of n individuals from this population which can be represented by a $k \times n$ matrix X. Typical multivariate analysis assumes independence among the individuals. In many situations this assumption may be too restrictive. For example, many data collection and sample designs involve some overlapping between interviewer workload and the sampling units (clusters). A proportion of the measurement variance which is due to interviewers is reflected to some degree in the sampling variance calculations. In the literature, the variable effects that interviewers have on respondent answers are sometimes labeled the *correlated response variance* [3]. Matrix variate distributions can be used to account for the dependence among individual vector observations.

*Corresponding author.

Email addresses: d-akdemir@onu.edu (D. Akdemir), gupta@bgsu.edu (A.K. Gupta)

Unarguably, the most commonly used family of distributions is the normal family. Under normality assumption, the matrix random variable X will have the following density [6]:

$$\phi_{k\times n}(X;M,AA',B'B)=\frac{etr(-\frac{1}{2}(AA')^{-1}(X-M)(B'B)^{-1}(X-M)')}{(2\pi)^{nk/2}|AA'|^{n/2}|B'B|^{k/2}} \tag{1}$$

where A is a $k\times k$ matrix, B is an $n\times n$ matrix such that AA' and $B'B$ are positive definite; M is a $k\times n$ matrix. Like the multivariate case, the matrix A determines how the variables are related, also the matrix B is introduced to account for dependence among individual observations. In this family, orthogonality of rows of the matrix A or the columns of the matrix B is equivalent to independence of rows or columns of the random matrix X.

In the remainder of this paper we will use $\phi(.)$ and $\Phi(.)$ for the density and the cdf of the normal random variables. When we want to refer to the multivariate or matrix variate forms of these functions, we will use subindices to describe the dimensions. For example, we will write $\Phi_{k\times n}(X;M,A,B)$ to represent the cdf of a matrix variate normal random variable with parameters M, A and B evaluated at X.

1.2. Multivariate Skew Distributions

A k dimensional random vector $\boldsymbol{x}$ with pdf $f(.)$ is *centrally symmetric* about $\mathbf{0}$ if $f(\boldsymbol{t})=f(-\boldsymbol{t})$ for all $\boldsymbol{t}$ in the domain of $\boldsymbol{x}$. In this section, we study a family of multivariate skew symmetric densities generated by centrally symmetric densities.

Theorem 1. *Let $g(.)$ be the k-dimensional jpdf for k independent variables centrally symmetric about $\mathbf{0}$, $H(.)$ be an absolutely continuous cumulative distribution function with $H'(.)$ symmetric about 0, $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\ldots,\alpha_k)'$ be a $k-$dimensional real vector, and $\boldsymbol{e}_j$ for $j=1,2,\ldots,k$ are the elementary vectors of $\mathbb{R}^k$. Then*

$$f(\boldsymbol{y},\boldsymbol{\alpha})=2^k g(\boldsymbol{y})\prod_{j=1}^{k}H(\alpha_j\boldsymbol{e}_j'\boldsymbol{y}) \tag{2}$$

defines a probability density function of $\boldsymbol{y}$.

Proof. First note that $f(\boldsymbol{y})\geq 0$ for all $\boldsymbol{y}\in\mathbb{R}^k$. We need to show that

$$k(\alpha_1,\alpha_2,...,\alpha_k)=\int_{R^k}2^k g(\boldsymbol{y})\prod_{j=1}^{k}H(\alpha_j\boldsymbol{e}_j'\boldsymbol{y})\mathbf{dy}=1.$$

Observe that,

$$\begin{aligned}\frac{\partial}{\partial\alpha_\ell}k(\alpha_1,\alpha_2,...,\alpha_k)&=\int_{R^k}\frac{d}{d\alpha_\ell}2^k g(\boldsymbol{y})\prod_{j=1}^{k}H(\alpha_j\boldsymbol{e}_j'\boldsymbol{y})\mathbf{dy}\\&=\int_{R^k}2^k y_\ell H'(\alpha_\ell\boldsymbol{e}_\ell'\boldsymbol{y})\prod_{j\neq\ell}^{k}H(\alpha_j\boldsymbol{e}_j'\boldsymbol{y})g(\boldsymbol{y})\mathbf{dy}\end{aligned}$$

$$= 0.$$

The first equality is true because of Lebesgue dominated convergence theorem; the last equality is first due to independence by seeing

$$E_{g(\boldsymbol{y})}(y_\ell H'(\alpha_\ell \boldsymbol{e}_\ell' \boldsymbol{y}) \prod_{j\neq \ell}^{k} H(\alpha_j \boldsymbol{e}_j' \boldsymbol{y})) = E_{g(\boldsymbol{y})}(y_\ell H'(\alpha_\ell \boldsymbol{e}_\ell' \boldsymbol{y})) E_{g(\boldsymbol{y})}(\prod_{j\neq \ell}^{k} H(\alpha_j \boldsymbol{e}_j' \boldsymbol{y}))$$

and because $g(.)$ is centrally symmetric about $\boldsymbol{0}$, $y_\ell H'(\alpha_\ell \boldsymbol{e}_\ell' \boldsymbol{y})$ is an odd function of y_ℓ, $E_{g(\boldsymbol{y})}(y_\ell H'(\alpha_\ell \boldsymbol{e}_\ell' \boldsymbol{y})) = 0$.

Hence, $k(\alpha_1, \alpha_2, ..., \alpha_k)$ is constant as a function of α_j for all $j = 1, 2, \ldots, k$; and when all $\alpha_i = 0$, $k(\alpha_1, \alpha_2, ..., \alpha_k) = 1$. This concludes the proof.

We will write $\boldsymbol{z} \sim ss_k^{g,H}(\boldsymbol{z}; \boldsymbol{0}, I_k, \boldsymbol{\alpha})$ or $\boldsymbol{z} \sim ss_k^{g,H}(\boldsymbol{\alpha})$ for a random variable with density given by (2). For the random vector $\boldsymbol{y} = A\boldsymbol{z} + \boldsymbol{\mu}$ where A is a nonsingular matrix we will write $\boldsymbol{y} \sim ss_k^{g,H}(\boldsymbol{z}; \boldsymbol{\mu}, A, \boldsymbol{\alpha})$ or $\boldsymbol{y} \sim ss_k^{g,H}(\boldsymbol{\mu}, A, \boldsymbol{\alpha})$.

In the next theorem, we relate the distribution of the even powers of a skew symmetric random variable to those of its kernel's.

Theorem 2. *Let $\boldsymbol{x}$ be a random vector with probability density function $g(\boldsymbol{x})$, and $\boldsymbol{y}$ be the random vector with probability density function*

$$f(\boldsymbol{y}, \boldsymbol{\alpha}) = 2^k g(\boldsymbol{y}) \prod_{j=1}^{k} H(\alpha_j \boldsymbol{e}_j' \boldsymbol{y})$$

where $g(.)$, $H(.)$ and $\boldsymbol{\alpha}$ are defined as in Theorem 1. Then,

1. *the even moments of $\boldsymbol{y}$ and $\boldsymbol{x}$ are the same, i.e $E(\boldsymbol{y}\boldsymbol{y}')^p = E(\boldsymbol{x}\boldsymbol{x}')^p$ for p even and $E(\boldsymbol{y}'\boldsymbol{y})^m = E(\boldsymbol{x}'\boldsymbol{x})^m$ for any natural number m,*

2. *$\boldsymbol{y}'\boldsymbol{y}$ and $\boldsymbol{x}'\boldsymbol{x}$ have the same distribution.*

Proof. It suffices to show that

$$E(x_1^{n_1} x_2^{n_2} \ldots x_k^{n_k}) = E(y_1^{n_1} y_2^{n_2} \ldots y_k^{n_k})$$

for $n_1, n_2, \ldots, n_k$ even.

Let $\Psi_y(\boldsymbol{t})$ be the characteristic function of $\boldsymbol{y}$. Then,

$$\Psi_y(\boldsymbol{t}) = \int_{R^k} e^{i\mathbf{t}'y} 2^k g(\boldsymbol{y}) \prod_{j=1}^{k} H(\alpha_j \boldsymbol{e}_j' \boldsymbol{y}) \mathbf{dy}. \tag{3}$$

Let $n_1 + n_2 + \ldots + n_k = n$. Taking the n_j^{th} partial derivatives of (3) with respect to t_j for $j = 1, 2, \ldots, k$ and putting $\boldsymbol{t} = \boldsymbol{0}$

$$\begin{aligned}
\frac{\partial^n \psi_{\boldsymbol{y}}(t)}{\partial t_1^{n_1}\partial t_2^{n_2}\dots\partial t_k^{n_k}|_{\boldsymbol{t}=\boldsymbol{0}}} &= \int_{R^k} \frac{\partial^n}{\partial t_1^{n_1}\partial t_2^{n_2}\dots\partial t_k^{n_k}} e^{i\boldsymbol{t}'\boldsymbol{y}} 2^k \\
&\times \prod_{j=1}^{k} H(\alpha_j e_j' \boldsymbol{y}) g(\boldsymbol{y}) \mathbf{dy}|_{\boldsymbol{t}=\boldsymbol{0}} \\
&= \int_{R^k} [e^{i\boldsymbol{t}'\boldsymbol{y}} 2^k i^n \prod_{j=1}^{k} H(\alpha_j e_j' \boldsymbol{y})] \\
&\times [\prod_{\ell=1}^{k} y_\ell^{n_\ell}] g(y) \mathbf{dy}|_{\boldsymbol{t}=\boldsymbol{0}} \\
&= \int_{R^k} [2^k i^n \prod_{j=1}^{k} H(\alpha_j e_j' y)][\prod_{\ell=1}^{k} y_\ell^{n_\ell}] g(y) \mathbf{dy}. \qquad (4)
\end{aligned}$$

Taking derivative of (4) with respect to α_m,

$$\begin{aligned}
&\frac{\partial(\int_{R^k}[2^k i^n \prod_{j=1}^k H(\alpha_j \boldsymbol{e}_j' \boldsymbol{y})][\prod_{\ell=1}^k y_\ell^{n_\ell}] g(\boldsymbol{y}) \mathbf{d}\boldsymbol{y})}{\partial \alpha_m} \\
&= 2^k i^n E_{g(\boldsymbol{y})}[y_m^{(n_m+1)} H'(\alpha_m \boldsymbol{e}_m' \boldsymbol{y})[\prod_{j\neq m} y_\ell^{n_\ell} H(\alpha_j \boldsymbol{e}_j' \boldsymbol{y})] \\
&= 2^k i^n E_{g(\boldsymbol{y})}[y_m^{(n_m+1)} H'(\alpha_m \boldsymbol{e}_m' \boldsymbol{y})] E_{g(\boldsymbol{y})}[\prod_{j\neq m} y_\ell^{n_\ell} H(\alpha_j \boldsymbol{e}_j' \boldsymbol{y})] \\
&= 0
\end{aligned}$$

The first equality is true because of Lebesgue dominated convergence theorem, the second equality due to the independence of components. The last equality is due to the fact that

$$y_m^{(n_m+1)} H'(\alpha_m y_m)$$

is an odd function of y_m and $g(.)$ is centrally symmetric about $\mathbf{0}$.

Therefore, for $n_1, n_2, \dots, n_k$ even, $E(y_1^{n_1} y_2^{n_2} \dots y_k^{n_k})$ is constant as a function of α_m.

If all $\alpha_m = 0$ then $f(\boldsymbol{x}) = g(\boldsymbol{x})$ and therefore

$$E(x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}) = E(y_1^{n_1} y_2^{n_2} \dots y_k^{n_k})$$

Finally,

$$E(x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}) = E(y_1^{n_1} y_2^{n_2} \dots y_k^{n_k})$$

is true for all α_m. The required results follow immediately.

A skew normal density is obtained from normal kernel in the following example.

Example 1. *In Theorem 1 above, let $g(.) = \phi_k(.)$, where $\phi_k(.)$ is the k-dimensional standard normal density function. Also let $H(.)$ and $\boldsymbol{\alpha}$ be defined as in Theorem 1. We can construct a density for k-dimensional joint p.d.f's of the form*

$$f(\boldsymbol{y},\boldsymbol{\alpha}) = 2^k \phi_k(\boldsymbol{y}) \prod_{j=1}^{k} H(\alpha_j \boldsymbol{e}_j' \boldsymbol{y}) \tag{5}$$

This density will be called the generalized skew normal probability density function and will be represented by $sn_k^H(y;\mathbf{0},I,\boldsymbol{\alpha})$. For the random vector $\boldsymbol{y}$ with this density we will write $\boldsymbol{y} \sim sn_k^H(\mathbf{0},I,\boldsymbol{\alpha})$. If $H(.)$ is taken as $\Phi(.)$, the cdf of the standard normal variable, then we will drop the super index H and this defines the skew normal probability density function and the skew normal random vector.

Using Theorem 2 we can relate some properties of the $sn_k(\mathbf{0},I,\boldsymbol{\alpha})$ random vector with its kernel, the standard multivariate normal random vector with density $\phi_k(.)$. Let $\boldsymbol{x} \sim \phi_k(\boldsymbol{x})$, and $\boldsymbol{y} \sim sn_k(\mathbf{0},I,\boldsymbol{\alpha})$. Then,

1. *the even moments of $\boldsymbol{y}$ and $\boldsymbol{x}$ are the same, i.e $E(\boldsymbol{y}\boldsymbol{y}')^p = E(\boldsymbol{x}\boldsymbol{x}')^p$ for p even and $E(\boldsymbol{y}'\boldsymbol{y})^m = E(\boldsymbol{x}'\boldsymbol{x})^m$ for any natural number m,*
2. *$\boldsymbol{y}'\boldsymbol{y}$ and $\boldsymbol{x}'\boldsymbol{x}$ both have χ_k^2 distribution.*

2. Matrix Variate Skew Distributions

Chen and Gupta extend the matrix normal distribution to accommodate skewness in the following form [4]:

$$f_1(X;\boldsymbol{\Sigma},\Psi,\boldsymbol{b}) = c_1^* \phi_{k\times n}(X;\mathbf{0},\boldsymbol{\Sigma},\Psi)\Phi_n(X'\boldsymbol{b};\mathbf{0},\Psi) \tag{6}$$

where $c_1^* = (\Phi_n(\mathbf{0};\mathbf{0},(1+\boldsymbol{b}'\boldsymbol{\Sigma}\boldsymbol{b})\Psi))^{-1}$. A drawback of this definition is that it allows independence only over its rows or columns, but not both. Harrar and Gupta [7] give two more definitions for the matrix variate skew normal density:

$$f_2(X;\boldsymbol{\Sigma},\Psi,\boldsymbol{b},\Omega) = c_2^* \phi_{k\times n}(X;\mathbf{0},\boldsymbol{\Sigma},\Psi)\Phi_n(X'\boldsymbol{b};\mathbf{0},\Omega) \tag{7}$$

and

$$f_3(X;\boldsymbol{\Sigma},\Psi,\boldsymbol{b},B) = c_3^* \phi_{k\times n}(X;\mathbf{0},\boldsymbol{\Sigma},\Psi)\Phi(tr(B'X),0,1) \tag{8}$$

where $c_2^* = (\Phi_n(\mathbf{0},(\Omega+\boldsymbol{b}'\boldsymbol{\Sigma}\boldsymbol{b})\Psi))^{-1}$, $c_3^* = 2$; $\boldsymbol{\Sigma}$, Ψ, and Ω are positive definite covariance matrices of dimensions k, n and n respectively, B is a matrix of dimension $k\times n$. Note that if $\Omega = \Psi$ then f_2 is the same as f_1. Although, more general than f_1, the density f_2 still does not permit independence of rows and columns simultaneously.

A very general definition of skew symmetric variable for the matrix case can be obtained from matrix variate selection models. Suppose X is a $k\times n$ random matrix with density $f(X)$, let $g(X)$ be a weight function. A weighted form of density $f(X)$ is given by

$$h(X) = \frac{f(X)g(X)}{\int_{\mathbb{R}^{k\times n}} g(X)f(X)dX}. \tag{9}$$

When the sample is only a subset of the population then the associated model would be called a selection model.

In the next section, a construction for a family of matrix variate skew-symmetric densities that allows for independence among both variables and individuals is studied.

2.1. Matrix variate Skew Symmetric Distribution

To define a matrix variate distribution from the multivariate skew symmetric distribution first assume that $\boldsymbol{z}_i \sim ss_k^{g,H}(\mathbf{0}, I_k, \boldsymbol{\alpha}_i)$ for $i = 1, 2, \ldots, n$ are independently distributed random variables. Write $Z = (\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_n)$. We can write the density of X as a product as follows,

$$\prod_{i=1}^{n} 2^k g^*(Z) \prod_{j=1}^{k} H(\alpha_{ji} \boldsymbol{e}_j' \boldsymbol{z}_i).$$

This is equal to

$$2^{nk} g^{**}(Z) \prod_{i=1}^{n} \prod_{j=1}^{k} H(\alpha_{ji} \boldsymbol{e}_j' Z \boldsymbol{c}_i).$$

Let A, and B be nonsingular symmetric matrices of order k, and n respectively, also assume M is a $k \times n$ matrix. Define the matrix variate skew symmetric variable as $X = AZB + M$.

Definition 1. *(Matrix Variate Skew-Symmetric Density) Let $g(.)$ be a density function symmetric about 0, $H(.)$ be an absolutely continuous cumulative distribution function with $H'(.)$ symmetric about 0. A variable X has matrix variate skew symmetric distribution if it has probability density function*

$$\frac{2^{nk} g^{**}(A^{-1}(X - M)B^{-1}) \prod_{i=1}^{n} \prod_{j=1}^{k} H(\alpha_{ji} \boldsymbol{e}_j' (A^{-1}(X - M)B^{-1}) \boldsymbol{c}_i)}{|A|^n |B|^k} \tag{10}$$

where α_{ji} are real scalars, $M \in \mathbb{R}^{k \times n}$, A and B be nonsingular symmetric matrices of order k and n respectively. Finally, $g^{**}(X) = \prod_{i=1}^{n} \prod_{j=1}^{k} g(y_{ij})$. The density is called matrix variate skew-symmetric density with location parameter M, scale parameters (A, B), and shape parameter $\Delta = (\alpha_{ji})$, and it is denoted by $mss_{k \times n}^{g,H}(M, A, B, \Delta)$.

Let $Z \sim mss_{k \times n}^{g,H}(\mathbf{0}_{k \times n}, I_k, I_n, \Delta)$. The moment generating function of Z evaluated at $T_{k \times n} \in \mathbb{R}^{k \times n}$ is $M_Z(T_{k \times n})$ and can be obtained as follows:

$$\begin{aligned}
M_Z(T_{k \times n}) &= E(etr(T_{k \times n}' Z)) \\
&= \int_{\mathbb{R}^{k \times n}} etr(T_{k \times n}' Z) 2^k g^{**}(Z) \prod_{i=1}^{n} \prod_{j=1}^{k} H(\alpha_{ji} \boldsymbol{e}_j' Z \boldsymbol{c}_i) dZ \\
&= E_{g^{**}(Z)}(etr(T_{k \times n}' Z) \prod_{i=1}^{n} \prod_{j=1}^{k} H(\alpha_{ji} \boldsymbol{e}_j' Z \boldsymbol{c}_i)).
\end{aligned}$$

Let $X = AZB + M$ where A $(k \times k)$, B $(n \times n)$ and M $(k \times n)$ are constant matrices. Then moment generating function of X evaluated at $T_{k\times n}$ is $M_X(T_{k\times n})$:

$$\begin{aligned} M_X(T_{k\times n}) &= etr(T'_{k\times n}M)M_Z(A'T_{k\times n}B') \\ &= etr(T'_{k\times n}M)E_{g^{**}(Z)}(etr((BT'_{k\times n}A)Z)\prod_{i=1}^{n}\prod_{j=1}^{k}H(\alpha_{ji}\boldsymbol{e}'_j Z\boldsymbol{c}_i)). \end{aligned}$$

Definition 2. *(Matrix Variate Skew Symmetric Distributions) Let $g(.)$ be a density function symmetric about 0, $H(.)$ be an absolutely continuous cumulative distribution function with $H'(.)$ symmetric about 0. Let $z_{ij} \sim f(z_{ij}, \alpha_{ji}) = 2g(z_{ij})H(\alpha_{ji}z)$ for $i = 1, 2, \ldots, n$, and $j = 1, 2, \ldots, k$ be independent variables. Then the matrix variate random variable $Z = (z_{ij})$ has density*

$$2^{nk}g^{**}(Z)\prod_{i=1}^{n}\prod_{j=1}^{k}H(\alpha_{ji}\boldsymbol{e}'_j Z\boldsymbol{c}_i)$$

where $g^{**}(\boldsymbol{z}) = \prod_{i=1}^{n}\prod_{j=1}^{k}g(z_{ij})$, and $\boldsymbol{e}'_j$ and $\boldsymbol{c}'_i$ are the elementary vectors of the coordinate system $\mathbb{R}^k$ and $\mathbb{R}^n$ respectively. Let $X = AZB + M$ where A $(k \times k)$, B $(n \times n)$ and M $(k \times n)$ are constant matrices. Then the random variable $X = AZB + M$ has matrix variate skew symmetric distribution with location parameter M, scale parameters (A, B), and shape parameter $\Delta = (\alpha_{ji})$. We denote this by $X \sim MSS_{k\times n}^{g,H}(M, A, B, \Delta)$.

2.2. Matrix Variate Skew-Normal Distribution

Definition 3. *(Matrix Variate Skew Normal Density). We call*

$$\frac{2^{kn}\phi_{k\times n}(A^{-1}(X-M)B^{-1})\prod_{j=1}^{k}\prod_{i=1}^{n}\Phi(\alpha_{ji}\boldsymbol{e}'_j(A^{-1}(X-M)B^{-1})\boldsymbol{c}_i)}{|A|^n|B|^k} \tag{11}$$

the matrix variate skew normal density with location parameter M, scale parameters (A, B), and shape parameter Δ. We denote it by $msn_{k\times n}(M, A, B, \Delta)$.

We will need the following lemmas: See Zacks [9] and Chen and Gupta [4].

Lemma 1. *Let $\boldsymbol{z} \sim \phi_k(\boldsymbol{z})$. For scalar b, $\boldsymbol{a} \in \mathbb{R}^k$, and for $\boldsymbol{\Sigma}$ a positive definite matrix of order k $E(\Phi(b + \boldsymbol{a}'\boldsymbol{\Sigma}^{1/2}\boldsymbol{z})) = \Phi(\frac{b}{(1+\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a})^{1/2}})$.*

Lemma 2. *Let $Z \sim \phi_{k\times n}(Z)$. For scalar $\boldsymbol{b}, \boldsymbol{a} \in \mathbb{R}^k$, and for A and B positive definite matrices of order k and n respectively, $E(\Phi_n(\boldsymbol{b} + \boldsymbol{a}'AZB)) = \Phi_n(\boldsymbol{a}, (1 + \boldsymbol{a}'AA'\boldsymbol{a})^{1/2}B)$.*

Let $Z \sim msn_{k\times n}(\boldsymbol{0}_{k\times n}, I_k, I_n, \Delta)$. Then, the moment generating function of Z evaluated at $T_{k\times n}$ is $M_Z(T_{k\times n})$. It can be obtained as follows:

$$M_Z(T_{k\times n}) = E_{\phi_{k\times n}(Z)}(etr(T'_{k\times n}Z)\prod_{i=1}^{n}\prod_{j=1}^{k}\Phi(\alpha_{ji}\boldsymbol{e}'_j Z\boldsymbol{c}_i))$$

$$= \frac{2^k}{(2\pi)^{k/2}} \prod_{i=1}^{n} \int_{R^k} e^{-\frac{1}{2}\mathbf{z}_i'\mathbf{z}_i + \mathbf{t}_i{}'\mathbf{z}_i} \prod_{j=1}^{k} \Phi(\alpha_{ji}\mathbf{e}_j'\mathbf{z}_i)\mathbf{dz}$$

$$= \frac{2^k}{(2\pi)^{k/2}} \prod_{i=1}^{n} \int_{R^k} e^{-\frac{1}{2}(\mathbf{z}_i'\mathbf{z}_i - 2\mathbf{t}_i{}'\mathbf{z}_i)} \prod_{j=1}^{k} \Phi(\alpha_{ji}\mathbf{e}_j'\mathbf{z}_i)\mathbf{dz}$$

$$= \prod_{i=1}^{n} \frac{2^k e^{\frac{1}{2}\mathbf{t}_i{}'\mathbf{t}_i}}{(2\pi)^{k/2}} \int_{R^k} e^{-\frac{1}{2}(\mathbf{z}_i'\mathbf{z}_i - 2\mathbf{t}_i{}'\mathbf{z}_i + \mathbf{t}_i{}'\mathbf{t}_i)} \prod_{j=1}^{k} \Phi(\alpha_{ji}\mathbf{e}_j'\mathbf{z}_i)\mathbf{dz}$$

$$= \prod_{i=1}^{n} \frac{2^k e^{\frac{1}{2}\mathbf{t}_i{}'\mathbf{t}_i}}{(2\pi)^{k/2}} \int_{R^k} e^{-\frac{1}{2}(\mathbf{z}_i - \mathbf{t}_i)'(\mathbf{z}_i - \mathbf{t}_i)} \prod_{j=1}^{k} \Phi(\alpha_{ji}\mathbf{e}_j'\mathbf{z}_i)\mathbf{dz}$$

$$= \prod_{i=1}^{n} \frac{2^k e^{\frac{1}{2}\mathbf{t}_i{}'\mathbf{t}_i}}{(2\pi)^{k/2}} \prod_{j=1}^{k} \int_{R} e^{-\frac{1}{2}(z_{ij} - (\mathbf{t})_{ij})^2} \Phi(\alpha_{ji}\mathbf{z}_{ij}) dz_{ij}$$

$$= \prod_{i=1}^{n} 2^k e^{\frac{1}{2}\mathbf{t}_i{}'\mathbf{t}_i} \prod_{j=1}^{k} \int_{R} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}(y_{ij})^2} \Phi(\alpha_{ji}\mathbf{y}_{ij} + \alpha_{ji}(\mathbf{t})_{ij}) dy_{ij}$$

$$= \prod_{i=1}^{n} 2^k e^{\frac{1}{2}\mathbf{t}_i{}'\mathbf{t}_i} \prod_{j=1}^{k} \Phi\left(\frac{\alpha_{ji}(\mathbf{t})_{ij}}{\sqrt{(1+\alpha_{ji}{}^2)}}\right)$$

$$= 2^{nk} etr\left(\frac{1}{2} T_{k\times n}{}' T_{k\times n}\right) \prod_{i=1}^{n} \prod_{j=1}^{k} \Phi\left(\frac{\alpha_{ji}(T)_{ij}}{\sqrt{(1+\alpha_{ji}{}^2)}}\right).$$

Let $X = AZB + M$ for constant $(k \times k)$ matrix $A, (n \times n)$ matrix B and $k \times n$ dimensional constant matrix M. Then the moment generating function of X evaluated at $T_{k\times n} \in \mathbb{R}^{k\times n}$ is $M_X(T_{k\times n})$, this can be obtained as follows:

$$\begin{aligned} M_X(T_{k\times n}) &= 2^{nk} etr\left(T_{k\times n}' M + \frac{1}{2}(A' T_{k\times n} B')' A' T_{k\times n} B'\right) \\ &\times \prod_{i=1}^{n} \prod_{j=1}^{k} \Phi\left(\frac{\alpha_{ji}(A' T_{k\times n} B')_{ij}}{\sqrt{(1+(\alpha_{ji}^2)}}\right). \end{aligned}$$

Hence the following definition and theorems.

Definition 4. *(Matrix Variate Skew Normal Random Variable) Let*

$$z_{ij} \sim 2\phi(z_{ij})\Phi(\alpha_{ji} z_{ij})$$

for $i = 1, 2, \ldots, n$, and $j = 1, 2, \ldots, k$ be independent univariate skew normal random variables. Then the matrix variate random variable $Z = (z_{ij})$ has density

$$2^{nk}\phi_{k\times n}(Z) \prod_{i=1}^{n} \prod_{j=1}^{k} \Phi(\alpha_{ji}\mathbf{e}_j' Z \mathbf{c}_i)$$

where $\phi_{k\times n}(Z)=\prod_{i=1}^{n}\prod_{j=1}^{k}\phi(z_{ij})$, *and* $\boldsymbol{e}_j$ *and* $\boldsymbol{c}_i$ *are the elementary vectors of the coordinate system* $\mathbb{R}^k$ *and* $\mathbb{R}^n$ *respectively. Let* A *be a* $k\times k$ *constant matrix,* B *be a* $n\times n$ *constant matrix and* M *be a* $k\times n$*-dimensional constant matrix. A random variable* $X=AZB+M$ *is distributed with respect to matrix variate skew symmetric distribution with location parameter* M, *scale parameters* (A,B), *and shape parameter* $\Delta=(\alpha_{ji})$. *We denote this by* $X\sim MSN_{k\times n}(M,A,B,\Delta)$. *If the density exists it is given in Equation (11). We denote this case by writing* $X\sim msn_{k\times n}(M,A,B,\Delta)$.

Theorem 3. *If* X *has multivariate skew-normal distribution* $MSN_{k\times n}(M,A,B,\Delta)$ *then the moment generating function of* X *evaluated at* $T_{k\times n}$ *is given by*

$$\begin{aligned}M_X(T_{k\times n})&=2^{nk}etr(T'_{k\times n}M+\frac{1}{2}(A'T_{k\times n}B')'A'T_{k\times n}B')\\&\times\prod_{i=1}^{n}\prod_{j=1}^{k}\Phi(\frac{\alpha_{ji}(A'T_{k\times n}B')_{ij}}{\sqrt{(1+\alpha_{ji}{}^2)}}).\end{aligned}\tag{12}$$

By Definition 5 we can write $Z\sim MSN_{k\times n}(\mathbf{0},I_k,I_n,\Delta)$, and prove the following theorems.

Theorem 4. *Assume that* $Y\sim MSN_{k\times n}(M,A,B,\Delta)$ *and* $X=CYD+N$ *where* C, D *and* N *are matrices of order* $k'\times k$, $n\times n'$ *and* $k'\times n'$ *respectively. Then* $X\sim MSN_{k'\times n'}(CMD+N,CA,BD,\Delta)$.

Proof. From assumption, we have $Y=AZB+M$, and so $X=CAZBD+(CMD+N)$, i.e., $X\sim MSN_{k'\times n'}(CMD+N,CA,BD,\Delta)$.

Theorem 5. *Let* $\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots\boldsymbol{x}_n$ *be independent, where* $\boldsymbol{x}_i$ *is distributed according to* $sn_k(\mathbf{0},\boldsymbol{\Sigma}^{1/2},\boldsymbol{\alpha})$. *Then,*

$$\sum_{j=1}^{n}\boldsymbol{x}'_j\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_j\sim\chi^2_{kn}.$$

Proof. Let $\boldsymbol{y}\sim N_k(\boldsymbol{\mu}=0,\boldsymbol{\Sigma})$. Then $\mathbf{y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y}\sim\chi^2_k$, and $\boldsymbol{x}'_j\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_j$ and $\mathbf{y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y}$ have the same distribution from Theorem 2. Moreover, $\boldsymbol{x}'_j\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_j$ are independent. Then the desired property is proven by the addition property of χ^2 distribution.

It is well known that if $X\sim\phi_{k\times n}(M,AA',\Psi=I_n)$ then the matrix variate XX' has the Wishart distribution with the moment generating function given as $|(I-2(AA')T)|^{-n/2}$, $(AA')^{-1}-2T$ being a positive definite matrix. The following theorem implies that the decomposition for a Wishart matrix is not unique.

Theorem 6. *If a* $k\times n$ *matrix variate random variable* X *has* $msn_{k\times n}(\mathbf{0}_{k\times n},A,I_n,\Delta)$ *distribution for constant positive definite matrix* A *of order* k *then* $XX'\sim W_k(n,AA')$.

Proof. The moment generating function of the quadratic form XX' can be obtained as follows, for any $T\in\mathbb{R}^{k\times k}$, with $(AA')^{-1}-2T$ being a positive definite matrix,

$$E(etr(XX'T))=\int_{\mathbb{R}^{k\times n}}etr(XX'T)dFX$$

$$= \int_{\mathbb{R}^{k\times n}} \frac{2^{nk} etr(-\frac{1}{2}(AA')^{-1}XX' + XX'T) \prod_{i=1}^{n} \prod_{j=1}^{k} \Phi(\alpha_{ji}\boldsymbol{e}_j' A^{-1} X \mathbf{c}_i)}{(2\pi)^{nk/2}|A|^n} dX$$

$$= \int_{\mathbb{R}^{k\times n}} \frac{2^{nk} etr(-\frac{1}{2}X'((AA')^{-1} - 2T)X) \prod_{i=1}^{n} \prod_{j=1}^{k} \Phi(\alpha_{ji}\boldsymbol{e}_j' A^{-1} X \mathbf{c}_i)}{(2\pi)^{nk/2}|A|^n} dX$$

$$= \int_{\mathbb{R}^{k\times n}} \frac{2^{nk} etr(-\frac{1}{2}Z'Z) \prod_{i=1}^{n} \prod_{j=1}^{k} \Phi(\alpha_{ji}\boldsymbol{e}_j' A^{-1}((AA')^{-1} - 2T)^{1/2} Z \mathbf{c}_i)}{(2\pi)^{nk/2}|(I - 2(AA')T)|^{n/2}} dZ$$

$$= \frac{2^{nk} \prod_{i=1}^{n} \prod_{j=1}^{k} E_z(\Phi(cz))}{(2\pi)^{nk/2}|(I - 2(AA')T)|^{n/2}}$$

($z \sim \phi(z)$ and $c \in \mathbf{R}$ is a constant)

$$= \frac{2^{nk}(\frac{1}{2})^{nk}}{|(I - 2(AA')T)|^{n/2}}$$

$$= |(I - 2(AA')T)|^{-n/2}.$$

3. Generalized Matrix Variate Skew Normal Distribution

Definition 5. *(Generalized Matrix Variate Skew Normal Density). Let $X = AZB + M$ where A $(k \times k)$, B $(n \times n)$ and M $(k \times n)$ are constant matrices and let A, B be nonsingular. We call the density*

$$\frac{2^{kn}\phi_{k\times n}(A^{-1}(X - M)B^{-1}) \prod_{j=1}^{k} H(\alpha_{ji}\boldsymbol{e}_j'(A^{-1}(X - M)B^{-1})\mathbf{c}_i)}{|A|^n|B|^k} \tag{13}$$

the matrix variate skew normal density and denote this by $gmsn^H_{k\times n}(M, A, B, \Delta)$.

Let $Z \sim gmsn^H_{k\times n}(\mathbf{0}_{k\times n}, I_k, I_n, \Delta)$. The moment generating function of Z evaluated at $T_{k\times n}$ is $M_Z(T_{k\times n})$, it can be obtained as follows:

$$M_Z(T_{k\times n}) = E_{H_{k\times n}(Z)}(etr(T'_{k\times n}Z) \prod_{i=1}^{n} \prod_{j=1}^{k} H(\alpha_{ji}\boldsymbol{e}_j' Z \mathbf{c}_i))$$

$$= \frac{2^k}{(2\pi)^{k/2}} \prod_{i=1}^{n} \int_{R^k} e^{-\frac{1}{2}\boldsymbol{z}_i'\boldsymbol{z}_i + \boldsymbol{t}_i{}'\boldsymbol{z}_i} \prod_{j=1}^{k} H(\alpha_{ji}\boldsymbol{e}_j'\boldsymbol{z}_i)\mathbf{dz}$$

$$= \frac{2^k}{(2\pi)^{k/2}} \prod_{i=1}^{n} \int_{R^k} e^{-\frac{1}{2}(\boldsymbol{z}_i'\boldsymbol{z}_i - 2\boldsymbol{t}_i{}'\boldsymbol{z}_i)} \prod_{j=1}^{k} H(\alpha_{ji}\boldsymbol{e}_j'\boldsymbol{z}_i)\mathbf{dz}$$

$$
\begin{aligned}
&= \prod_{i=1}^{n} \frac{2^k e^{\frac{1}{2}\boldsymbol{t}_i{}'\boldsymbol{t}_i}}{(2\pi)^{k/2}} \int_{R^k} e^{-\frac{1}{2}(\boldsymbol{z}_i'\boldsymbol{z}_i - 2\boldsymbol{t}_i{}'\boldsymbol{z}_i + \boldsymbol{t}_i{}'\boldsymbol{t}_i)} \prod_{j=1}^{k} H(\alpha_{ji}\boldsymbol{e}_j'\boldsymbol{z}_i)\mathbf{dz} \\
&= \prod_{i=1}^{n} \frac{2^k e^{\frac{1}{2}\boldsymbol{t}_i{}'\boldsymbol{t}_i}}{(2\pi)^{k/2}} \int_{R^k} e^{-\frac{1}{2}(\boldsymbol{z}_i - \boldsymbol{t}_i)'(\boldsymbol{z}_i - \boldsymbol{t}_i)} \prod_{j=1}^{k} H(\alpha_{ji}\boldsymbol{e}_j'\boldsymbol{z}_i)\mathbf{dz} \\
&= \prod_{i=1}^{n} \frac{2^k e^{\frac{1}{2}\boldsymbol{t}_i{}'\boldsymbol{t}_i}}{(2\pi)^{k/2}} \prod_{j=1}^{k} \int_{R} e^{-\frac{1}{2}(z_{ij} - (\boldsymbol{t})_{ij})^2} H(\alpha_{ji}\boldsymbol{z}_{ij}) dz_{ij} \\
&= \prod_{i=1}^{n} 2^k e^{\frac{1}{2}\boldsymbol{t}_i{}'\boldsymbol{t}_i} \prod_{j=1}^{k} \int_{R} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}(y_{ij})^2} H(\alpha_{ji}\boldsymbol{y}_{ij} + \alpha_{ji}(\boldsymbol{t})_{ij}) dy_{ij} \\
&= \prod_{i=1}^{n} 2^k e^{\frac{1}{2}\boldsymbol{t}_i{}'\boldsymbol{t}_i} \prod_{j=1}^{k} H(\frac{\alpha_{ji}(\boldsymbol{t})_{ij}}{\sqrt{(1+\alpha_{ji}{}^2)}}) \\
&= 2^{nk} etr(\frac{1}{2} T_{k\times n}{}' T_{k\times n}) \prod_{i=1}^{n}\prod_{j=1}^{k} H(\frac{\alpha_{ji}(T)_{ij}}{\sqrt{(1+\alpha_{ji}{}^2)}}).
\end{aligned}
$$

Let $X = AZB + M$ for constant $(k \times k)$ matrix A, $(n \times n)$ matrix B and $k \times n$ dimensional constant matrix M. Then the moment generating function of X evaluated at $T_{k\times n} \in \mathbb{R}^{k\times n}$ is $M_X(T_{k\times n})$, this can be obtained as follows:

$$
\begin{aligned}
M_X(T_{k\times n}) = 2^{nk} etr(T_{k\times n}'M + \frac{1}{2}(A'T_{k\times n}B')'A'T_{k\times n}B') \\
\times \prod_{i=1}^{n}\prod_{j=1}^{k} H(\frac{\alpha_{ji}(A'T_{k\times n}B')_{ij}}{\sqrt{(1+\alpha_{ji}{}^2)}}).
\end{aligned}
$$

Hence the following definition and theorems.

Definition 6. *(Matrix Variate Generalized Skew Normal distribution) Let*

$$z_{ij} \sim 2\phi(z_{ij})H(\alpha_{ji}z_{ij})$$

for $i = 1, 2, \ldots, n$, and $j = 1, 2, \ldots, k$ be independent univariate generalized skew normal random variables. The matrix variate random variable $Z = (z_{ij})$ has density

$$2^{nk}\phi_{k\times n}(Z)\prod_{i=1}^{n}\prod_{j=1}^{k} H(\alpha_{ji}\boldsymbol{e}_j'Z\mathbf{c}_i)$$

where $\phi_{k\times n}(Z) = \prod_{i=1}^{n}\prod_{j=1}^{k}\phi(z_{ij})$, and $\boldsymbol{e}_j'$ and $\mathbf{c}_i'$ are the elementary vectors of the coordinate system $\mathbb{R}^k$ and $\mathbb{R}^n$ respectively. Let A be a $k \times k$ constant matrix, B be a $n \times n$ constant matrix and M be a $k \times n$-dimensional constant matrix. Then the random matrix $X = AZB + M$ is distributed with respect to generalized matrix variate skew symmetric distribution with location *parameter M, scale parameters (A,B), and shape parameter $\Delta = (\alpha_{ji})$. We denote this by $X \sim GMSN^H_{k\times n}(M,A,B,\Delta)$. If the density exists it is given in Equation (13). We denote this case by writing $X \sim gsn^H_{k\times n}(M,A,B,\Delta)$.*

Theorem 7. *If X has generalized matrix variate skew-normal distribution, $GMSN^H_{k\times n}(M,A,B,\Delta)$, then the moment generating function of X evaluated at $T_{k\times n}$ is given by*

$$M_X(T_{k\times n}) = 2^{nk} etr(T'_{k\times n}M + \frac{1}{2}(A'T_{k\times n}B')'A'T_{k\times n}B') \times \prod_{i=1}^{n}\prod_{j=1}^{k} H(\frac{\alpha_{ji}(A'T_{k\times n}B')_{ij}}{\sqrt{(1+\alpha_{ji}{}^2)}}). \tag{14}$$

By Definition 6 we can write $Z \sim GMSN^H_{k\times n}(\mathbf{0}, I_k, I_n, \Delta)$, and prove the following theorems.

Theorem 8. *Assume that $Y \sim GMSN^H_{k\times n}(M,A,B,\Delta)$ and $X = CYD + N$ where C, D and N are matrices of order $k' \times k$, $n \times n'$ and $k' \times n'$ respectively. Then $X \sim GMSN^H_{k'\times n'}(CMD + N, CA, BD, \Delta)$.*

Proof. From assumption we have $Y = AZB + M$, and so $X = CAZBD + (CMD + N)$, i.e., $X \sim MSN_{k'\times n'}(CMD + N, CA, BD, \Delta)$.

Theorem 9. *Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_n$ be independent, where $\boldsymbol{x}_i$ is distributed according to $gsn^H_k(\mathbf{0}, \boldsymbol{\Sigma}^{1/2}, \boldsymbol{\alpha})$. Then,*

$$\sum_{j=1}^{n} \boldsymbol{x}'_j \boldsymbol{\Sigma}^{-1} \boldsymbol{x}_j \sim \chi^2_{kn}$$

.

Proof. Let $\boldsymbol{y} \sim N_k(\boldsymbol{\mu} = 0, \boldsymbol{\Sigma})$. Then $\mathbf{y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} \sim \chi^2_k$, and $\boldsymbol{x}'_j\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_j$ and $\mathbf{y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y}$ have the same distribution from Theorem 2. Moreover, $\boldsymbol{x}'_j\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_j$ are independent. Then the desired property is proved by the addition property of χ^2 distribution.

4. Extensions

An odd function, say $w(x)$, can be used to replace the term in the form $\alpha_{ji}x$ in the skewing function to give more flexible families of densities. We can take $w(x_{ji}) = \frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}$, to obtain a matrix variate form of the skew symmetric family introduced by Arellanno-Valle at al. [1]; if we take $w(x) = \alpha x + \beta x^3$, we obtain a matrix variate form of the skew symmetric family introduced by Ma and Genton [8]; or take $w(x) = sign(x)|x|^{\alpha/2}\lambda(2/\alpha)^{1/2}$ to obtain a matrix variate form of the skew symmetric family introduced by DiCiccio and Monti [5].

The skewing function of the matrix variate skew normal density in Equation (11), i.e.

$$\prod_{j=1}^{k}\prod_{i=1}^{n} \Phi(\alpha_{ji}\boldsymbol{e}'_j(A^{-1}(X-M)B^{-1})\boldsymbol{c}_i), \tag{15}$$

can be replaced by

$$\Phi_{k^*\times n^*}(\Gamma A^{-1}(X-M)B^{-1}\Lambda;\mathbf{0},C,D) \tag{16}$$

where Γ, Λ, C (positive definite) and D (positive definite) are matrices of dimensions $k^*\times k$, $n\times n^*$, $k^*\times k^*$ and $n^*\times n^*$ correspondingly. In this case, the normalizing constant 2^{kn} in (11) will have to be changed to $E_Z(P(Y<\Gamma A^{-1}(Z-M)B^{-1}\Lambda|Z))$ for $Y\sim\phi_{k^*\times n^*}(\mathbf{0},C,D)$ and $Z\sim\phi_{k\times n}(M,AA',B'B)$. If all of Γ, Λ, C and D are positive definite diagonal matrices then (16) can be written in the same form as (15), therefore the latter density is more general. Armando [2] introduced the matrix variate closed skew-normal distribution based on marginal representation or hidden truncation. The density of the matrix variate closed skew-normal distribution has the expression (16) as for its skewing function. The class of distributions defined by $msn_{k\times n}(M,A,B,\Delta)$ is a subclass of the matrix variate closed skew-normal distribution. The extension $gmsn^H_{k\times n}(M,A,B,\Delta)$ is not a subclass of the matrix variate closed skew-normal distribution.

References

[1] R.B. Arellano-Valle, H.W. Gomez, and F.A. Quintana. A new class of skew-normal distributions. *Communications in Statistics-Theory and Methods*, 33(7):1465–1480, 2004.

[2] J. Armando, G.F. Domínguez-Molina, R. Ramos-Quiroga, and A.K. Gupta. A matrix variate closed skew-normal distribution with applications to stochastic frontier analysis. *Communications in Statistics - Theory and Methods*, 36(9):1691–1703, 2007.

[3] L. Bailey, T.F. Moore, and B.A. Bailar. An interviewer variance study for the eight impact cities of the national crime survey cities sample. *Journal of the American Statistical Association*, 73(1):16–23, 1978.

[4] J.T. Chen and A.K. Gupta. Matrix variate skew normal distributions. *Statistics*, 39(3):247–253, 2005.

[5] T.J. DiCiccio and A.C. Monti. Inferential aspects of the skew exponential power distribution. *Journal of American Statistical Association*, 99(466):439–450, 2004.

[6] A.K. Gupta and D.K. Nagar. *Matrix Variate Distributions*. Chapman & Hall/CRC, 2000.

[7] S.W. Harrar and A.K. Gupta. On matrix variate skew-normal distributions. *Statistics*, 42(2):179–194, 2008.

[8] Y.Y. Ma and M.G. Genton. Flexible class of skew-symmetric distributions. *Scandinavian Journal of Statistics*, 31(3):459–468, 2004.

[9] S. Zacks. *Parametric Statistical Inference*. Pergamon Press, New York, 1981.