Statistica & Applicazioni - Vol. XVIII, n. 2, 2020, pp. 161-175
First published online 16.10.2020

GENERALIZED ALPHA SKEW NORMAL DISTRIBUTION

Tareq Alodat*
Shahjahan Khan**
Moh'd T. Alodat***
Muhammed A Memon****

SUMMARY

This paper extends the work of Elal-Olivero (2010) on the alpha-skew normal distribution. The extension is a multivariate version of Elal-Olivero's univariate case. Then we study the statistical properties of the new extension such as marginal and conditional distribution, closure under convolution with normal random variate. Furthermore, we illustrate the performance of the distribution using simulated data obtained from the generalized distribution via the Metropolis-Hasting algorithm.

Keywords: Skew-normal Distribution, Skewness, Characteristics Function, Quadratic Form.
DOI: 10.26350/999999_000028
ISSN: 18246672 (print) 2283-6659 (digital)

1. INTRODUCTION

The commonly used normal distribution is considered to be a strong assumption for various data sets that arise in different fields of applications as environment, finance, medicine, etc. For example, the distribution of stock returns is skewed and has heavier tails than the normal distribution. So, one of the challenges in data modeling is to find a flexible class of distributions that capture both skewness and kurtosis of real data. The quest on developing such classes of distributions was started by Azzalini (1985) who proposed a class of univariate distributions possessing a skewness parameter. Azzalini and Dalla Valle (1996) extended the earlier work of Azzalini (1985) to the following multivariate skew normal distribution:

$$f_X(x) = 2\varphi_n(x;\mu,\Sigma)\Phi\left(\alpha^{\mathrm{T}}(x-\mu)\right), \qquad x \in R^n \tag{1}$$

* Department Mathematics and Statistics - La Trobe University - Victoria, 3086 - AUSTRALIA (e-mail: t.alodat@latrobe.edu.au).
** School of Sciences - University of Southern Queensland - AUSTRALIA (e-mail: Shahjahan.Khan@usq.edu.au).
*** Department of Statistics-Sultan Qaboos University - MUSCAT (e-mail: ✉ m.alodat@squ.edu.om).
**** School of Medicine - University of Queensland - AUSTRALIA (e-mail: mmemon@yahoo.com).

where $\alpha \in R^n$ is the skewness parameter, $\varphi_n(x; 0, \Sigma)$ is the n-dimensional probability density function (PDF) of multivariate normal distribution with zero mean vector, $\Phi(.)$ is the cumulative distribution function (CDF) of the standard normal distribution, and a positive definite covariance matrix Σ. The distribution in (1) is denoted by $X \sim SN_n\ (\mu, \Sigma, \alpha)$.

Over the last two decades, there was a considerable interest in defining new multivariate families of distributions that possess skewness parameters. A rich survey on skew distributions and its ramifications can be found in the book by Genton (2004). It has been shown that most multivariate skew normal distributions enjoy several nice statistical properties that coincide with/close to the properties of the multivariate normal distribution (Pewsey, 2000; Genton, 2004; Azzalini and Capitanio, 2014).

Although modeling data skewness via multivariate skew normal distributions has found considerable attentions in recent years, still this family of distributions suffering from computational challenges. These challenges are due to the presence of the CDF of the multivariate normal distribution in their definition. In particular, these challenges become obstacle when making inferences from higher n-dimensional data. As an alternative to the family of skew distributions, Elal-Olivero (2010) introduced a univariate family of distributions called alpha-skew-normal (ASN). Its PDF (f_X) is

$$f_X(x; \alpha) = \frac{1 + (1 - \alpha x)^2}{2 + \alpha^2} \varphi(x; 0, 1), \quad x \in R \tag{2}$$

where $\alpha \in R$. The random variable X in the above PDF is an alpha-skew-normal random variable, and α is the underlying parameter. We denote this as $X \sim ASN\ (\alpha)$.

The ASN family possesses a skewness parameter and does not include the univariate normal CDF in its definition. Furthermore, the skewness parameter has an effect on distribution's modality. Therefore, the ASN family of distributions is flexible enough to support easy computation as well as both unimodal and bimodal shapes. Although bimodal data can be approached using mixture distributions, there is still some considerable attention in literature of the bimodal distributions. For example, Sturrok (2008) has analyzed the solar neutrino capture-rate measurements that arise from the GALLEX experiment, which shows significant bimodality. Based on a probability-transform procedure, he introduced an index called "bimodality index" in order to assess the significance of bimodality in two experiments. Bayramoglu (2020) has studied several bivariate and multivariate distributions with bimodal marginals. Also, she motivated her study by data from hydrology, biology, medicine, astronomy, etc. Such studies give us a motivation to define and study a multivariate version of ASN family.

As a simple generalization to (2), Handam (2012) defined the following univariate distribution

$$g_X(x; \alpha) = \frac{1 + (1 - \alpha x)^{2n}}{A(\alpha)} \varphi(x; 0, 1), \qquad x \in R \tag{3}$$

where n is a positive integer, $\alpha \in R$ and $A(\alpha)$ is a normalizing constant. The bivariate of version of (2) has been defined by Louzada, Ara and Fernandes (2017). They have studied its distributional properties and the asymptotic properties of the maximum likelihood estimators. Although the ASN family has not given much attention in the literature, in particular in the multivariate case, the aforementioned properties make it a promising candidate for developing further statistical inferences. In this paper, we extend the work of Elal-Olivero (2010) to multivariate case and we study the properties of the new family. This extension to the multivariate case is considered as another generalization of the univariate skew normal distribution, which enjoys properties similar to those of the univariate skew normal one.

The rest of this paper is organized as follows. In Section 2, we define a multivariate family called the generalized alpha skew normal family. Then we define the location-scale of the new family. In Section 3, we study its statistical properties such as the moment generating and characteristic functions. In Section 4, we derive their marginal and conditional distributions, distribution of linear combination as well as the convolution with multivariate normal distribution. In Section 5, we use a simulated data to illustrate the performance of the generalized alpha skew normal distribution. In Section 6, we state our conclusions.

2. GENERALIZED ALPHA SKEW NORMAL DISTRIBUTION

In this section, we generalize (1) to the multivariate case. Also, location-scale version for the new generalization is defined.

DEFINITION 1

A random vector X is said to have an n-dimesnaional generalized alpha skew-normal distribution (GASN) if it has the following PDF:

$$f_X(x;\gamma,\theta,\alpha) = c\Big(\gamma + \big(\theta - \alpha^{\mathrm{T}}x\big)^2\Big)\varphi_n(x;0,\mathrm{I}_n), \quad x \in R^n \tag{4}$$

where $\theta \in R$ $\gamma > 0$, $\alpha \in R^n$ and $c^{-1} = \gamma + \theta^2 + \alpha^T\alpha$ is the normalizing constant, and $\varphi_n(x;0,I_n)$ stands for the n-dimensional PDF of the multivariate standard normal distribution.

Hereafter, we adopt the notation $X \sim GASN_n\ (\gamma,\ \theta,\alpha)$ for the distribution given by (4). A more general form of (4) is obtained by introducing a location and scale parameters μ and, where the matrix Σ is positive definite. If $\Sigma^{\frac{1}{2}}$ denotes the square root matrix of Σ, then the PDF of $Y = \mu + \Sigma^{\frac{1}{2}}X$ takes the form

$$f_Y(y;\gamma,\theta,\alpha,\mu,\Sigma) = \frac{\gamma + \left(\theta - \alpha^{\mathrm{T}}\Sigma^{-\frac{1}{2}}(y-\mu)\right)^2}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\varphi_n(y;\mu,\Sigma), \quad y \in R^n. \tag{5}$$

Here after, we adopt the notation $Y \sim GASN_n\ (\gamma,\ \theta,\alpha,\mu,\Sigma)$ to mean that Y follows an n-dimensional location-scale alpha skew-normal distribution with parameters $\gamma,\ \theta, \alpha, \mu$ and Σ. By taking the limits either as $\gamma \to 0$ or as $\gamma \to \infty$, one can show that the closure of the family defined by (5) contains the family of normal distributions. Similar conclusion can be obtained as either as $\theta \to 0$ or as $\theta \to \infty$.

The Figures 1 and 2 represent the PDF's and their contours plots for $GASN_2(\gamma,\ \theta, \alpha, 0, \Sigma_0)$ with $\gamma = 1,\ \ \theta = 1,\ \Sigma_0 = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$ and different values of the vector α.

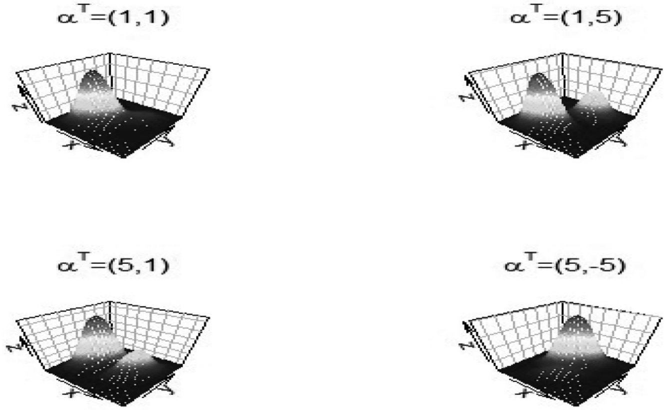

FIGURE 1. - *Plots for the PDF of* $\text{GASN}_2(\gamma, \theta, \alpha, 0, {}_0)$ *with* $\gamma = 1, \theta = 1$ *and different values of the vector* α

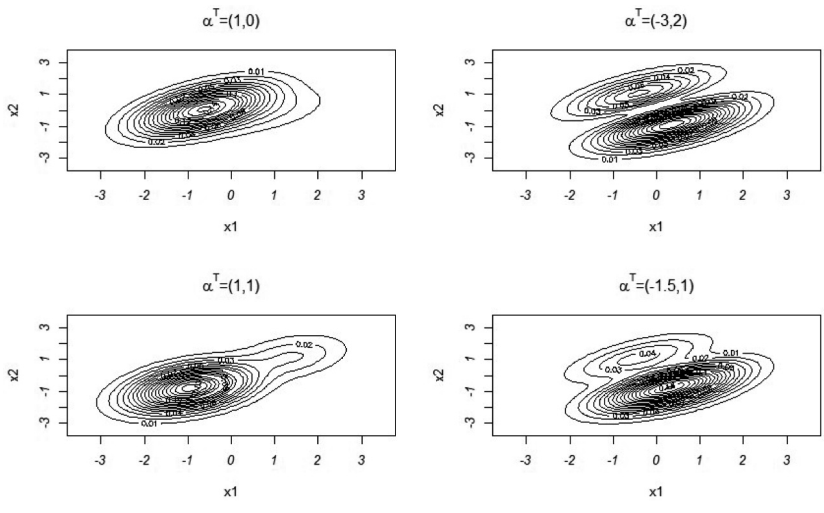

FIGURE 2. - *A 15 contour plots for the PDF of* $GASN_2(\gamma,\ \theta, \alpha, 0, \Sigma_0)$ *with* $\gamma = 1,\ \ \theta = 1$ *and different values of the vector* α

It is clear, from the above figures, that the parameter α affects the skewness, kurtosis and number of modes of the $GASN_2(\gamma, \ \theta, \alpha)$ distribution.

3. STATISTICAL PROPERTIES

This section deals with some properties of the new generalization.

THEOREM 1

If $X \sim GASN_n(\gamma, \ \theta, \alpha)$, then

(i) The moment generating function (MGF) of X is

$$M_X(t) = \left\{ 1 - \alpha^T t \frac{(2\theta - \alpha^T t)}{\gamma + \theta^2 + \alpha^T \alpha} \right\} e^{\frac{1}{2} t^T t}$$

(ii) The characteristic function (c.f) of X is

$$\varphi_X(t) = 1 - \alpha^T it \frac{(2\theta - \alpha^T it)}{\gamma + \theta^2 + \alpha^T \alpha} e^{-\frac{1}{2} t^T t}$$

(iii) The mean vector and covariance matrix of X, respectively are

$$E(X) = -\frac{2\theta\alpha}{\gamma + \theta^2 + \alpha^T \alpha}$$

and

$$Cov(X) = \frac{(\gamma + \theta^2 + \alpha^T \alpha) I_n + 2\alpha\alpha^T}{\gamma + \theta^2 + \alpha^T \alpha} - \frac{4\theta^2 \alpha\alpha^T}{\left(\gamma + \theta^2 + \alpha^T \alpha\right)^2}$$

PROOF:

(i) By the definition of the MGF, we have

$$\begin{aligned} M_X(t) = E\big(\exp(\mathrm{t^T X})\big) &= \int_{R^n} \exp(t^{\mathrm{T}} x) \frac{\gamma + (\theta - \alpha^T x)^2}{\gamma + \theta^2 + \alpha^{\mathrm{T}} \alpha} \varphi_n(x; 0, I_n) dx, \\ &= C \int_{R^n} \left\{ \gamma + \left(\theta - \alpha^{\mathrm{T}} x\right)^2 \right\} e^{-\frac{1}{2}(x^{\mathrm{T}} x - 2t^{\mathrm{T}} x)} \ dx, \end{aligned}$$

where C is given by

$$C^{-1} = \left(\gamma + \theta^2 + \alpha^{\mathrm{T}} \alpha\right) (2\pi)^{\frac{n}{2}}.$$

Let $\theta^* = \theta - \alpha^{\mathrm{T}} t$. Then we write

$$M_X(t) = Ce^{\frac{t^{\mathrm{T}} t}{2}} \int_{R^n} \left\{ \gamma + \left(\theta^* - \alpha^{\mathrm{T}} (x - t)\right)^2 \right\} e^{-\frac{1}{2}\left((x-t)^{\mathrm{T}} (x-t)\right)} \ dx.$$

More simplification gives

$$M_X(t) = \left(\gamma + \theta^{*2} + \alpha^{\mathrm{T}}\alpha\right)(2\pi)^{\frac{n}{2}}\ Ce^{\frac{t^{\mathrm{T}}t}{2}} \int_{R^n} \frac{\left\{\gamma + (\theta^* - \alpha^{\mathrm{T}}(x-t))^2\right\}}{\left(\gamma + \theta^{*2} + \alpha^{\mathrm{T}}\alpha\right)(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\left((x-t)^{\mathrm{T}}(x-t)\right)}\ dx.$$

So we get

$$M_X(t) = \left(\gamma + \theta^{*2} + \alpha^{\mathrm{T}}\alpha\right)(2\pi)^{\frac{n}{2}}\ Ce^{\frac{1}{2}t^{\mathrm{T}}t}.$$

Using the expression of C then

$$M_X(t) = \frac{\gamma + \theta^{*2} + \alpha^{\mathrm{T}}\alpha}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha} e^{\frac{t^{\mathrm{T}}t}{2}} = \left(1 - \alpha^{\mathrm{T}}t\ \frac{(2\theta - \alpha^{\mathrm{T}}t)}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\right) e^{\frac{1}{2}t^{\mathrm{T}}t}.$$

(ii) It is analogous to the proof of (i)

(iii) The mean and covariance matrix can obtain by differentiating the moment generating function two times directly and setting t by 0.

REMARK 1

Notice that it is easy to generalize (i)-(iii) in Theorem 1 immediately to the location-scale version. So, these results can be summarized as follows:

$$(i)' \qquad M_Y(t) = \left\{1 - \alpha^{\mathrm{T}}\Sigma^{\frac{1}{2}}t\frac{\left(2\theta\ - \alpha^{\mathrm{T}}\Sigma^{\frac{1}{2}}t\right)}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\right\} \mathrm{e}^{t^{\mathrm{T}}\mu + \frac{1}{2}t^{\mathrm{T}}\Sigma t}$$

$$(ii)' \quad \varphi_Y(t) = 1 - \alpha^{\mathrm{T}}\mathrm{i}\Sigma^{\frac{1}{2}}t\frac{\left(2\theta - \alpha^{\mathrm{T}}\mathrm{i}^{\frac{1}{2}}t\right)}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\mathrm{e}^{t^{\mathrm{T}}\mu - \frac{1}{2}t^{\mathrm{T}}\Sigma t}$$

$$(iii)' \quad E(Y) = \mu - \frac{2\theta\alpha}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha} \ \ \textit{and}$$

$$Cov(Y) = \Sigma + \frac{2\Sigma^{\frac{1}{2}}\alpha\alpha^{\mathrm{T}}\Sigma^{\frac{1}{2}}}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha} - \left(\mu - \frac{2\theta\alpha}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\right)\left(\mu - \frac{2\theta\alpha}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\right)^{\mathrm{T}}$$

4. MARGINAL AND CONDITIONAL DISTRIBUTIONS

In this section, we show that the generalized multivariate alpha skew normal distribution is closed under marginal and conditional distributions. Before starting these properties, we need the following notation. Let Σ be a positive definite matrix and $\Sigma^{\frac{1}{2}}$ is the square root of Σ. Consider the following partitions for Σ and $\Sigma^{\frac{1}{2}}$:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \text{ and } \Sigma^{\frac{1}{2}} = A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

where Σ_{11} and A_{11} are matrices of size $m \times m$, Σ_{22} and A_{22} are matrices of size $(n-m) \times (n-m)$, Σ_{12} and A_{12} are vectors of size $m \times (n-m)$, $\Sigma_{21} = \Sigma_{12}^T$ and $A_{21} = A_{12}^T$.

Then the inverses of Σ and $\Sigma^{\frac{1}{2}}$ can be written as

$$\Sigma^{-1} = B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \text{ and } \Sigma^{-\frac{1}{2}} = C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix},$$

where

$$B_{11} = \left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1}, \quad B_{12} = -\Sigma_{11}^{-1}\Sigma_{12}\left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)^{-1},$$

$$B_{21} = -\Sigma_{22}^{-1}\Sigma_{21}\left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1} \text{ and } B_{22} = \left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)^{-1}.$$

With the above notations, we give the following theorem.

THEOREM 2

Let $X \sim GASN_n\ (\gamma,\ \theta, \alpha, \mu, \Sigma)$. *If we consider the partitions of* $X,\ \mu$ *and* α*:*

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix},\ \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \text{ and } \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix},$$

where $X_1,\ \mu_1$ *and* α_1 *are vectors of size* $m \times 1$ *and* $X_2,\ \mu_2$ *and* α_2 *are vectors of size* $(n-m) \times 1$*, then*

(i) The marginal distribution of X_1 *is*

$$X_1 \sim GASN_m\ (\gamma^*,\ \theta, \alpha^*, \mu_1, \Sigma_{11}),$$

where

$\alpha^* = \Sigma_{11}^{\frac{1}{2}}\lambda_1, \quad \gamma^* = \gamma + \lambda_2^T B_{22}^{-1}\lambda_2,\ \lambda_1 = C_{11}\alpha_1 + C_{12}\alpha_2 + \Sigma_{11}^{-1}\Sigma_{12}\lambda_2,$
and $\lambda_2 = C_{21}\alpha_1 + +C_{22}\alpha_2$.

(ii) The conditional distribution of X_2 *given* $X_1 = x_1$ *is*

$$(X_2 | X_1 = x_1) \sim GASN_{n-m}\left(\gamma,\ \tilde{\theta}, B_{22}^{-\frac{1}{2}}\lambda_2,\ \mu_{1.2}, B_{11}^{-1}\right)$$

where $\tilde{\theta} = \theta - \lambda_1^T(x_1 - \mu_1)$ *and* $\mu_{1.2} = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1)$.

PROOF:

(i) Let $t^{\mathrm{T}} = (t_1^{\mathrm{T}}\ \ t_2^{\mathrm{T}})$ where t_1^{T} and t_2^{T} are vectors of dimensions $1 \times m$ and $1 \times (n-m)$, respectively. Using the MGF and set $t_2 = 0$, one obtains

$$M_{X_1}(t_1) = \left\{ 1 - (A_{11}\alpha_1 + A_{12}\alpha_2)^{\mathrm{T}} t_1 \frac{\left(2\theta - (A_{11}\alpha_1 + A_{12}\alpha_2)^{\mathrm{T}} t_1\right)}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha} \right\} e^{t_1^{\mathrm{T}}\mu_1 + \frac{1}{2}t_1^{\mathrm{T}}\Sigma_{11}t_1}.$$

Using the arguments in the theorem, one writes the MGF of X_1 as

$$M_{X_1}(t_1) = \left\{1 - \alpha^{*\mathrm{T}}\Sigma_{11}^{\frac{1}{2}}t_1 \frac{\left(2\theta - \alpha^{*\mathrm{T}}\Sigma_{11}^{\frac{1}{2}}t_1\right)}{\gamma^* + \theta^2 + \alpha^{*\mathrm{T}}\alpha^*}\left(\frac{\gamma^* + \theta^2 + \alpha^{*\mathrm{T}}\alpha^*}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\right)\right\} e^{t_1^{\mathrm{T}}\mu_1 + \frac{1}{2}t_1^{\mathrm{T}}\Sigma_{11}t_1}.$$

Note that

$$\frac{\gamma^* + \theta^2 + \alpha^{*\mathrm{T}}\alpha^*}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha} = \frac{\gamma + \theta^2 + \lambda_1^{\mathrm{T}}\ \Sigma_{11}\lambda_1 + \lambda_2^{\mathrm{T}}B_{22}^{-1}\lambda_2}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}.$$

Therefore, to complete the proof, it is enough to show that above ratio equals 1, i.e. $\lambda_1^{\mathrm{T}}\ \Sigma_{11}\lambda_1 + \lambda_2^{\mathrm{T}}B_{22}^{-1}\lambda_2 = \alpha^{\mathrm{T}}\alpha$. To this end, we write $\lambda_1^{\mathrm{T}}\ \Sigma_{11}\lambda_1 + \lambda_2^{\mathrm{T}}B_{22}^{-1}\lambda_2$ as

$$\lambda_1^{\mathrm{T}}\ \Sigma_{11}\lambda_1 + \lambda_2^{\mathrm{T}}B_{22}^{-1}\lambda_2 = \begin{pmatrix}\alpha_1^{\mathrm{T}} & \alpha_2^{\mathrm{T}}\end{pmatrix} K \begin{pmatrix}\alpha_1 \\ \alpha_2\end{pmatrix},$$

where K is given by

$$K = \begin{pmatrix} C_{11}\ \Sigma_{11} + \ C_{12}\ \Sigma_{21} \\ C_{21}\Sigma_{11} + \ \ C_{22}\Sigma_{21}\end{pmatrix}\left(C_{11} + \Sigma_{11}^{-1}\ \Sigma_{12}\ C_{21}\ C_{12} + \Sigma_{11}^{-1}\ \Sigma_{12}\ C_{22}\right) +$$

$$\begin{pmatrix} C_{12}B_{22}^{-1} \\ C_{22}B_{22}^{-1}\end{pmatrix}\left(C_{21}\ C_{22}\right).$$

So, it is sufficient to show that K is the identity matrix. Now the matrix K can be written as

$$K = \begin{pmatrix} C_{12}B_{22}^{-1}\ C_{21} & C_{12}B_{22}^{-1}\ C_{22} \\ C_{22}B_{22}^{-1}\ C_{21} & C_{12}B_{22}^{-1}\ C_{22}\end{pmatrix} +$$

$$\begin{pmatrix} (C_{11}\ \Sigma_{11} + \ C_{12}\ \Sigma_{21})\left(C_{11} + \Sigma_{11}^{-1}\ \Sigma_{12}\ C_{21}\right) & (C_{11}\ \Sigma_{11} + \ C_{12}\ \Sigma_{21})\left(C_{12} + \Sigma_{11}^{-1}\ \Sigma_{12}\ C_{22}\right) \\ (C_{21}\Sigma_{11} + \ \ C_{22}\Sigma_{21})\left(C_{11} + \Sigma_{11}^{-1}\ \Sigma_{12}\ C_{21}\right) & (C_{21}\Sigma_{11} + \ C_{22}\Sigma_{21})\left(C_{12} + \Sigma_{11}^{-1}\ \Sigma_{12}\ C_{22}\right)\end{pmatrix}.$$

After some simple algebra of matrices, we conclude that

$$K = \begin{pmatrix} C_{12} & C_{12} \\ C_{21} & C_{22}\end{pmatrix}\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}\begin{pmatrix} C_{12} & C_{12} \\ C_{21} & C_{22}\end{pmatrix} = \Sigma^{-\frac{1}{2}}\Sigma\Sigma^{-\frac{1}{2}} = I_n$$

which completes the proof.

(ii) Consider the same arguments that used in proof of part (i), we can write

$$f(x_1, x_2) = \frac{\gamma + \left(\theta - \lambda_1^{\mathrm{T}}(x_1 - \mu_1) - \lambda_2^{\mathrm{T}}(x_2 - \mu_{1.2})\right)^2}{\left(\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha\right)(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \times$$

$$e^{-\frac{1}{2}\left((x_1-\mu_1)^{\mathrm{T}}\Sigma_{11}^{-1}(x_1-\mu_1) + (x_2-\mu_{1.2})^{\mathrm{T}}\ B_{22}(x_2-\mu_{1.2})\right)}.$$

So, the PDF of $(X_2|X_1 = x_1)$ can be found as folows:

$$f_{(X_2|X_1=x_1)}(x_1,x_2) \propto \left\{\gamma + \left(\theta - \lambda_1^{\mathrm{T}}(x_1-\mu_1) - \lambda_2^{\mathrm{T}}(x_2-\mu_{1.2})\right)^2\right\}\times$$
$$e^{-\frac{1}{2}(x_2-\mu_{1.2})^{\mathrm{T}}\, B_{22}(x_2-\mu_{1.2})}$$
$$\propto \left\{\gamma + \left(\theta - \lambda_1^{\mathrm{T}}(x_1-\mu_1) - \lambda_2^{\mathrm{T}} B_{22}^{-\frac{1}{2}} B_{22}^{\frac{1}{2}}(x_2-\mu_{1.2})\right)^2\right\}\times$$
$$e^{-\frac{1}{2}(x_2-\mu_{1.2})^{\mathrm{T}}\, B_{22}(x_2-\mu_{1.2})}.$$

So, the proof is completed by writing the PDF of $(X_2|X_1 = x_1)$.

For $\mu = 0$ and $\Sigma = I_n$, we have the following corollary.

COROLLARY 1

Let $X \sim GASN_n\ (\gamma,\ \theta, \alpha)$. *If we partition* X *and* α *as*

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \ and\ \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix},$$

then

(i) The marginal distribution of X_1 *is*

$$X_1 \sim GASN_m\ (\gamma^*,\ \theta, \alpha_1),$$

where $\gamma^* = \gamma + \alpha_2^T \alpha_2$.

(ii) The conditional distribution of X_2 *given* $_1 = x_1$ *is*

$$(X_2|X_1 = x_1) \sim GASN_{n-m}\left(\gamma,\ \tilde{\theta}, \alpha_2\right),$$

where $\tilde{\theta} = \theta - \alpha_1^T x_1$.

PROOF:

(i) Let $t^{\mathrm{T}} = (t_1^{\mathrm{T}} \quad t_2^{\mathrm{T}})$ such that t_1^{T} and t_2^{T} are two vectors of dimension $1 \times m$ and $1 \times n - m$, respectevly. Using the MGF of X and set $t_2 = 0$, we write

$$M_{X_1}(t_1) = \left\{1 - \alpha_1^{\mathrm{T}} t_1 \frac{\left(2\theta - \alpha_1^{\mathrm{T}} t_1\right)}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\right\} e^{\frac{1}{2} t_1^{\mathrm{T}} t_1},$$
$$= \left\{1 - \alpha_1^{\mathrm{T}} t_1 \frac{\left(2\theta - \alpha_1^{\mathrm{T}} t_1\right)}{\gamma + \theta^2 + \alpha_1^{\mathrm{T}}\alpha_1 + \alpha_2^{\mathrm{T}}\alpha_2}\right\} e^{\frac{1}{2} t_1^{\mathrm{T}} t_1},$$
$$= \left\{1 - \alpha_1^{\mathrm{T}} t_1 \frac{\left(2\theta - \alpha_1^{\mathrm{T}} t_1\right)}{\gamma^* + \theta^2 + \alpha_1^{\mathrm{T}}\alpha_1}\right\} e^{\frac{1}{2} t_1^{\mathrm{T}} t_1},$$

which completes the proof.

(ii) Using (4) and the arguments in the corollary, then we have

$$f(x_1,x_2) = \frac{\gamma + \left(\theta - \alpha_1^{\mathrm{T}} x_1 - \alpha_2^{\mathrm{T}} x_2\right)^2}{\left(\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha\right)(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\left(x_1^{\mathrm{T}} x_1 + x_2^{\mathrm{T}} x_2\right)}.$$

The conditional PDF of X_2 given $X_1 = x_1$ ($f_{X_2|X_1}(x_2|x_1)$) can be found as

$$\begin{aligned} f_{X_2|X_1}(x_2|x_1) &= \frac{\gamma + \left(\theta - \alpha_1^{\mathrm{T}} x_1 - \alpha_2^{\mathrm{T}} x_2\right)^2}{\left(\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha\right)(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\left(x_1^{\mathrm{T}} x_1 + x_2^{\mathrm{T}} x_2\right)}, \\ &\propto \left(\gamma + \left(\theta - \alpha_1^{\mathrm{T}} x_1 - \alpha_2^{\mathrm{T}} x_2\right)^2\right) e^{-\frac{1}{2}\left(x_2^{\mathrm{T}} x_2\right)}, \\ &\propto \left(\gamma + \left(\tilde{\theta} - \alpha_2^{\mathrm{T}} x_2\right)^2\right) e^{-\frac{1}{2}\left(x_2^{\mathrm{T}} x_2\right)}, \\ &= \frac{\left(\gamma + \left(\tilde{\theta} - \alpha_2^{\mathrm{T}} x_2\right)^2\right)}{\left(\gamma + \tilde{\theta}^2 + \alpha_2^{\mathrm{T}}\alpha_2\right)(2\pi)^{\frac{n-m}{2}}} e^{-\frac{1}{2}\left(x_2^{\mathrm{T}} x_2\right)}. \end{aligned}$$

The new generalization is closed under affine transformation as it is shown in the Theorem 3.

THEOREM 3
Let c be an $m \times 1$ real valued vector and A is an $m \times n$ matrix of rank m, where $m < n$.

Define the random vector $W = c + AY$, where $Y \sim GASN_n(\gamma, \theta, \alpha, \mu, \Sigma)$. Then $W \sim GASN_m\left(\gamma, \theta, \dot{\alpha}, c + A\mu, A\Sigma A^T\right)$, where $\dot{\alpha} = \left(A\Sigma A^T\right)^{-\frac{1}{2}} A\Sigma^{\frac{1}{2}}\alpha$.

PROOF:
The MGF of $W = c + AY$ is

$$\begin{aligned} M_W(t) &= E\left(e^{t^{\mathrm{T}} W}\right) = e^{t^{\mathrm{T}} c} M_Y\left(A^{\mathrm{T}} t\right), \\ &= e^{t^{\mathrm{T}} c}\left\{1 - \alpha^{\mathrm{T}}\Sigma^{\frac{1}{2}} A^{\mathrm{T}} t \frac{\left(2\theta - \alpha^{\mathrm{T}}\Sigma^{\frac{1}{2}} A^{\mathrm{T}} t\right)}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\right\} e^{t^{\mathrm{T}} A\mu + \frac{1}{2} t^{\mathrm{T}} A\Sigma A^{\mathrm{T}} t}, \\ &= \left\{1 - \alpha^{\mathrm{T}}\Sigma^{\frac{1}{2}} A^{\mathrm{T}} t \frac{\left(2\theta - \alpha^{\mathrm{T}}\Sigma^{\frac{1}{2}} A^{\mathrm{T}} t\right)}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\right\} e^{t^{\mathrm{T}}(c + A\mu) + \frac{1}{2} t^{\mathrm{T}} A\Sigma A^{\mathrm{T}} t}, \\ &= \left\{1 - \dot{\alpha}^{\mathrm{T}}\left(A\Sigma A^{\mathrm{T}}\right)^{\frac{1}{2}} t \frac{\left(2\theta - \dot{\alpha}^{\mathrm{T}}\left(A\Sigma A^{\mathrm{T}}\right)^{\frac{1}{2}} t\right)}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\right\} e^{t^{\mathrm{T}}(c + A\mu) + \frac{1}{2} t^{\mathrm{T}} A\Sigma A^{\mathrm{T}} t}. \end{aligned}$$

By noting the orthogonality of the matrix $\left(A\Sigma A^{\mathrm{T}}\right)^{-\frac{1}{2}} A\Sigma^{\frac{1}{2}}$, we can write $\dot{\alpha}^{\mathrm{T}}\dot{\alpha} = \alpha^{\mathrm{T}}\Sigma^{\frac{1}{2}} A^{\mathrm{T}}\left(A\Sigma A^{\mathrm{T}}\right)^{-\frac{1}{2}}\left(A\Sigma A^{\mathrm{T}}\right)^{-\frac{1}{2}} A\Sigma^{\frac{1}{2}}\alpha = \alpha^{\mathrm{T}}\alpha$. So, the proof is completed.

The quadratic forms of skew normal distributions are of central interest in literature. In Theorem 4, we derive the MGF for the quadratic form of $X \sim GASN_n(\gamma, \theta, \alpha)$.

THEOREM 4

Let $X \sim GASN_n\ (\gamma,\ \theta, \alpha, \mu, \Sigma)$. Define $Q = (X-\mu)^T A(X-\mu)$, where A is an $n \times n$ positive definite matrix. For all $t \in R$ such that the matrix $\Sigma^{-1} - 2tA$ is positive definite, the MGF of Q is

$$M_Q(t) = \frac{\left(\gamma + \theta^2 + \alpha^T\left(I_n - 2t\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}\right)^{-1}\alpha\right)\left|\Sigma^{-1} - 2tA\right|^{-\frac{1}{2}}}{\left(\gamma + \theta^2 + \alpha^T\alpha\right)|\Sigma|^{\frac{1}{2}}}.$$

PROOF:

Using the definition of the MGF, we write

$$M_Q(t) = E(\exp(tQ)) = \int_{R^n} \exp(tQ)\frac{\gamma + \left(\theta - \alpha^{\mathrm{T}}\Sigma^{-\frac{1}{2}}(x-\mu)\right)^2}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\varphi_n(x;\mu,\Sigma)dx.$$

Since $Q = (X-\mu)^{\mathrm{T}}A(X-\mu)$, then

$$M_Q(t) = \int_{R^n} \frac{\gamma + (\theta - \alpha^{\mathrm{T}}\Sigma^{-\frac{1}{2}}(x-\mu))^2}{\left(\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha\right)(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}\left((x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu) - 2t(x-\mu)^{\mathrm{T}}A(x-\mu)\right)}dx,$$

$$= \int_{R^n} \frac{\gamma + (\theta - \alpha^{\mathrm{T}}\Sigma^{-\frac{1}{2}}(x-\mu))^2}{\left(\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha\right)(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}\left((x-\mu)^{\mathrm{T}}\left(\Sigma^{-1} - 2tA\right)(x-\mu)\right)}dx.$$

To simplify the calculation, let $B_t^{-1} = \Sigma^{-1} - 2tA$. Then

$$M_Q(t) = \int_{R^n} \frac{\gamma + (\theta - \alpha^{\mathrm{T}}\Sigma^{-\frac{1}{2}}(x-\mu))^2}{\left(\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha\right)(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}\left((x-\mu)^{\mathrm{T}}B_t^{-1}(x-\mu)\right)}dx.$$

The integral can be simplified as follows:

$$M_Q(t) = \frac{|B_t|^{\frac{1}{2}}}{\left(\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha\right)|\Sigma|^{\frac{1}{2}}}\int_{R^n} \frac{\gamma + (\theta - \alpha^{\mathrm{T}}\Sigma^{-\frac{1}{2}}B_t^{\frac{1}{2}}B_t^{-\frac{1}{2}}(x-\mu))^2}{(2\pi)^{\frac{n}{2}}|B_t|^{\frac{1}{2}}} \times$$

$$e^{-\frac{1}{2}\left((x-\mu)^{\mathrm{T}}B_t^{-1}(x-\mu)\right)}dx,$$

$$= \frac{\left(\gamma + \theta^2 + \alpha^{\mathrm{T}}\Sigma^{-\frac{1}{2}}B_t\Sigma^{-\frac{1}{2}}\alpha\right)|B_t|^{\frac{1}{2}}}{\left(\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha\right)|\Sigma|^{\frac{1}{2}}}\int_{R^n} \frac{\gamma + (\theta - \alpha^{\mathrm{T}}\Sigma^{-\frac{1}{2}}B_t^{\frac{1}{2}}B_t^{-\frac{1}{2}}(x-\mu))^2}{\left(\gamma + \theta^2 + \alpha^{\mathrm{T}}\Sigma^{-\frac{1}{2}}B_t\Sigma^{-\frac{1}{2}}\alpha\right)(2\pi)^{\frac{n}{2}}|B_t|^{\frac{1}{2}}} \times$$

$$e^{-\frac{1}{2}\left((x-\mu)^{\mathrm{T}}B_t^{-1}(x-\mu)\right)}dx.$$

Notice that the integrand is the PDF of $X \sim GASN_n\left(\gamma, \theta, B_t^{\frac{1}{2}}\Sigma^{-\frac{1}{2}}\alpha, \mu, B_t\right)$. So, the $M_Q(.)$ is

$$M_Q(t) = \frac{\gamma + \theta^2 + \alpha^{\mathrm{T}}\Sigma^{-\frac{1}{2}}B_t\Sigma^{-\frac{1}{2}}\alpha}{(\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha)|\Sigma|^{\frac{1}{2}}}|B_t|^{\frac{1}{2}}.$$

Using the definition of B_t, then

$$\Sigma^{-\frac{1}{2}}B_t\Sigma^{-\frac{1}{2}} = \Sigma^{-\frac{1}{2}}\left(\Sigma^{-1} - 2tA\right)^{-1}\Sigma^{-\frac{1}{2}} = \left(I_n - 2t\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}\right)^{-1}.$$

Hence

$$M_Q(t) = \frac{\gamma + \theta^2 + \alpha^{\mathrm{T}}\left(I_n - 2t\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}\right)^{-1}\alpha}{(\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha)|\Sigma|^{\frac{1}{2}}}\left|\Sigma^{-1} - 2tA\right|^{-\frac{1}{2}}.$$

So, the proof is completed.

The generalized alpha skew-normal distribution is closed under the convolution with the multivariate normal distribution as shown in the following theorem.

THEOREM 5
Let $X_1 \sim GASN_n(\gamma, \theta, \alpha, \mu, A)$ be independent of $X_2 \sim N_n(v, \Sigma)$. Define $S = X_1 + X_2$. Then $S \sim GSN_n(\gamma^*, \theta, \dot{\alpha}, \mu + v, \Sigma + A)$, where $\gamma^* = \gamma + \breve{\alpha}^T\breve{\alpha}$, $\breve{\alpha}^T = \alpha^{-\frac{1}{2}}(\Sigma^{-1} + A^{-1})^{-\frac{1}{2}}$ and $\dot{\alpha}^T = \alpha^T A^{\frac{1}{2}}(\Sigma + A)^{-\frac{1}{2}}$.

PROOF:
Since X_1 and X_2 are independent then by using their MGF one writes the MGF of S as

$$M_S(t) = M_{X_1}(t)M_{X_2}(t),$$

where

$$M_{X_1}(t) = \left\{1 - \alpha^{\mathrm{T}}A^{\frac{1}{2}}t\frac{\left(2\theta - \alpha^{\mathrm{T}}A^{\frac{1}{2}}t\right)}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\right\}e^{t^{\mathrm{T}}\mu + \frac{1}{2}t^{\mathrm{T}}At}$$

and

$$M_{X_2}(t) = e^{t^{\mathrm{T}}v + \frac{1}{2}t^{\mathrm{T}}\Sigma t}.$$

Therefore, one can easily write

$$\begin{aligned} M_S(t) &= \left\{1 - \alpha^{\mathrm{T}}A^{\frac{1}{2}}t\frac{\left(2\theta - \alpha^{\mathrm{T}}A^{\frac{1}{2}}t\right)}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\right\}e^{t^{\mathrm{T}}(\mu+v) + \frac{1}{2}t^{\mathrm{T}}(A+\Sigma)t}, \\ &= \left\{1 - \dot{\alpha}^{\mathrm{T}}(A+\Sigma)^{\frac{1}{2}}t\frac{\left(2\theta - \dot{\alpha}^{\mathrm{T}}(A+\Sigma)^{\frac{1}{2}}t\right)}{\gamma + \theta^2 + \alpha^{\mathrm{T}}\alpha}\right\}e^{t^{\mathrm{T}}(\mu+v) + \frac{1}{2}t^{\mathrm{T}}(A+\Sigma)t}. \end{aligned}$$

Note that the term $\alpha^{\mathrm{T}}\alpha$ in the denominator can be written as

$$\alpha^{\mathrm{T}}\alpha = \alpha^{\mathrm{T}}\left(I_n - A^{-\frac{1}{2}}\left(\Sigma^{-1}+A^{-1}\right)^{-1}A^{-\frac{1}{2}} + A^{-\frac{1}{2}}\left(\Sigma^{-1}+A^{-1}\right)^{-1}A^{-\frac{1}{2}}\right)\alpha,$$

$$= \alpha^{T}\left(I_n - A^{-\frac{1}{2}}(\Sigma^{-1}+A^{-1})^{-1}A^{-\frac{1}{2}}\right)\alpha + \breve{\alpha}^{T}\breve{\alpha}.$$

Now using the identity $\left(\Sigma^{-1}+A^{-1}\right)^{-1} = A - A(\Sigma+A)^{-1}A$, we get

$$\alpha^{T}\alpha = \dot{\alpha}^{T}\dot{\alpha} + \breve{\alpha}^{T}\breve{\alpha}.$$

Hence, the $M_S(t)$ takes the form

$$M_S(t) = \left\{1 - \dot{\alpha}^{\mathrm{T}}(A+\Sigma)^{\frac{1}{2}}t\frac{\left(2\theta - \dot{\alpha}^{\mathrm{T}}(A+\Sigma)^{\frac{1}{2}}t\right)}{\gamma^{*}+\theta^{2}+\dot{\alpha}^{\mathrm{T}}\dot{\alpha}}\right\}e^{t^{\mathrm{T}}(\mu+\nu)+\frac{1}{2}t^{\mathrm{T}}(A+\Sigma)t},$$

which completes the proof.

5. MARGINAL AND CONDITIONAL DISTRIBUTIONS

In this section, we consider a data set, which consists of $n = 500$ observations simulated from $GASN_n\ (\gamma,\ \theta,\alpha,\mu,\Sigma)$ with $\gamma = 5,\ \theta = 2,\ \alpha = (2,\ 1)^{T},\ \mu = (4,\ 2)^{T}$ and $\Sigma = \begin{pmatrix} 3 & 1 \\ 1 & 5 \end{pmatrix}$. We used the following Metropolis-Hasting algorithm to simulate from $GASN_n\ (\gamma,\ \theta,\alpha,\mu,\Sigma)$:

(i) Given a starting point $x^{(t-1)}$, simulate $Y^{(t)} \sim N_2\ \left(x^{(t-1)};\ \mu,\Sigma\right)$

(ii) $X^{(t)} = \begin{cases} Y^{(t)} & \text{with probability} \quad \rho\left(X^{(t)},Y^{(t)}\right), \\ X^{(t)} & \text{with probability} \quad 1-\rho\left(X^{(t)},Y^{(t)}\right), \end{cases}$

where

$$\rho\left(X^{(t)},Y^{(t)}\right) = \min\left(1,\ \frac{f_Y\left(Y^{(t)};\gamma,\theta,\alpha,\mu,\Sigma\right)}{f_Y\left(X^{(t)};\gamma,\theta,\alpha,\mu,\Sigma\right)}\right).$$

Then the maximum likelihood estimators of the parameters $\gamma,\ \theta,\alpha,\mu$ and are obtained by maximizing the following loglikhood function via the R function *auglag()*. The R function *MultiStart()* is used to search for suitable initial parameter values.

$$\log f_X(x;\gamma,\theta,\alpha,\mu,\Sigma) = -\ \frac{n}{2}\log\det(\Sigma) - \frac{n}{2}\log\det(2\pi) - n\log\left(\gamma+\theta^{2}+\alpha^{\mathrm{T}}\alpha\right)$$

$$-\frac{1}{2}\sum_{j=1}^{n}\left(x_j-\mu\right)^{T}\Sigma^{-1}\left(x_j-\mu\right) + \sum_{j=1}^{n}\log(\gamma+(\theta-\alpha^{\mathrm{T}}\Sigma^{-\frac{1}{2}}(x_j-\mu))^{2})$$

where $x_j, j = 1,\ 2,\ldots,500$, are the simulated points.

The data produced the following estimates: $\hat{\gamma} = 5.2158$, $\hat{\theta} = 2.013$, $\hat{\alpha} = (1.1082,\ 1.0464)^{T}$, $\hat{\mu} = (3.9198,\ 1.8961)^{T}$ and $\hat{\Sigma} = \begin{pmatrix} 3.0862 & 0.9058 \\ 0.9058 & 4.7007 \end{pmatrix}$. In

Figure 3, we plot the data points together with eight contours of the $GASN_2(\hat{\gamma},\ \hat{\theta}, \hat{\alpha}, \hat{\mu}, \hat{\Sigma})$ and its marginal distributions. As it can be seen that the majority of the data points are distributed along the contours of $GASN_2(\hat{\gamma},\ \hat{\theta}, \hat{\alpha}, \hat{\mu}, \hat{\Sigma})$ distribution. For further comparison, we obtained the modes for both data set and for the distribution $GASN_2(\hat{\gamma},\ \hat{\theta}, \hat{\alpha}, \hat{\mu}, \hat{\Sigma})$. The data modes are estimated as the mid-point of the modal class. The bivariate mode for the data was $(2.65,\ 1.8)$ and the modes for the two marginal frequency histograms were 2.52 and 1.64, respectively. Also we obtained the mode of the $GASN_2(\hat{\gamma},\ \hat{\theta}, \hat{\alpha}, \hat{\mu}, \hat{\Sigma})$ by maximizing its joint pdf with respect to x_1 and x_2 which produced the estimate $(2.451,\ 1.675)$. It can be noticed, form these numerical results, that the data measures are close to the distribution measures, which reflect an accepted fit for the data by the $GASN_2$.

It can be noticed, from the Figure 3 and the parameters estimates, that the distribution *GASN* captures almost all the variation in the data set and other characteristics such as location and scale parameters, in additions to modes. This recommends the *GASN* distribution to be a good candidate for analyzing skewed data sets.

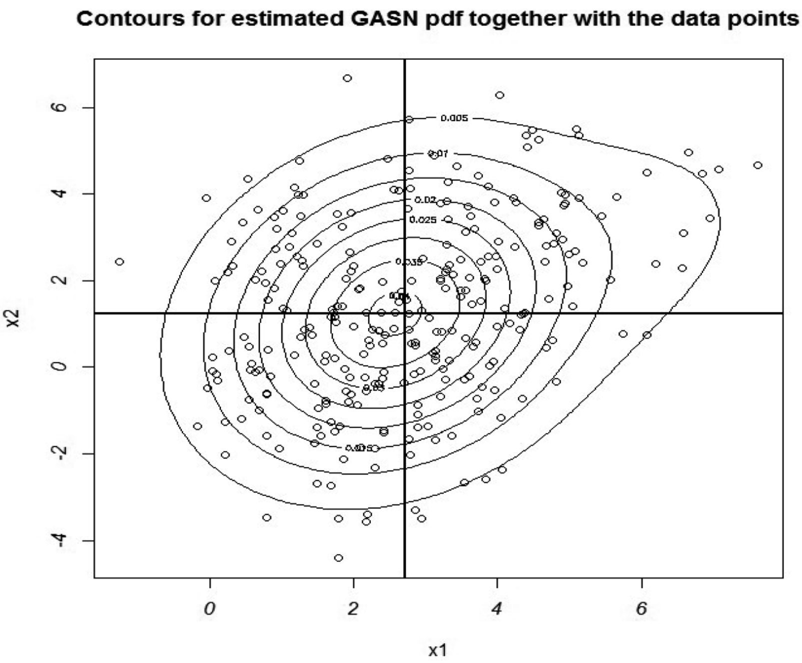

FIGURE 3. - *Eight contours of* $GASN_2\left(\hat{\gamma},\ \hat{\theta}, \hat{\alpha}, \hat{\mu}, \hat{\Sigma}\right)$*, together with the data points*

6. CONCLUSIONS

We generalized the work of Elal-Olivero (2010) on the ASN distribution to multivariate version. We have shown that the GASN distribution such enjoys several nice statistical properties such as the marginal, the conditional distribution and the closure under convolution with normal random variate are studied. Also, the generalization of Alpha skew normal distribution defined by Handam (2012) can be applied for the distribution given in (4). Another generalization to (4) can be obtained by using matrices instead of the vectors α and x, so that a matrix variate version of (4) can be obtained.

ACKNOWLEDGEMENT

The authors are also grateful to the editor and referees for their suggestions that helped to improve the article.

REFERENCES

Azzalini A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, **12**, 171-178.

Azzalini A., Capitanio A. (2014). *The Skew-Normal and Related Families*, with the collaboration of Capitanio A, IMS monographs. Cambridge University Press, Cambridge.

Azzalini A., Dalla Valle A. (1996). The multivariate skew-normal distribution. *Biometrika*, **83**, 715-726.

Bayramoglu I. (2020). Bivariate and multivariate distributions with bimodal marginals. *Communications in Statistics-Theory and Methods*, **29**(2), 361-384.

Elal-Olivero A. (2010). Alpha-*skew*-normal distribution. *Proyecciones Journal of Mathematics,* **29**(3), 224-240.

Genton M.G. (Ed). (2004). *Skew-Elliptical Distributions and their Applications: A Journey Beyond Normality*. Chapman & Hall/CRC.

Handam A.H. (2012). A note on generalized alpha-skew-normal distribution. *International Journal of Pure and Applied Mathematics*, **74**(4), 491-496.

Louzada F., Ara A., Fernandes G. (2017). The bivariate alpha-skew-normal distribution. *Communication in Statistics-Theory and Methods*, **46**(14), 7147-7156.

Pewsey A. (2000). Problems of inference for Azzalini's skew-normal distribution. *Journal of Applied Statistics*, **27**(7), 859-870.

Sturrock P.A. (2008). Analysis of bimodality in histograms formed from GALLEX and GNO solar neutrino data. *Solar Physics*, **249**(1), 1-9.

Verbeek M. (2004). *A Guide to Modern Econometrics*, John Wiley and Sons, Chichester.