On Fundamental Skew Distributions

Reinaldo B. Arellano-Valle(*) and Marc G. Genton(**)

(*) Departamento de Estadística, Pontificia Universidad Católica de Chile
Casilla 306, Santiago 22, Chile.

(**) Department of Statistics, 209-D, North Carolina State University
Raleigh, NC 27695-8203, USA.

August 19, 2003

Abstract

A new class of multivariate skew-normal distributions, fundamental skew-normal distributions, is developed. It contains the product of independent univariate skew-normal distributions as a special case. Stochastic representations and other main properties of the associated distribution theory of linear and quadratic forms are considered. A unified procedure for extending this class to other families of skew distributions such as the skew-spherical and skew-elliptical class of distributions is also discussed.

Key words: Fundamental generalized skew distribution; marginal and conditional distributions; moment generating function; normal, spherical and elliptical distributions; quadratic forms; stochastic representation.

1 Introduction

During the last decade, there has been an increasing interest in finding more flexible methods to represent features of the data as adequately as possible and to reduce unrealistic assumptions. The motivation originates from data sets, including environmental, financial and biomedical ones, which often do not satisfy some standard assumptions such as independence and normality. In order to model departures from these assumptions, and particularly from the normal assumption, which played an overwhelming role in multivariate analysis, several approaches are available in the literature.

From a practical viewpoint, maybe the most commonly adopted approach is transformation of the variables in order to achieve multivariate normality (or symmetry). However, the transformed variables are more difficult to interpret, especially when each variable is transformed by using different functions.

An alternative approach for data modeling from a parametric standpoint, which is particularly appropriate for the treatment of continuous multivariate observations, consists in constructing
flexible parametric classes of multivariate distributions that exhibit skewness and kurtosis which is different from the normal distribution. The class of elliptical distributions is probably the most popular example of this approach. This class was first introduced by Kelker (1970) and systematically discussed by Cambanis et al. (1981). For a comprehensive review, see Fang et al. (1990). The elliptical class includes a vast set of known distributions, for example, normal, compound normal, Student-$t$, power exponential and Pearson type II, among others. Its main advantage is that it represents a natural extension of the concept of symmetry in the multivariate setting. Although elliptical models provide alternatives to the normal model, these can only be applied in practical situations where the symmetry seems reasonable. Therefore, the construction of parametric families of asymmetric distributions which are analytically tractable, can accommodate practical values of skewness and kurtosis, and strictly include the normal distribution, can be useful for data modeling, statistical analysis, and robustness studies of normal theory methods. This work is focused on the study of multivariate skew distributions from a unified approach, and also examines the main properties of such models.

1.1 Azzalini’s approach and its extensions

Although the idea of modeling skewness by means of the construction of a mathematically tractable family including the normal distribution was proposed early by other authors (see, for example, O’Hagan and Leonard, 1976), the formal definition of the univariate skew-normal (SN, hereafter) family is due to Azzalini (1985). He said that a random variable $Z$ has an SN distribution with asymmetric parameter $\lambda$, which is denoted by $Z \sim SN(\lambda)$, if its density is

$$f(z \mid \lambda) = 2 \phi(z) \Phi(\lambda z), \quad z \in \mathbb{R}, \quad \lambda \in \mathbb{R},$$

where $\phi$ and $\Phi$ are the $N(0,1)$ probability density function (pdf) and cumulative distribution function (cdf), respectively. The case $\lambda = 0$ reduces (1.1) to the $N(0,1)$ density. Further properties are studied by Azzalini (1985, 1986) and Henze (1986). They show, in particular, that if $Z \sim SN(\lambda)$, then $Z^2 \sim \chi_1^2$ and

$$Z \overset{d}{=} \frac{\lambda}{\sqrt{1+\lambda^2}} |X| + \frac{1}{\sqrt{1+\lambda^2}} Y,$$

where $X$ and $Y$ are iid $N(0,1)$ random variables and the notation $X \overset{d}{=} Y$ means that $X$ and $Y$ have the same distribution. Later, Azzalini and Dalla Valle (1996) use the stochastic representation (1.2) to extend (1.1) to the multivariate $SN$ family of densities, which are given by

$$f(z \mid \lambda) = 2 \phi_k(z) \Phi_k(\lambda^T z), \quad z \in \mathbb{R}^k, \quad \lambda \in \mathbb{R}^k,$$

where $\phi_k$ and $\Phi_k$ are the pdf and cdf of the $k$-dimensional normal distribution $N_k(0, I_k)$, respectively. Again, the case with $\lambda = 0$, reduces (1.3) to the $N_k(0, I_k)$ density, so that $\lambda$ is interpreted as a shape vector of parameters. Extensions to multivariate location-scale $SN$ distributions are also considered in Azzalini and Dalla Valle (1996). Further properties of the multivariate $SN$ distribution are studied in Azzalini and Capitanio (1999). Subsequently, Genton et al. (2001)
derive the moments of a random vectors with multivariate SN distributions and their quadratic forms (see also Loperfido, 2001).

Generalizations of these ideas have been proposed by many authors. For instance, multivariate distributions such as skew-Cauchy (Arnold and Beaver, 2000), skew-t (Branco and Dey, 2001; Azzalini and Capitanio, 2003; Jones and Faddy, 2003; Sahu et al., 2003), skew-logistic (Wahed and Ali, 2001), and other skew-elliptical ones (Azzalini and Capitanio, 1999; Branco and Dey, 2001; Arnold and Beaver, 2002). Sahu et al. (2003) provided a more general way to obtain a family of multivariate skew-elliptical distributions, from which a multivariate version of the univariate SN distribution with independent SN marginals can be obtained. Recently, Genton and Loperfido (2002) introduced a class of generalized skew-spherical (elliptical) distributions defined by densities of the form

$$ f(z|Q) = 2f_k(z)Q(z), \quad z \in \mathbb{R}^k, \quad (1.4) $$

where $f_k$ is the density corresponding to a $k$-dimensional spherical (elliptical) distribution (see Fang et al., 1990) and $Q$ is a skewing function, which is such that $Q(z) \geq 0$ and $Q(-z) = 1 - Q(z)$, for all $z \in \mathbb{R}^k$. They show that many of the SN properties can be extend to any distribution in this class.

**Remark 1.1** Strictly speaking, we must note in (1.4) that $Q(z) = v(u(z))$, for some function $u : \mathbb{R}^k \to \mathbb{R}$ and some non-negative function $v : \mathbb{R} \to \mathbb{R}$, which are such that $u(-z) = -u(z)$, for all $z \in \mathbb{R}^k$, and $v(-u) = 1 - v(u)$, for all $u \in \mathbb{R}$. This alternative representation is used by Azzalini and Capitanio (2003) but it is not unique.

More recently, Wang et al. (2002) extended (1.4) to any symmetric density $f_k$ in $\mathbb{R}^k$, that is, assuming that $f_k$ satisfies the condition $f_k(-z) = f_k(z)$ for all $z \in \mathbb{R}^k$. Further properties and characterizations of these distributions are also discussed in their work. For example, it is shown that under (1.4) the associated distribution theory of linear and quadratic forms remains largely valid. Thus, many multivariate extensions of the univariate SN distribution of Azzalini can be obtained as special cases of (1.4), for example, the multivariate SN distribution defined in (1.1) by Azzalini and Dalla Valle (1996) and that introduced by Azzalini and Capitanio (1999). Moreover, (1.4) generalizes the class of skew-spherical (elliptical) distributions considered in Branco and Dey (2001). However, from (1.4) it is not possible to obtain neither the multivariate family of the skew-spherical (elliptical) distributions considered by Sahu et al. (2003), nor the class of multivariate SN distributions considered by Gupta et al. (2001) and fully discussed by González-Farías et al. (2002) (see also Liseo and Loperfido, 2002), whose canonical version is defined from its density by

$$ f(z|\Phi_m,D) = \frac{\phi_k(z)\Phi_m(Dz)}{\Phi_m(0|I_m + DD^T)}, \quad z \in \mathbb{R}^k, \quad (1.5) $$

where $\phi_p(\cdot|\mu,\Sigma)$ and $\Phi_p(\cdot|\mu,\Sigma)$ are respectively the pdf and cdf of the $N_p(\mu,\Sigma)$ distribution, $\phi_p(\cdot|\Sigma)$ and $\Phi_p(\cdot|\Sigma)$ denote these functions when $\mu = 0$, and $D$ is a matrix of dimension $m \times k$. 
Note that (1.5) contains also the multivariate SN family defined by (1.3) when \( m = 1 \) and reduces to the product of univariate SN distributions when \( m = k \) and \( D \) is diagonal.

Important additional results are given in Arellano-Valle et al. (2002), where a general class of skew-symmetric distributions is introduced starting from the definition of a special \( C \)-class of symmetric distributions (or random vectors). Two equivalent stochastic representations, the conditional and marginal representations, are given there for any skew random vector obtained from that \( C \)-class. Following the results of these authors, we consider here the following general procedure to obtain the density of an arbitrary skew distribution:

\[
f(z|Q_m) = K_m^{-1} f_k(z) Q_m(z), \quad z \in \mathbb{R}^k,
\]

where

\[
K_m = \Pr(X > 0) \quad \text{and} \quad Q_m(z) = \Pr(X > 0|Z = z),
\]

for some random vectors \( X \) and \( Z \) with dimensions \( m \times 1 \) and \( k \times 1 \), respectively, and with joint distribution such that \( Z \) has marginal density \( f_k \). Note that (1.6) corresponds to the conditional density of \( [Z, X > 0] \). Thus, as was noted by Arellano-Valle et al. (2002), if \( X \) is a \( C \)-random vector, then \( K_m = \Pr(X > 0) = 2^{-m} \), so that (1.6) reduces to

\[
f(z, Q_m) = 2^m f_k(z) Q_m(z), \quad z \in \mathbb{R}^k.
\]

If \( f_k \) is a symmetric density on \( \mathbb{R}^k \), then \( Q_m \) can be interpreted as a skewing function, which in general does not satisfy the condition given by Genton and Loperfido (2002) for the family defined by (1.4). In fact, these conditions are satisfied for \( m = 1 \) only, and in this particular case (1.8) is contained in (1.4). For \( m > 1 \), (1.8) is more general than (1.4). We note that (1.5) may be obtained from (1.6)-(1.7) by taking \( X \sim N_m(0, I_m + DD^T) \) and \( Z \sim N_k(0, I_k) \), with \( \text{Cov}(X, Z) = D \). Thus, \( f_k(z) = \phi_k(z), X|Z = z \sim N_m(Dz, I_m) \), so that \( Q_m(z) = \Pr(X > 0|Z = z) = \Phi_m(Dz) \) and \( K_m = \Pr(X > 0) = \Phi_m(0|I_m + DD^T) \). Note that \( K_m = 2^{-m} \) if \( DD^T \) is a diagonal matrix. Similarly, if we suppose that \( X \sim N_m(0, I_m) \) and \( Z \sim N_k(0, \Omega + DD^T) \), with \( \text{Cov}(Z, X) = D \), then from (1.8) we have that

\[
f(z, \Phi_m, \Omega, D) = 2^m \phi_k(z|\Omega + DD^T) \Phi_m(D^T(\Omega + DD^T)^{-1}z|I_m - D^T(\Omega + DD^T)^{-1}D),
\]

which generalizes the results of Sahu et al. (2003) SN distribution with null location vector. In fact, these authors consider (1.9) with \( m = k \) and assume that \( \Omega \) and \( D \) are diagonal matrices.

Remark 1.2 Note in (1.6)-(1.8) that \( Q_m(z) = v(u(z)) \), for some function \( u: \mathbb{R}^k \to \mathbb{R}^m \), which is such that \( u(-z) = -u(z), \) for all \( z \in \mathbb{R}^k \), and some non-negative function \( v: \mathbb{R}^m \to \mathbb{R} \).

1.2 Outline

The main objective of this work is to study the family of skew distributions that follows from (1.6) when in (1.7) the random vectors \( X \) and \( Z \) have a particular symmetric joint distribution. We start by assuming that \( X \) and \( Z \) are normally distributed. Thus, we introduce in Section
the so called multivariate fundamental skew-normal distributions ("fundamental" because it generalizes existing definitions of $SN$ distributions that are special cases of the form given in (1.8) with normalizing constant $K_m = 2^{-m}$). We derive for this class several results associated with the distributional theory of linear, marginal, conditional and quadratic forms. A marginal stochastic representation for a fundamental skew-normal random vector is also considered here. Such representation is used to obtain the moments of these skew-normal random vectors, and may be used also to develop simulation studies related to inferential aspects for these models. Some extensions of these ideas are considered in Section 3, where we show that the general family of skew density defined by (1.6)-(1.7) is closed under marginalization and conditioning. We explore also the special case when the random vector $Z$ is assumed to be the linear combination $AX + BY$, where $X$ and $Y$ are uncorrelated and symmetrically distributed random vectors. Finally, in Section 4 we apply the results obtained in Section 3 to the case where $X$ and $Y$ are spherically distributed, thus defining the class of fundamental skew-spherical distributions.

2 The fundamental skew-normal distribution

In this section, we introduce a new class of $SN$ distributions and we study their main properties. Consider first the univariate family of $SN$ distributions defined by means of its pdf in (1.1). Let $\delta = \lambda/\sqrt{1 + \lambda^2}$. Since the relation between $\lambda \in \mathbb{R}$ and $\delta \in (-1, 1)$ is one to one, we have that \{2$\phi(x)$/$\Phi(\lambda x), x \in \mathbb{R}; \lambda \in \mathbb{R}$\} and \{2$\phi(x)$/$\Phi(\delta x|1 - \delta^2), x \in \mathbb{R}; |\delta| < 1$\} are equivalent families of $SN$-pdfs. Here, the first parametrization will be denoted by \{$SN(\lambda); \lambda \in \mathbb{R}$\} and the second by \{$FSN(\delta); |\delta| < 1$\}.

Let $Z^*$ be a $k \times 1$ random vector and let $\Delta$ be a $k \times m$ correlation matrix (which is a symmetric matrix when $m = k$) such that $I_m - \Delta^T \Delta$ and $I_k - \Delta \Delta^T$ are positive definite matrices, i.e., $\|\Delta a\| < 1$ and $\|\Delta^T b\| < 1$ for any unitary pair of vectors $a \in \mathbb{R}^m$ and $b \in \mathbb{R}^k$, where $\| \cdot \|$ denotes the norm of a vector.

**Definition 2.1** We say that $Z^*$ has a $k$-variate fundamental skew-normal ($FSN$) distribution with a $k \times m$ skewness matrix $\Delta$, which will be denoted by $Z^* \sim FSN_{k,m}(\Delta)$ and by $Z^* \sim FSN_k(\Delta)$ when $m = k$, if its density is given by

$$f_{Z^*}(z) = 2^m \phi_k(z) \Phi_m(\Delta^T z | I_m - \Delta^T \Delta), \quad z \in \mathbb{R}^k,$$

where $\Delta$ is such that $\|\Delta a\| < 1$, for all unitary vectors $a \in \mathbb{R}^m$.

**Remark 2.1** Note that the matrix $\Delta$ can be constructed as $\Delta = \Lambda(I_m + \Lambda^T \Lambda)^{-1/2}$, for some $k \times m$ real matrix $\Lambda$ with finite entries. In such case, the matrix identities $I_m - \Delta^T \Delta = (I_m + \Lambda^T \Lambda)^{-1}$ and $I_k - \Delta \Delta^T = (I_k + \Lambda \Lambda^T)^{-1}$ imply that (2.1) can be rewritten as

$$f_{Z^*}(z) = 2^m \phi_k(z) \Phi_m((I_m + \Lambda^T \Lambda)^{-1/2} \Lambda^T z | (I_m + \Lambda^T \Lambda)^{-1}), \quad z \in \mathbb{R}^k.$$

Hence, an alternative notation for referring to the $FSN$ distribution defined in (2.1) is $FSN_{k,m}(\Lambda)$. 

5
The following lemma will be used frequently.

**Lemma 2.1** Let \( Z \sim N_k(0, I_k) \). Then,

\[
E[\phi_m(u + AZ|\Omega)] = \phi_m(u \Omega + AA^T),
\]

and

\[
E[\Phi_m(a + AZ|\Omega)] = \Phi_m(a|\Omega + AA^T).
\]

**Proof:** In fact, \( E[\phi_m(u + AZ|\Omega)] = \int_{\mathbb{R}^k} \phi_m(u - Az, \Omega) \phi_k(z) \, dz = \int_{\mathbb{R}^k} f_U Z = \Phi_k(u) \), where \( U|Z = z \sim N_m(-Az, \Omega) \) and \( Z \sim N_k(0, I_k) \) implying \( U \sim N_m(0, \Omega + AA^T) \).

\( \square \)

Note that Lemma 2.1 guarantees in particular that (2.1) is just a density function on \( \mathbb{R}^k \), since \( \int_{\mathbb{R}^k} \phi_k(z) \Phi_m(\Delta^T z I_m - \Delta^T \Delta) \, dz = E[\Phi_m(\Delta^T Z|I_m - \Delta^T \Delta)] = \Phi_m(0) = 2^{-m}. \)

As it is shown in the next result, the cdf of the FSN distribution has a simple form.

**Proposition 2.1** If \( Z^* \sim FSN_{k,m}(\Delta) \), then its cdf is given by

\[
F_{Z^*}(z) = 2^m \Phi_{k+m}(z, 0|\Omega), \quad z \in \mathbb{R}^k, \quad \text{where} \quad \Omega = \begin{pmatrix} I_k & -\Delta \\ -\Delta^T & I_m \end{pmatrix}.
\]

**Proof:** By (2.1),

\[
F_{Z^*}(z) = 2^m \int_{u \leq 0} \int_{v \leq 0} \phi_k(u + z) \phi_m(v + \Delta^T (u + z)|I_m - \Delta^T \Delta) \, dv \, du = 2^m P(U \leq 0, V \leq 0),
\]

where \( V|U = u \sim N_m(-\Delta^T (u + z), I_m - \Delta^T \Delta) \) and \( U \sim N_k(-z, I_k) \). Thus, the proof follows from the fact that

\[
\begin{pmatrix} U \\ V \end{pmatrix} \sim N_{k+m} \left( \begin{pmatrix} -z \\ 0 \end{pmatrix}, \begin{pmatrix} I_k & -\Delta \\ -\Delta^T & I_m \end{pmatrix} \right).
\]

\( \square \)

### 2.1 Special cases

Some important special cases of the FSN family defined by (2.1) are the following:

1. If \( m = 1 \), with \( \Delta = (\delta_1, \ldots, \delta_k)^T \), then (2.1) reduces to the Azzalini and Dalla Valle (1996) SN density in (1.3), with

\[
\lambda = \begin{pmatrix} \delta_1 / \sqrt{1 - \sum_{i=1}^k \delta_i^2} \\ \vdots \\ \delta_k / \sqrt{1 - \sum_{i=1}^k \delta_i^2} \end{pmatrix}^T;
\]
2. If \( m = k \) and \( \Delta = \text{Diag}(\delta_1, \ldots, \delta_k) \), then (2.1) reduces to the product of \( k \) univariate SN marginals, that is,

\[
f_{\mathbf{z}^*}(\mathbf{z}) = \prod_{i=1}^{k} 2\phi_1(z_i)\Phi_1(\lambda_i z_i), \quad \text{with} \quad \lambda_i = \frac{\delta_i}{\sqrt{1 - \delta_i^2}}.
\]

Thus, for any univariate SN random sample \( Z_i \sim SN(\lambda), \ i = 1, \ldots, n \), we have that \( \mathbf{Z}^* = (Z_1^*, \ldots, Z_n^*)^T \sim FSN_n(\delta \mathbf{I}_n) \), with \( \delta = \lambda/\sqrt{1 + \lambda^2} \);

3. If the matrix \( \Delta^T \Delta \) is diagonal, i.e., \( \Delta^T \Delta = \sum_{i,j} \delta_{ij} \mathbf{I} \), then (2.1) reduces to

\[
f_{\mathbf{z}^*}(\mathbf{z}) = 2^m \phi_k(\mathbf{z})\Phi_m(\Lambda^T \mathbf{z}), \quad \text{with} \quad \Lambda = \Delta(I_m - \Delta^T \Delta)^{-1/2}.
\]

Thus, by letting \( D = \Lambda^T = (I_m - \Delta^T \Delta)^{-1/2} \Delta \) in (1.5), it follows that a necessary and sufficient condition to obtain (2.1) as special case of (1.5) is that \( DD^T = \Delta^T \Delta \) be a diagonal matrix. Since this condition is not required by Definition 2.1, then we have in general that (1.5) and (2.1) define different families of SN distributions.

### 2.2 Stochastic representations

The following proposition presents conditional and marginal stochastic representations for the FSN random vector \( \mathbf{Z}^* \) introduced in Definition 2.1.

**Proposition 2.2** Let \( \mathbf{Z}^* \sim FSN_{k,m}(\Delta) \), where \( \|\Delta^T \mathbf{b}\| < 1 \), for any unitary vector \( \mathbf{b} \in \mathbb{R}^k \). Let also \( \mathbf{Z} = \Delta \mathbf{X} + (I_k - \Delta \Delta^T)^{1/2} \mathbf{Y} \), where \( \mathbf{X} \sim N_m(0, \mathbf{I}_m) \) and \( \mathbf{Y} \sim N_k(0, \mathbf{I}_k) \), which are independent. Then,

\[
\mathbf{Z}^* \overset{d}{=} [\mathbf{Z} | \mathbf{X} > 0],
\]

which is called conditional representation. Moreover,

\[
\mathbf{Z}^* \overset{d}{=} \Delta [\mathbf{X}] + (I_k - \Delta \Delta^T)^{1/2} \mathbf{Y}, \quad \text{where} \quad \mathbf{X} = (|X_1|, \ldots, |X_m|)^T,
\]

which is called marginal representation.

**Proof:** Theorem 5.1 in Arellano-Valle et al. (2002) establishes that the conditional random vector \( [\mathbf{Z} | \mathbf{X} > 0] \) has a density as in (1.6), where \( f_k(\mathbf{z}) \) is the marginal density of \( \mathbf{Z} \) and, as is indicated in (1.7), \( K_m = P(\mathbf{X} > 0) \) and \( Q_m(z) = P(\mathbf{X} > 0 \mathbf{Z} = \mathbf{z}) \). Here, by the assumptions it is clear that \( \mathbf{Z} \sim N_k(0, \mathbf{I}_k) \), so that the joint distribution of \( \mathbf{X} \) and \( \mathbf{Z} \) is

\[
\begin{pmatrix} \mathbf{X} \\ \mathbf{Z} \end{pmatrix} \sim N_{m+k} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \mathbf{I}_m & \Delta^T \\ \Delta & \mathbf{I}_k \end{pmatrix} \right)
\]

and so \( [\mathbf{X} | \mathbf{Z} = \mathbf{z}] \sim N_m(\Delta^T \mathbf{z}, \mathbf{I}_m - \Delta \Delta^T) \). Thus, \( K_m = 2^{-m}, f_k(\mathbf{z}) = \phi_k(\mathbf{z}) \) and \( Q_m(\mathbf{z}) = \Phi_m(\Delta^T \mathbf{z}| \mathbf{I}_m - \Delta \Delta^T) \), from where the conditional representation follows. The marginal representation comes from its equivalence with the conditional representation, which is established in Theorem 3.1 of Arellano-Valle et al. (2002). \( \square \)
Remark 2.2 Since \( \Delta = \Lambda(I_m + \Lambda^T \Lambda)^{-1/2} \), for some real \( k \times m \) matrix \( \Lambda \), it follows from the matrix identity \((I_k + \Delta \Lambda^T)^{-1} = I_k - \Lambda (I_m + \Lambda^T \Lambda)^{-1/2} \Lambda^T = I_k - \Delta \Delta^T \) that \( \mathbf{Z} = \Lambda (I_m + \Lambda^T \Lambda)^{-1/2} \mathbf{X} + (I_k + \Delta \Lambda^T)^{-1/2} \mathbf{Y} \). Note also that \( \text{Cov}(\mathbf{Z}, \mathbf{X}) = \Delta \). Thus, since \( \text{V}(\mathbf{Z}) = I_k \) and \( \text{V}(\mathbf{X}) = I_m \), we have for any given \( k \times m \) real matrix \( \Lambda \) that \( \Delta = \Lambda (I_m + \Lambda^T \Lambda)^{-1/2} \) is a correlation matrix.

Corollary 2.1 If \( \mathbf{Z}^* \sim \text{FSN}_{k,m}(\Delta) \), then

\[
E(\mathbf{Z}^*) = \sqrt{\frac{2}{\pi}} \Delta \mathbf{1}_m \quad \text{and} \quad V(\mathbf{Z}^*) = I_k - \frac{2}{\pi} \Delta \Delta^T,
\]

where \( \mathbf{1}_n \) is a \( n \times 1 \) vector of ones.

Proof: Using that \( E[\mathbf{X}] = \sqrt{\frac{2}{\pi}} \mathbf{1}_m \) and \( V[\mathbf{X}] = (1 - 2/\pi)I_m \), where \( \mathbf{X} \sim \text{N}_m(0, I_m) \), the proof is direct from the marginal stochastic representation.

\[ \square \]

Note from Corollary 2.1 that if \( \mathbf{Z}^* = (Z_1^*, \ldots, Z_k^*)^T \sim \text{FSN}_{k,m}(\Delta) \), then

\[
\text{Cov}(Z_i^*, Z_j^*) = -\frac{2}{\pi} \Delta_i^T \Delta_j = -\frac{2}{\pi} \sum_{p=1}^{m} \delta_{ip} \delta_{jp}, \quad i \neq j,
\]

where \( \Delta_i^T = (\delta_{i1}, \ldots, \delta_{im}) \), \( i = 1, \ldots, k \), are the rows of \( \Delta \). Similarly, if \( \mathbf{Z}^* \) and \( \Delta \) are partitioned in the form of

\[
\mathbf{Z}^* = \begin{pmatrix} \mathbf{Z}_1^* \\ \mathbf{Z}_2^* \end{pmatrix} \quad \text{and} \quad \Delta = \begin{pmatrix} \Delta_1 \\ \Delta_2 \end{pmatrix},
\]  

(2.2)

where \( \mathbf{Z}_i^* \) and \( \Delta_i \) have dimensions \( k_i \times 1 \) and \( k_i \times m \), \( i = 1, 2 \), respectively, and \( k_1 + k_2 = k \), then

\[
\text{Cov}(\mathbf{Z}_1^*, \mathbf{Z}_2^*) = -\frac{2}{\pi} \Delta_1 \Delta_2^T.
\]

2.3 Moment generating function

In the next result, we derive the moment generating function (mgf) of the FSN distribution. Additional properties of this distribution are obtained from its mgf.

Proposition 2.3 If \( \mathbf{Z}^* \sim \text{FSN}_{k,m}(\Delta) \), then its mgf is given by

\[
M_{\mathbf{Z}^*}(\mathbf{t}) = 2^m e^{(1/2)\mathbf{t}^T \Delta^T \mathbf{t}} \Phi_m(\Delta^T \mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^k.
\]

Proof: Note first that \( e^{\mathbf{t}^T \phi_k(\mathbf{z})} = e^{(1/2)\mathbf{t}^T \mathbf{z}} \phi_k(\mathbf{z} - \mathbf{t}) \). Hence, using the variable change \( \mathbf{y} = \mathbf{z} - \mathbf{t} \) after applying (2.1), we have that

\[
M_{\mathbf{Z}^*}(\mathbf{t}) = E[e^{\mathbf{t}^T \mathbf{Z}^*}] = 2^m e^{(1/2)\mathbf{t}^T \mathbf{t}} \int_{u \leq \Delta^T \mathbf{t}} E[\phi_m(u + \Delta^T \mathbf{Y} I_m - \Delta^T \Delta)]d\mathbf{u}.
\]
where \( Y \sim N_k(0, I_k) \). Thus, the proof follows from Lemma 2.1.

An important byproduct of Proposition 2.3 is the following additive property of the FSN distribution.

**Corollary 2.2** Let \( Z_i^* \sim FSN_{k, m_i}(\Delta_i) \), \( i = 1, \ldots, n \), which are \( k \times 1 \) independent random vectors. Then,

\[
Z^* = \frac{1}{n} \sum_{i=1}^{n} Z_i^* \sim FSN_{k, m}(\Delta), \quad \text{where} \quad m = \sum_{i=1}^{n} m_i \quad \text{and} \quad \Delta = \frac{1}{n} \sum_{i=1}^{n} \Delta_i.
\]

**Proof:** Since \( \prod_{i=1}^{n} \phi_{m_i}((1/n)\Delta_i^T t) = \phi_m(\Delta^T t) \), where \( m = \sum_{i=1}^{n} m_i \) and \( \Delta = (1/n) \sum_{i=1}^{n} \Delta_i \), the proof follows from Proposition 2.3.

Now, we characterize the distribution of an arbitrary linear transformation \( AZ^* + b \) by means of its mgf and also by means of its pdf when \( A \) is a nonsingular matrix. The proof of such results is direct from Proposition 2.3 by noting that \( M_{AZ^*+b}(t) = e^{t^T b} M_{Z^*}(A^T t) \) and from (2.1) by using that \( f_{AZ^*+b}(y) = |\det(A)|^{-1} f_{Z^*}(A^{-1}(y-b)) \).

**Proposition 2.4** If \( Z^* \sim FSN_{k, m}(\Delta) \), then for any \( n \times k \) matrix \( A \),

\[
M_{AZ^*+b}(t) = 2^m e^{t^T b + (1/2)t^T A A^T t} \phi_m(\Delta^T A^T t), \quad t \in \mathbb{R}^k.
\]

Moreover, if for \( n = k \) the matrix \( A \) is nonsingular, then

\[
f_{AZ^*+b}(y) = 2^m |\det(A)|^{-1} \phi_k(A^{-1}(y-b)) \phi_m(\Delta^T A^{-1}(y-b) I_m - \Delta^T \Delta), \quad y \in \mathbb{R}^k.
\]

By using (2.3) with \( b = 0 \), we have the following additional properties of the FSN distribution.

**Corollary 2.3** Let \( Z^* \sim FSN_{k, m}(\Delta) \). Then:
(i) \( -Z^* \sim FSN_{k, m}(-\Delta) \);
(ii) \( a^T Z^* \sim FSN_{1, m}(a^T \Delta) \), for any unitary vector \( a \in \mathbb{R}^k \);
(iii) \( AZ^* \sim FSN_{k, m}(A \Delta) \), for any \( k \times k \) orthogonal matrix \( A \).

To end this section, we characterize next the distribution of quadratic forms of an FSN random vector by means of their mgf's.

**Proposition 2.5** If \( Z^* \sim FSN_{k, m}(\Delta) \) and \( A \) is a given \( k \times k \) symmetric matrix, then the mgf of \( \Psi = Z^T A Z^* \) is

\[
M_{\Psi}(t) = 2^m |I_k - 2tA|^{-1/2} \phi_m(0 | I_m + 2t \Delta^T (I_k - 2tA)^{-1/2} A (I_k - 2tA)^{-1/2} \Delta), \quad t \in \mathbb{R},
\]

which for \( A = I_k \) is reduced to

\[
M_{\Psi}(t) = 2^m (1 - 2t)^{-k/2} \phi_m(0 | I_m + 2t(1 - 2t)^{-1} \Delta^T \Delta), \quad t \in \mathbb{R}.
\]
Proof: Note first that $e^{t^T z^* A z} \phi_k(z) = |I_k - 2tA^{-1/2} \phi_k(z(I_k - 2tA)^{-1})$. Now, consider the random vector $Z \sim N_k(0, (I_k - 2tA)^{-1})$. The proof follows by using (2.1) and the fact that by Lemma 2.1, $E[e^{t^T \Psi}] = 2^m E[\Phi_m(\Delta^T Z|I_m - \Delta^T \Delta)] = \Phi_m(0)|I_m - \Delta^T \Delta + \Delta^T (I_k - 2tA)^{-1} \Delta|.

From Proposition 2.5, we can obtain the moments of the quadratic form $\Psi = Z^T A Z^*$ by deriving its mgf, following Genton et al. (2001). For instance, we note that the Corollary 2.1 implies that
\[ E(\Psi) = \text{tr}(A V(Z^*)) + E(Z^T A E(Z^*) = \text{tr}(A) + \frac{2}{\pi} [1_m^T \Delta^T A \Delta 1_m - \text{tr}(\Delta^T A \Delta)]. \]
Alternatively, since $A$ is a $k \times k$ symmetric matrix with eigenvalues $\lambda_i$, $i = 1, \ldots, k$, then we can use that $A = P^T A P$, where $A = \text{Diag}(\lambda_1, \ldots, \lambda_k)$ and $P$ is a $k \times k$ orthogonal matrix, so that
\[ \Psi = Z^T P^T \Lambda P Z^* \overset{d}{=} Y^T \Lambda Y^* = \sum_{i=1}^k \lambda_i Y_i^2, \]
where, from part (iii) of Corollary 2.3, $Y^* = P Z^* \sim FSN_{k,n}(P \Delta)$, and from part (ii) of the same corollary, $Y_i^* = P_i^T Z^* \sim FSN_{1,n}(P_i \Delta)$, $i = 1, \ldots, k$, where $P_1, \ldots, P_k$ are the columns of $P$ (i.e., $P_i$ is the eigenvector of $A$ corresponding to the eigenvalue $\lambda_i$, $i = 1, \ldots, k$). Thus, the results given in Section 4 by Arellano-Valle et al. (2002) can be used to obtain the moments of $\Psi$.

Another important byproduct of Proposition 2.5 is the following characterization for the distribution of the squared length of an $FSN$ random vector.

Corollary 2.4 If $Z^* \sim FSN_{k,m}(\Delta)$, with $\Delta \neq O$, then $|Z^*|^2 \sim \chi_k^2$ if and only if the matrix $\Delta^T \Delta$ is diagonal.

2.4 Marginal distributions and independence

Denote by $(Z_i^*, i = 1, 2)$ and by $(\Delta_i, i = 1, 2)$ the partition of the random vector $Z^*$ and the induced row-partition of the matrix $\Delta$ described in (2.2), respectively. Considering this notation, the next result shows that the $FSN$ distribution is closed under marginalization.

Proposition 2.6 Assume that $Z^* = (Z_i^*, i = 1, 2) \sim FSN_{k,m}(\Delta)$. Then, $Z_i^* \sim FSN_{k_i,m}(\Delta_i)$, i.e.,
\[ f_{Z_i^*}(z_i) = 2^m \phi_{k_i}(z_i)|\Phi_m(\Delta_i^T z_i|I_m - \Delta_i^T \Delta_i)|, z_i \in \mathbb{R}^{k_i}, i = 1, 2. \]

Proof: It is direct from (2.3) by taking $A = [I_{k_1}, O]$ for $i = 1$ and $A = [0, I_{k_2}]$ for $i = 2$. Alternatively, since $I_m - \Delta^T \Delta = I_m - \Delta_1^T \Delta_1 - \Delta_2^T \Delta_2$, we have from (2.1) that $f_{Z_i^*}(z_1) = 2^m \phi_{k_1}(z_1) \int_{u \leq \Delta_1^T z_1} f_{Z_2^*}(z_2) \phi_m(u + \Delta_2^T z_2|I_m - \Delta^T \Delta)dz_2du$, which yields (2.5) with $i = 1$ by using the Lemma 2.1. In a similar way, we obtain (2.5) for $i = 2$. □
Remark 2.3 If $\Delta = \Lambda (I_m + \Lambda \Lambda)^{-1}$, then $\Delta_i = \Lambda_i (I_m + \Lambda^T \Lambda)^{-1}$, where $(\Lambda_i, i = 1, 2)$ is the corresponding partition on the matrix $\Lambda$.

Suppose now that $m > 1$ and let us partition the matrices $\Delta_i$, $i = 1, 2$, indicated in (2.2) as $\Delta_i = (\Delta_{ij}, i, j = 1, 2)$, where $\Delta_{ij}$ has dimension $k_i \times m_j$, $j = 1, 2$, and $m_1 + m_2 = m$.

Proposition 2.7 Assume that $Z^* = (Z^*_i, i = 1, 2) \sim FSN_{k,m}(\Delta)$, with $m > 1$ and $\Delta \neq 0$. Then, under each of the following conditions on the shape matrix $\Delta$, the random vectors $Z^*_i$ and $Z^*_2$ are independent:

(i) $\Delta_{12} = 0$ and $\Delta_{21} = 0$. In this case, $Z^*_i \sim FSN_{k,m}(\Delta_{ii})$, $i = 1, 2$; or

(ii) $\Delta_{ii} = 0$, $i = 1, 2$. In this case, $Z^*_1 \sim FSN_{k_1,m_1}(\Delta_{11})$ and $Z^*_2 \sim FSN_{k_2,m_2}(\Delta_{22})$.

Proof: The proof follows by noting that (2.1) implies

$$\int Z^*_1 Z^*_2(dz_1,dz_2) = 2^{m_1 + m_2} \phi_{k_1}(z_1) \phi_{k_2}(z_2) \Phi_{m_1 + m_2}(\Delta^T z_1 + \Delta^T z_2, I_{m_1 + m_2} - \Delta^T \Delta_1 - \Delta^T \Delta_2),$$

where from the partition $\Delta_i = (\Delta_{ij}, i, j = 1, 2)$, $i = 1, 2$,

$$\Delta^T z_1 + \Delta^T z_2 = \begin{pmatrix} \Delta^T_{11} z_1 + \Delta^T_{12} z_2 \\ \Delta^T_{21} z_1 + \Delta^T_{22} z_2 \end{pmatrix}$$

and

$$I_{m_1 + m_2} - \Delta^T \Delta_1 - \Delta^T \Delta_2 = \begin{pmatrix} I_{m_1} - (\Delta^T_{11} \Delta_{11} + \Delta^T_{12} \Delta_{21}) & -(\Delta^T_{11} \Delta_{12} + \Delta^T_{12} \Delta_{22}) \\ -(\Delta^T_{21} \Delta_{11} + \Delta^T_{22} \Delta_{21}) & I_{m_2} - (\Delta^T_{21} \Delta_{12} + \Delta^T_{22} \Delta_{22}) \end{pmatrix}.$$ 

Thus, the last term in (2.6) factorizes into (i) $\Phi_{m_1}(\Delta^T_{11} z_1, I_{m_1} - \Delta^T_{11} \Delta_{11}) \Phi_{m_2}(\Delta^T_{12} z_2, I_{m_2} - \Delta^T_{22} \Delta_{22})$, if $\Delta_{12} = 0$ and $\Delta_{21} = 0$, and into (ii) $\Phi_{m_1}(\Delta^T_{21} z_1, I_{m_1} - \Delta^T_{12} \Delta_{12}) \Phi_{m_2}(\Delta^T_{22} z_2, I_{m_2} - \Delta^T_{22} \Delta_{22})$, if $\Delta_{11} = 0$ and $\Delta_{22} = 0$, thus concluding the proof. \hfill $\Box$

Now, denote by $(\delta_{ij}; i = 1, \ldots, k, j = 1, \ldots, m)$ and by $\Delta^T_{ii} = (\delta_{ij}, \ldots, \delta_{i m_k})$, $i = 1, \ldots, k$, the entries and the rows of the shape matrix $\Delta$, respectively. Let also $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ be the $k \times 1$ vector with 1 at the $i$-th position, $i = 1, \ldots, m$.

Proposition 2.8 Let $Z^* = (Z^*_1, \ldots, Z^*_k)^T \sim FSN_{k,m}(\Delta)$. Then:

(i) $Z^*_i \sim FSN_{1,m}(\Delta^T_{ii})$, $i = 1, \ldots, k$;

(ii) $Z^*_1, \ldots, Z^*_k$ are independent if and only if $m = k$ and $\Delta_{i j} = \delta_{ij} e_j$, for some $j = 1, \ldots, k$, with $j_i \neq j_v$ for all $i \neq v$, $i = 1, \ldots, k$. Moreover, in such case $Z^*_i \sim FSN_1(\delta_{i j})$, $i = 1, \ldots, k$.

Proof: Note first that $e_i^T Z^* = Z^*_i$ and $e_i^T \Delta = \Delta^T_{ii}$, where $\|e_i\| = 1$, $i = 1, \ldots, k$. Thus, by the part (ii) of Corollary 2.3, $M_{Z^*_i}(t_i) = 2^m e^{(1/2) \ell_i^T \phi_m(\Delta_{i i} t_i)}$, $i = 1, \ldots, k$, which proves the part (i). To prove (ii), note first that

$$\prod_{i = 1}^k M_{Z^*_i}(t_i) = 2^{km} e^{(1/2) \sum_{i = 1}^k \ell_i^T \phi_m(\Delta_{i i} t_i)}.$$ 

Now, by assumption and Proposition 2.3, $M_{Z^*_1, \ldots, Z^*_k}(t_1, \ldots, t_k) = 2^m e^{(1/2) \sum_{i = 1}^k \ell_i^T \phi_m(\sum_{i = 1}^k \Delta_{i i} t_i)}$. 

\hfill 11
Thus, we have that $M_{Z_1^* \ldots Z_k^*}(t_1, \ldots, t_k) = \prod_{i=1}^k M_{Z_i^*}(t_i)$, for all $(t_1, \ldots, t_k)$, if and only if $\Phi_m(\sum_{i=1}^k \Delta_i t_i) = 2^k \prod_{i=1}^k \Phi_m(\Delta_i t_i)$, for all $(t_1, \ldots, t_k)$, which holds if and only if $m = k$ and $\Delta_i = \delta_{ji} e_{ji}$, for some $j_i, 1, \ldots, k$, with $ji \neq j_i'$ for all $i \neq i'$, $i = 1, \ldots, k$. □

Note that for $m = 1$, it is not possible to obtain independence for the components of an FSN random vector $Z^*$.

### 2.5 Conditional distributions

As is shown in the following result, the FSN distributions are not closed under conditioning. However, their conditional distributions have the form of the more general family defined by (1.6).

**Proposition 2.9** If $Z^* = (Z_i^*, i = 1, 2) \sim FSN_{k,m}(\Delta)$, then the conditional density of $Z_1^*$ given $Z_2^* = z_2$ is

$$f_{Z_1^*|Z_2^* = z_2}(z_1) = (K_m(z_2))^{-1} \phi_{k_1}(z_1) \Phi_m(\Delta_1^T z_1| \Delta_1^T z_2, I_m - \Delta^T \Delta),$$

where $K_m(z_2) = \Phi_m(\Delta_2^T z_2| I_m - \Delta_2^T \Delta_2)$.

**Proof:** Since $\Phi_m(\Delta^T z| I_m - \Delta^T \Delta) = \Phi_m(\Delta_1^T z_1 + \Delta_2^T z_2| I_m - \Delta^T \Delta|), \Phi_m(\Delta_1^{T}\tilde{z}_1| I_m - \Delta_1^T \tilde{z}_2, I_m - \Delta^T \Delta)$, the proof follows from the fact that (2.5) and (2.6) imply

$$f_{Z_1^*|Z_2^* = z_2}(z_1) = \frac{\phi_{k_1}(z_1) \Phi_m(\Delta_1^T \tilde{z}_1| I_m - \Delta_1^T \tilde{z}_2, I_m - \Delta^T \Delta)}{\Phi_m(\Delta_2^T \tilde{z}_2| I_m - \Delta_2^T \Delta_2)}.$$

In order to obtain the conditional moments of $Z_1^*$ given $Z_2^* = z_2$, we need the following preliminary result. Let $Z_1^* \sim N_{k_1}(0, I_{k_1})$ and let $g$ be a real function such that $E[|g(Z_1^*)|] < \infty$. Let also $X$ be a $m \times 1$ random vector with density $f(x)$ and let

$$f_A(x|x_0) = \frac{f(x)}{P(X \leq x_0)} I(x \leq x_0) \quad (2.7)$$

be the conditional density of $X$ given $\{X \leq x_0\}$ for some fixed vector $x_0 \in \mathbb{R}^m$, where $I_A$ is the indicator of $A$.

**Proposition 2.10** Let $Z^* = (Z_i^*, i = 1, 2) \sim FSN_{k,m}(\Delta)$. Then, for any integrable real function $g$, it follows that

$$E[g(Z_1^*)|Z_2^* = z_2] = E[h(X)|X \leq \Delta_2^T z_2],$$

where $[Z_1^* X = x] \sim N_m(-\Delta_1(I_m - \Delta_2^T \Delta_2)^{-1} x, I_m - \Delta_1(I_m - \Delta_2^T \Delta_2)^{-1} \Delta_1^T)$ and $X \sim N_m(0, I_m - \Delta_2^T \Delta_2)$.\[12\]
\textbf{Proof:} Using that \( g \) is an integrable real function and considering the fact that \( \Phi_m(\Delta_T^T z_1 - \Delta_T^T z_2, I_m - \Delta_T^T \Delta) = \Phi_m(\Delta_T^T z_2 - \Delta_T^T z_1, I_m - \Delta_T^T \Delta) \), we have from Proposition 2.9 that
\[
E[g(Z_1^*)|Z_2^* = z_2] = \frac{1}{\Phi_m(\Delta_T^T z_2|I_m - \Delta_T^T \Delta)} \int_{\mathbb{R}^k} g(z_1) \phi_k(z_1) \Phi_m(\Delta_T^T z_2|I_m - \Delta_T^T \Delta) dz_1
\]
\[
= \frac{1}{\Phi_m(\Delta_T^T z_2|I_m - \Delta_T^T \Delta)} \int_{\mathbb{R}^k} g(z_1) \phi_k(z_1) \int_{x \leq \Delta_T^T z_2} \phi_m(x|\Delta_T^T z_1, I_m - \Delta_T^T \Delta) dx dz_1
\]
\[
= \frac{1}{\Phi_m(\Delta_T^T z_2|I_m - \Delta_T^T \Delta)} \int_{x \leq \Delta_T^T z_2} \int_{\mathbb{R}^k} g(z_1) \phi_k(z_1) \phi_m(x|\Delta_T^T z_1, I_m - \Delta_T^T \Delta) dz_1 dx
\]
where \( f_{Z_1}(z_1) = \phi_k(z_1) \) and \( f_X|Z_1 = z_1(x) = \phi_m(x|\Delta_T^T z_1, I_m - \Delta_T^T \Delta) \), i.e., \( Z_1 \sim N_k(0, I_k) \) and \( [X|Z_1 = z_1] \sim N_m(\Delta_T^T z_1, I_m - \Delta_T^T \Delta) \). Using that \( f_{Z_1}(z_1) f_X|Z_1 = z_1(x) = f_X(x) f_{Z_1|x=x}(z_1) \), where \( X \sim N_m(0, I_m - \Delta_T^T \Delta) \) and so \( P(X \leq \Delta_T^T z_2) = \Phi_m(\Delta_T^T z_2|I_m - \Delta_T^T \Delta) \), and that \( [Z_1|X = x] \sim N_k(\Delta_T^T (I_m - \Delta_T^T \Delta)^{-1} x, \Delta_T^T (I_m - \Delta_T^T \Delta)^{-1} \Delta_T^T) \), we have that
\[
E[g(Z_1^*)|Z_2^* = z_2] = \frac{1}{\Phi_m(\Delta_T^T z_2|I_m - \Delta_T^T \Delta)} \int_{x \leq \Delta_T^T z_2} f_X(x) \int_{\mathbb{R}^k} g(z_1) f_{Z_1|x=x}(z_1) dz_1 dx
\]
\[
= \int_{\mathbb{R}^m} \frac{f_X(x)}{P(X \leq \Delta_T^T z_2)} I_{x \leq \Delta_T^T z_2} \int_{\mathbb{R}^k} g(z_1) f_{Z_1|x=x}(z_1) dz_1 dx
\]
\[
= \int_{\mathbb{R}^m} f_S(x|\Delta_T^T z_2) E[g(Z_1)|X = x] dx,
\]
where, as was defined in (2.7), \( f_S(x|\Delta_T^T z_2) \) is the conditional density of \( X \) given \( \{X \leq \Delta_T^T z_2\} \), which concludes the proof.

\[
\square
\]

\textbf{Corollary 2.5} If \( Z^* = (Z_i^*, i = 1, 2) \sim FS N_{k,m}(\Delta) \), then
\[
E[Z_1^*|Z_2^* = z_2] = -\Delta_1 (I_m - \Delta_2^T \Delta_2)^{-1} E[X|X \leq \Delta_T^T z_2] \quad (2.8)
\]
and
\[
V[Z_1^*|Z_2^* = z_2] = \Delta_{k_1} (I_m - \Delta_2^T \Delta_2)^{1} \Delta_T^T + \Delta_1 (I_m - \Delta_2^T \Delta_2)^{1} V[X|X \leq \Delta_T^T z_2] (I_m - \Delta_2^T \Delta_2)^{1} \Delta_T^T, \quad (2.9)
\]
where \( X \sim N_m(0, I_m - \Delta_2^T \Delta_2) \).
Proof: Let $g(Z_1^*) = a^T Z_1^*$, where $a \in \mathbb{R}^{k_1}$. Then, from Proposition 2.10, it follows that

$$E[a^T Z_1^* | Z_2^* = z_2] = \int_{\mathbb{R}^{k_1}} f_x(\mathbf{x} | z_2) E[a^T Z_1 | \mathbf{X} = \mathbf{x}] d\mathbf{x},$$

where $E[a^T Z_1 | \mathbf{X} = \mathbf{x}] = a^T E[Z_1 | \mathbf{X} = \mathbf{x}] = -a^T \Delta_1^T (I_m - \Delta_2 \Delta_2)^{-1} \mathbf{x}$, and the proof of (2.8) follows by noting that $\mathbf{a}$ is arbitrary. In a similar way, by taking $g(Z_1^*) = (a^T Z_1^*)^2$, we have that

$$E[(a^T Z_1^*)^2 | Z_2^* = z_2] = \int_{\mathbb{R}^{k_1}} f_x(\mathbf{x} | z_2) E[(a^T Z_1)^2 | \mathbf{X} = \mathbf{x}] d\mathbf{x},$$

where

$$E[(a^T Z_1)^2 | \mathbf{X} = \mathbf{x}] = a^T E[Z_1 Z_1^T | \mathbf{X} = \mathbf{x}] \mathbf{a} = a^T \{V[Z_1 | \mathbf{X} = \mathbf{x}] + E[Z_1 | \mathbf{X} = \mathbf{x}] E[Z_1^T | \mathbf{X} = \mathbf{x}]\} \mathbf{a}$$

$$= a^T \{I_{k_1} - \Delta_1 (I_m - \Delta_2 \Delta_2)^{-1} \Delta_1^T + \Delta_1 (I_m - \Delta_2 \Delta_2)^{-1} \mathbf{x} \mathbf{a} \} \mathbf{a},$$

and so

$$E[(a^T Z_1^*)^2 | Z_2^* = z_2] = a^T \{I_{k_1} - \Delta_1 (I_m - \Delta_2 \Delta_2)^{-1} \Delta_1^T + \Delta_1 (I_m - \Delta_2 \Delta_2)^{-1} \mathbf{x} \mathbf{a} \} \mathbf{a},$$

and the proof of (2.9) follows by noting that $E[\mathbf{X} \mathbf{X}^T | \mathbf{X} \leq \Delta_2^T \mathbf{z}_2] = V[\mathbf{X} \mathbf{X}^T | \mathbf{X} \leq \Delta_2^T \mathbf{z}_2] + E[\mathbf{X} | \mathbf{X} \leq \Delta_2^T \mathbf{z}_2] E[\mathbf{X}^T | \mathbf{X} \leq \Delta_2^T \mathbf{z}_2] = V[\mathbf{X} \mathbf{X}^T | \mathbf{X} \leq \Delta_2^T \mathbf{z}_2] + E[\mathbf{Z}_1 | \mathbf{Z}_2^* = \mathbf{z}_2] E[\mathbf{Z}_1^T | \mathbf{Z}_2^* = \mathbf{z}_2].$

Note that in general, in (2.8) and (2.9), it is not easy to obtain closed-form expressions for $E[\mathbf{X} | \mathbf{X} \leq \mathbf{x}_0]$ and $V[\mathbf{X} | \mathbf{X} \leq \mathbf{x}_0]$ for any given $\mathbf{x}_0$. However, as we show next, for the particular case where $\Delta_2 \Delta_2 = \text{Diag}(\delta_i^2, \ldots, \delta_m^2)$ we have closed-form solutions for these conditional moments. In fact, since $\mathbf{X} = (X_1, \ldots, X_m)^T \sim N_m(0, I_m - \Delta_2 \Delta_2)$, then $X_i \sim N(0, 1 - \delta_i^2)$, $i = 1, \ldots, m$, and are independent, so that

$$E[\mathbf{X} | \mathbf{X} \leq \mathbf{x}_0] = (E[X_1 | X_1 \leq x_{01}], \ldots, E[X_m | X_m \leq x_{0n}])^T$$

and

$$V[\mathbf{X} | \mathbf{X} \leq \mathbf{x}_0] = \text{Diag}(V[X_1 | X_1 \leq x_{01}], \ldots, V[X_m | X_m \leq x_{0n}]).$$

Now, we may use the following lemma (Johnson et al., 1994).

**Lemma 2.2** If $\mathbf{X} \sim N(\mu, \sigma^2)$, then for any given $\mathbf{a}$, it follows that

$$E[\mathbf{X} | \mathbf{X} \leq \mathbf{a}] = \mu - \frac{\phi(\frac{\mathbf{a} - \mu}{\sigma})}{\Phi(\frac{\mathbf{a} - \mu}{\sigma})}$$

and

$$V[\mathbf{X} | \mathbf{X} \leq \mathbf{a}] = \left\{1 - \frac{\phi(\frac{\mathbf{a} - \mu}{\sigma})}{\Phi(\frac{\mathbf{a} - \mu}{\sigma})} \right\} \sigma^2.$$
In particular, if \( m = 1 \), then \( X_1 \sim N(0, 1 - \delta_1^2) \), so that

\[
E[X_1 | X_1 \leq \Delta_z^T z_2] = \frac{\phi \left( \frac{\Delta_z^T z_2}{\sqrt{1 - \delta_1^2}} \right)}{\Phi \left( \frac{\Delta_z^T z_2}{\sqrt{1 - \delta_1^2}} \right)}
\]

and

\[
V[X_1 | X_1 \leq \Delta_z^T z_2] = \left\{ 1 - \frac{\phi \left( \frac{\Delta_z^T z_2}{\sqrt{1 - \delta_1^2}} \right)}{\Phi \left( \frac{\Delta_z^T z_2}{\sqrt{1 - \delta_1^2}} \right)} \right\} \left( \frac{\Delta_2^T z_2}{\sqrt{1 - \delta_1^2}} \right)^2 \left( \frac{\phi \left( \frac{\Delta_z^T z_2}{\sqrt{1 - \delta_1^2}} \right)}{\Phi \left( \frac{\Delta_z^T z_2}{\sqrt{1 - \delta_1^2}} \right)} \right)^2 \right\} (1 - \delta_1^2).
\]

The above results simplify also when \( \Delta_z^T z_2 = 0 \), since the fact that \( U = (I_m - \Delta_2^T \Delta_2)^{-1/2} X \sim N_m(0, I_m) \), with \( \Delta_2^T \Delta_2 \) being a diagonal matrix, implies

\[
E[X | X \leq 0] = (I_m - \Delta_2^T \Delta_2)^{-1/2} E[U | U \leq 0] = -(I_m - \Delta_2^T \Delta_2)^{-1/2} E[|U|] = -\sqrt{\frac{2}{\pi}}(I_m - \Delta_2^T \Delta_2)^{-1/2} 1_m
\]

and

\[
V[X | X \leq 0] = (I_m - \Delta_2^T \Delta_2)^{-1/2} V[|U|] (I_m - \Delta_2^T \Delta_2)^{-1/2} = (1 - 2/\pi)(I_m - \Delta_2^T \Delta_2).
\]

In such case, it follows from (2.7) and (2.8) that

\[
E[Z^*_1 | Z^*_2 = 0] = \sqrt{\frac{2}{\pi}} \Delta_1 (I_m - \Delta_2^T \Delta_2)^{-1/2} 1_m
\]

and

\[
V[Z^*_1 | Z^*_2 = 0] = I_{k_1} - \Delta_1 (I_m - \Delta_2^T \Delta_2)^{-1} \Delta_1^T + (1 - 2/\pi) \Delta_1 \Delta_1^T.
\]

### 2.6 Location-scale extension

From the above results, further properties of the \( FSN \) distribution defined by (2.1) can be studied. For instance, if its corresponding location-scale extension is denoted by \( FSN_{k,m}(\mu, \Sigma; \Delta) \), where \( \mu \) is a \( k \times 1 \) vector and \( \Sigma \) is a \( k \times k \) positive definite matrix, then it can easily be introduced by considering the linear transformation \( W^* = \mu + \Sigma^{1/2} Z^* \), with \( Z^* \sim FSN_{k,m}(\Delta) \), whose density and moment generating functions are given by

\[
f_{W^*}(w) = 2^m |\Sigma|^{-1/2} \phi_k((\Sigma^{-1/2}(w - \mu))\Phi_m(\Delta^T \Sigma^{-1/2}(w - \mu)|I_m - \Delta^T \Delta), \quad w \in \mathbb{R}^k,
\]

and

\[
M_{W^*}(t) = 2^m e^{t^T \mu + (1/2)t^T \Sigma t} \phi_m(\Delta^T \Sigma^{1/2} t), \quad t \in \mathbb{R}^k,
\]
respectively (see (2.4) and (2.3)). In such case, we say \( W^* \sim FSN_{k,m}(\mu, \Sigma; \Delta) \). Thus, it follows from Corollary 2.1 that

\[
E(W^*) = \mu + \sqrt{\frac{2}{\pi}} \Delta 1_m \quad \text{and} \quad V(W^*) = \Sigma - \frac{2}{\pi} \Sigma^{1/2} \Delta \Delta^T \Sigma^{1/2}
\]

where the matrix \( \Delta \) can be constructed as \( \Delta = \Lambda (I_m + \Lambda^T \Lambda)^{-1/2} \) for some real \( k \times m \) matrix \( \Lambda \) with finite entries.

An important special case follows when \( m = k \), with \( \Sigma = \text{Diag}(\sigma_1^2, \ldots, \sigma_k^2) \) and \( \Delta = \text{Diag}(\delta_1, \ldots, \delta_k) \), since under this situation we have that \( W_i^* \sim FSN_1(\mu_i, \sigma_i^2; \delta_i), i = 1, \ldots, k \), and are independent. Note finally that if in (2.10) we take \( \Sigma = \Omega + DD^T \) and \( \Delta = (\Omega + DD^T)^{-1/2}D \), then we obtain (1.9), which generalizes the Sahu et al. (2003) \( SN \) distribution.

3 The fundamental generalized skew distribution

In this section, we consider the most general class of skew distributions defined in terms of its density by (1.6)-(1.7), which will be called here as fundamental generalized skew distributions (FGS). We start with the study of some general properties of this class of distributions, considering the following definition based on (1.6)-(1.7).

**Definition 3.1** Let \( Z \) be a \( k \times 1 \) random vector with density \( f_Z \) and let \( X \) be a \( m \times 1 \) random vector. If \( f_Z \) is a symmetric density on \( \mathbb{R}^k \), we say that \( Z^* = [Z \mid X > 0] \) has a \( k \)-variate fundamental generalized skew (FGS) distribution, which will be denoted by \( Z^* \sim FGS_{k,m}(K_m, f_Z, Q_m) \), if its density is given by

\[
f_{Z^*}(z) = K_m^{-1} f_Z(z) Q_m(z),
\]

where \( K_m = P(X > 0) \) and \( Q_m(z) = P(X > 0 \mid Z = z) \).

Note that \( K_m \) is a normalizing constant and that the term \( Q_m \) may be interpreted as a skewing function. Thus, as was mentioned in Section 1.2, from (3.1) we can obtain different families of skew distributions. For instance, if we assume that \( X \sim N_m(\mu, \Sigma) \) and \( Z \sim N_k(-\nu, \Omega + D\Sigma D^T) \), with \( \text{Cov}(Z,X) = \Sigma D^T \), then the \( SN \) distribution introduced by Gonzáles-Farías et al. (2002) follows from (3.1), which reduces to (1.5) when \( \nu = 0 \) and \( \Sigma = \Omega = I_k \). A special case is obtained when \( K_m = 2^{-m} \), which is satisfied by any random vector \( X \in C \), where \( C \) is the class of all symmetric random vectors \( X \), with \( P(X = 0) = 0 \) and such that: \( |X| = (|X_1|, \ldots, |X_m|)^T \) and \( \text{sign}(X) = (W_1, \ldots, W_m)^T \) are independent, and \( \text{sign}(X) \sim U_m \), where \( W_i = -1 \), if \( X_i < 0 \), \( i = 1, \ldots, m \), and \( U_m \) is the uniform distribution on \( \{-1,1\}^m \) (see Arellano-Valle et al., 2002). In this case, we have, for example, a generalization of (1.9) by letting \( X \sim N_m(0, I_m) \) and \( Z \sim N_k(\mu, \Omega + DD^T) \), with \( \text{Cov}(Z,X) = D^T \), obtaining from (3.1) that

\[
f_{Z^*}(z) = 2^m \phi_k(z|\mu, \Omega + DD^T) \Phi_m(D^T(\Omega + DD^T)^{-1}(z - \mu)|I_m - D^T(\Omega + DD^T)^{-1}D),
\]
which generalizes the Sahu et al. (2003) SN distribution. Note that to obtain the SN distribution introduced by these authors, we need to assume that \( m = k \) and that both matrices, \( \Omega \) and \( D \), are diagonal.

### 3.1 Marginal and conditional distributions

In the next result, we characterize the marginal and conditional densities originated by (3.1), for which we consider the partition \( \mathbf{Z} = (\mathbf{Z}_i, i = 1, 2) \) and the induced partition on \( \mathbf{Z}^* = (\mathbf{Z}_i^*, i = 1, 2) \), where \( \mathbf{Z}_i \) and \( \mathbf{Z}_i^* \) are \( k_i \times 1 \) random vectors, \( i = 1, 2 \), with \( k_1 + k_2 = k \).

**Proposition 3.1** Let \( \mathbf{Z}^* \sim FGS_{k,m} \) and let \( f_{\mathbf{Z}_i} \) be the marginal density of \( \mathbf{Z}_i, i = 1, 2 \). Then, the marginal density of \( \mathbf{Z}_i^* \) is

\[
f_{\mathbf{Z}_i^*}(\mathbf{z}_i) = K_m^{-1} f_{\mathbf{z}_i}(\mathbf{z}_i) Q_{i,m}(\mathbf{z}_i), \text{ with } K_m = P(\mathbf{X} > 0) \text{ and } Q_{i,m}(\mathbf{z}_i) = P(\mathbf{X} > 0 | \mathbf{Z}_i = \mathbf{z}_i),
\]

i.e., \( \mathbf{Z}_i^* \sim FGS_{k_i,m}(K_m, f_{\mathbf{Z}_i}, Q_{i,m}), i = 1, 2 \).

**Proof:** We give the proof for \( i = 1 \). The proof for \( i = 2 \) is analogous. By (3.1),

\[
f_{\mathbf{Z}_1^*}(\mathbf{z}_1) = K_m^{-1} \int_{\mathbb{R}^{k_2}} f_{\mathbf{z}_1, \mathbf{z}_2}(\mathbf{z}_1, \mathbf{z}_2) Q_m(\mathbf{z}_1, \mathbf{z}_2) d\mathbf{z}_2
\]

\[
= K_m^{-1} f_{\mathbf{z}_1}(\mathbf{z}_1) \int_{\mathbb{R}^{k_2}} f_{\mathbf{z}_2, \mathbf{z}_1=\mathbf{z}_1}(\mathbf{z}_2) Q_m(\mathbf{z}_1, \mathbf{z}_2) d\mathbf{z}_2,
\]

where, by the properties of the conditional expectation,

\[
\int_{\mathbb{R}^{k_2}} f_{\mathbf{z}_2 | \mathbf{z}_1=\mathbf{z}_1}(\mathbf{z}_2) Q_m(\mathbf{z}_1, \mathbf{z}_2) d\mathbf{z}_2 = E(Q_m(\mathbf{Z}_1, \mathbf{Z}_2) | \mathbf{Z}_1 = \mathbf{z}_1)
\]

\[
= E(P(\mathbf{X} > 0 | \mathbf{Z}_1 = \mathbf{z}_1, \mathbf{Z}_2) | \mathbf{Z}_1 = \mathbf{z}_1)
\]

\[
= P(\mathbf{X} > 0 | \mathbf{Z}_1 = \mathbf{z}_1)
\]

\[
= Q_{1,m}(\mathbf{z}_1).
\]

\( \Box \)

**Corollary 3.1** The conditional density of \( \mathbf{Z}_1^* \) given \( \mathbf{Z}_2^* = \mathbf{z}_2 \) is

\[
f_{\mathbf{Z}_1^* | \mathbf{Z}_2^* = \mathbf{z}_2}(\mathbf{z}_1) = \frac{f_{\mathbf{z}_1, \mathbf{z}_2}(\mathbf{z}_1, \mathbf{z}_2) Q_m(\mathbf{z}_1, \mathbf{z}_2)}{f_{\mathbf{z}_2}(\mathbf{z}_2) Q_{2,m}(\mathbf{z}_2)} = (K_m(\mathbf{z}_2))^{-1} f_{\mathbf{z}_1 | \mathbf{z}_2=\mathbf{z}_2}(\mathbf{z}_1) Q_{m}(\mathbf{z}_1 | \mathbf{z}_2),
\]

where

\[
K_m(\mathbf{z}_2) = Q_{2,m}(\mathbf{z}_2) \quad \text{and} \quad Q_m(\mathbf{z}_1, \mathbf{z}_2) = Q_{m}(\mathbf{z}_1, \mathbf{z}_2),
\]

so that \( [\mathbf{Z}_1^* | \mathbf{Z}_2^* = \mathbf{z}_2] \sim FGS_{k_1,m}(Q_{2,m}, f_{\mathbf{Z}_1 | \mathbf{z}_2}, Q_{m}). \)
Corollary 3.2 Consider the partition \( X = (X_i, i = 1, 2) \), where \( X_i \) is a \( m_i \times 1 \) random vector, \( i = 1, 2 \), with \( m_1 + m_2 = m > 1 \). If \((X_1, Z_1)\) and \((X_2, Z_2)\) are independent, then \( Z_1^* \) and \( Z_2^* \) are also independent, and their marginal densities are

\[
f_{Z_i^*}(z_i) = K_{m_i}^{-1} f_{Z_i}(z_i)Q_{m_i}(z_i),
\]

with \( K_{m_i} = P(X_i > 0) \) and \( Q_{m_i}(z_i) = P(X_i > 0 | Z_i = z_i), i = 1, 2 \).

Proof: In fact, the independence assumption implies \( K_m = P(X_1 > 0)P(X_2 > 0) = K_{m_1}K_{m_2}, \)
\( Q_{1,m}(z_1) = P(X_2 > 0)P(X_1 > 0 | Z_1 = z_1) = K_{m_1}Q_{m_1}(z_1) \) and, similarly, \( Q_{2,m}(z_2) = K_{m_2}Q_{m_2}(z_2) \), and \( Q_m(z_1, z_2) = P(X_1 > 0 | Z_1 = z_1)P(X_2 > 0 | Z_2 = z_2) = Q_{m_1}(z_1)Q_{m_2}(z_2) \). Thus, considering the result in Proposition 3.1, we have the above marginal densities. The independence result is obtained from Corollary 3.1.

\[\square\]

3.2 Obtaining skew distributions from symmetric linear combinations

We consider now the case where \( f_Z \) is the density of a linear combination \( Z = AX + BY \), for any given random vector \((X, Y) \in \mathbb{R}^{m+n}\) with symmetric density function \( f_{X,Y} \), and where it is assumed (without loss of generality) that \( n = k \) and that \( B \) is a nonsingular matrix. Thus, the skew random vector \( Z^* = [Z \ X > 0] \) has a density function as in (3.1), which can be obtained in terms of the joint density \( f_{X,Y} \) by using the following result.

Proposition 3.2 Let \((X, Y)\) be a random vector with symmetric density \( f_{X,Y} \) and such that \( Z = AX + BY \) has a density \( f_Z \). Then, \( Z^* = [Z \ X > 0] \) has density as in (3.1) with:

\[
f_Z(z) = \frac{1}{|B|} \int_{\mathbb{R}^m} f_{X,Y}(x, B^{-1}z - B^{-1}Ax)dx
\]

and

\[
Q_m(z) = \frac{\int_{\mathbb{R}^m} f_{X,Y}(x, B^{-1}z - B^{-1}Ax)dx}{\int_{\mathbb{R}^m} f_{X,Y}(x, B^{-1}z - B^{-1}Ax)dx},
\]

that is,

\[
f_{Z^*}(z) = \frac{K_m}{|B|} \int_{\mathbb{R}^m} f_{X,Y}(x, B^{-1}z - B^{-1}Ax)dx.
\]

Proof: Let \( V = \begin{pmatrix} X \\ Z \end{pmatrix} \) and \( U = \begin{pmatrix} X \\ Y \end{pmatrix} \). Note that

\[
V = \begin{pmatrix} X \\ AX + BY \end{pmatrix} = \begin{pmatrix} I_m & O \\ A & B \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = CU,
\]

where, since \( k = n \) and \( |B| \neq 0 \), the matrix

\[
C = \begin{pmatrix} I_m & O \\ A & B \end{pmatrix} \quad \text{implies} \quad |C| = |B| \quad \text{and} \quad C^{-1} = \begin{pmatrix} I_m & O \\ -B^{-1}A & B^{-1} \end{pmatrix}.
\]
Thus,
\[ \mathbf{U} = C^{-1} \mathbf{V} = \begin{pmatrix} I_m & 0 \\ -B^{-1}A & B^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Z} \end{pmatrix} = \begin{pmatrix} \mathbf{X} \\ B^{-1} \mathbf{Z} - B^{-1} \mathbf{A} \mathbf{X} \end{pmatrix}. \]

By the Jacobian method, \( f_{\mathbf{V}}(\mathbf{v}) = \frac{1}{|\mathbf{v}|} f_{\mathbf{U}}(C^{-1} \mathbf{v}) \), that is,
\[ f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}, \mathbf{z}) = \frac{1}{|\mathbf{B}|} f_{\mathbf{X}, \mathbf{Y}}(\mathbf{x}, B^{-1} \mathbf{z} - B^{-1} \mathbf{A} \mathbf{x}), \]
from where the proof follows.
\[ \square \]

Note that if we assume that the random vectors \( \mathbf{X} \) and \( \mathbf{Y} \) are uncorrelated with zero mean vector and finite covariance matrices, then we have that
\[ V(\mathbf{Z}) = AV(\mathbf{X})A^T + BV(\mathbf{Y})B^T, \quad \text{Cov}(\mathbf{Z}, \mathbf{X}) = AV(\mathbf{X}) \quad \text{and} \quad \text{Cov}(\mathbf{Z}, \mathbf{Y}) = BV(\mathbf{Y}). \]

Hence, if we consider the canonical situation where \( V(\mathbf{X}) = I_m, V(\mathbf{Y}) = I_n \) and \( V(\mathbf{Z}) = I_k \), then we can choose the matrices \( A \) and \( B \) so that
\[ A = \text{Corr}(\mathbf{Z}, \mathbf{X}), \quad B = \text{Corr}(\mathbf{Z}, \mathbf{Y}) \quad \text{and} \quad AA^T + BB^T = I_k. \]

Thus, for any given \( k \times m \) correlation matrix \( A \), we have that \( BB^T = I_k - AA^T \). Moreover, since we are assuming that \( B \) is a \( k \times k \) matrix with \( \text{rank}(B) = k \), we can choose \( B = (I_k - AA^T)^{1/2} \).

Note finally that the matrix \( A \) may be constructed as \( A = \Lambda(I_m + \Lambda^T \Lambda)^{-1/2} \), so that \( B = (I_k + \Lambda \Lambda^T)^{-1/2} \), for any given \( k \times m \) matrix \( \Lambda \). Alternatively, by taking \( \Lambda = B^{-1}A \), we have that \( A = (I_k + \Lambda \Lambda^T)^{-1/2} \Lambda \) and \( B = (I_k + \Lambda \Lambda^T)^{-1/2} \).

The assumptions considered above are satisfied, for example, when \( (\mathbf{X}, \mathbf{Y}) \) is a \( C \)-random vector. In such case, as was shown in Arellano-Valle et al. (2002), we have that the normalizing constant is \( K_m = 2^{-m} \) and we have the following marginal stochastic representation for any skew random vector \( \mathbf{Z}^\ast \) considered in the previous proposition.

**Proposition 3.3** Let \( \mathbf{Z}^\ast = \{ \mathbf{Z} | \mathbf{X} > 0 \} \), where \( \mathbf{Z} = \mathbf{AX} + \mathbf{BY} \), with \( (\mathbf{X}, \mathbf{Y}) \in C \). Then,
\[ \mathbf{Z}^\ast \overset{d}{=} A|\mathbf{X}| + \mathbf{BY}. \]

Moreover, if we assume the existence of the necessary moments, then
\[ E(\mathbf{Z}^\ast) = AE(|\mathbf{X}|) \quad \text{and} \quad V(\mathbf{Z}^\ast) = AV(|\mathbf{X}|)A^T + BV(\mathbf{Y})B^T. \]

We can use here the notation \( \mathbf{Z}^\ast \sim GFS_{k,m}(K_m, f_\mathbf{z}, \Delta) \), where \( f_\mathbf{z} \) is the density of the symmetric random vector \( \mathbf{Z} = \mathbf{AX} + \mathbf{BY} \) and \( \Delta \) is a skewness \( k \times m \) matrix depending on \( A \) and \( B \) and such that \( ||\Delta a|| < 1 \), for any unitary \( k \times 1 \) vector \( a \).
4 The fundamental skew-spherical distribution

In this section, we consider the special case where the density of the skew random vector $\mathbf{Z}^*$ is obtained from the density of a symmetric $\mathcal{C}$-random vector $(\mathbf{X}, \mathbf{Y})$, which is spherically distributed. That is, we assume that

$$\mathbf{U} = \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim S_{m+n}(h),$$

for some density generator $h = h^{m+n}$. Thus, $f_{\mathbf{x},\mathbf{y}}(\mathbf{x},\mathbf{y}) = h^{m+n}(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$. Let $\mathbf{Z} = \mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{Y}$ and let $M = \mathbf{A}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T$. Assume also that rank$(M) = k$. Then, the properties of spherical (elliptical) distributions (see Fang et. al, 1990) imply $\mathbf{Z} \sim \mathcal{E}l_k(0, M; h)$ and has density

$$f_\mathbf{Z}(\mathbf{z}) = |M|^{-1/2} h^k(q(\mathbf{z})), \quad \text{with} \quad q(\mathbf{z}) = \mathbf{z}^T M^{-1} \mathbf{z},$$

where

$$h^k(u) = \int_0^\infty \frac{n^{(m+n-k)/2} \Gamma((m+n)/2) (m+n-k)/2-1}{u} h^{m+n}(u+v)dv$$

is a marginal generator. Moreover,

$$\mathbf{V} = \begin{pmatrix} \mathbf{X} \\ \mathbf{Z} \end{pmatrix} = \mathbf{C}\mathbf{U} \sim \mathcal{E}l_{m+k}(0, \mathbf{C}\mathbf{C}^T; h), \quad \text{where} \quad \mathbf{C}\mathbf{C}^T = \begin{pmatrix} \mathbf{I}_m & \mathbf{A}^T \\ \mathbf{A} & \mathbf{M} \end{pmatrix},$$

which implies

$$[\mathbf{X}|\mathbf{Z} = \mathbf{z}] \sim \mathcal{E}l_m(\mathbf{A}\mathbf{D}^{-1/2} \mathbf{z}, \mathbf{I}_m - \mathbf{A}\mathbf{D}^{-1/2}; h_{q(\mathbf{z})}),$$

whose conditional density is

$$f_{\mathbf{x}|\mathbf{z} = \mathbf{z}}(\mathbf{x}) = |\mathbf{I}_m - \mathbf{D}^{-1/2} h_{q(\mathbf{z})}^m((\mathbf{x} - \mathbf{D}^{-1/2} \mathbf{z})^T (\mathbf{I}_m - \mathbf{A}\mathbf{D}^{-1})^{-1} (\mathbf{x} - \mathbf{D}^{-1/2} \mathbf{z})), $$

where $\mathbf{D}^{-1/2} \mathbf{A}, \quad q(\mathbf{z}) = \mathbf{z}^T M^{-1} \mathbf{z}$, and

$$h^m_a(u) = \frac{h^{m+k}(u+a)}{h^k(a)} \quad (4.1)$$

is the conditional generator. Thus, considering the canonical form with $M = \mathbf{I}_k$, we get

$$[\mathbf{X}|\mathbf{Z} = \mathbf{z}] \sim \mathcal{E}l_m(\mathbf{D}^{-1/2} \mathbf{z}, \mathbf{I}_m - \mathbf{D}^{-1/2}; h_{q(\mathbf{z})}),$$

where $q(\mathbf{z}) = \|\mathbf{z}\|^2$ and $\mathbf{Z} \sim S_k(h)$. Now, denote by $H_{q(\mathbf{z})}^m(\mathbf{x}|\mathbf{I}_m - \mathbf{D}^{-1} \mathbf{D})$, $\mathbf{x} \in \mathbb{R}^m$, the cdf of $[\mathbf{X} - \mathbf{D}^{-1} \mathbf{Z}|\mathbf{Z} = \mathbf{z}] \sim \mathcal{E}l_m(0, \mathbf{I}_m - \mathbf{D}^{-1} \mathbf{D}; h_{q(\mathbf{z})})$ and consider the random vector defined by $\mathbf{Z}^* = [\mathbf{Z}|\mathbf{X} > 0]$. Note that $Q_m(\mathbf{z}) = P(\mathbf{X} > 0|\mathbf{Z} = \mathbf{z}) = H_{q(\mathbf{z})}^m(\mathbf{D}^{-1/2} \mathbf{z}) \mathbf{I}_m - \mathbf{D}^{-1/2} \mathbf{z})$ and $K_m = P(\mathbf{X} > 0) = 2^{-m}$, since $\mathbf{X} \sim S_m(h)$. Thus, considering this result, it follows from (3.1) that

$$f_{\mathbf{Z}^*}(\mathbf{z}) = 2^m h^k(q(\mathbf{z})) H_{q(\mathbf{z})}^m(\mathbf{D}^{-1/2} \mathbf{z}) \mathbf{I}_m - \mathbf{D}^{-1/2} \mathbf{z}), \quad \text{where} \quad q(\mathbf{z}) = \|\mathbf{z}\|^2. \quad (4.2)$$

In such case, we will say that $\mathbf{Z}^*$ has a fundamental skew-spherical distribution with generator $h$ and $k \times m$ shape matrix $\Delta$, which will be denoted by $\mathbf{Z}^* \sim \mathcal{F}\mathcal{S}S_{k,m}(\Delta, h)$.
Remark 4.1 Note from (4.1) that, for any \( a > 0 \),
\[
H_a^m(x, \Omega) = \frac{1}{h^k(a)} |\Omega|^{-1/2} \int_{v \leq x} h^{m+k}(q(v) + a)dv,
\]
where \( q(v) = v^T \Omega^{-1} v \), so that (4.2) can be rewritten as
\[
f_{z^*}(z) = 2^m I_m - \Delta^T \Delta^{-1/2} \int_{v \leq \Delta^T z} h^{m+k}(v^T (I_m - \Delta^T \Delta)^{-1} v + \|z\|^2)dv,
\]
which reduces to
\[
f_{z^*}(z) = 2^m \int_{u \leq (I_m - \Delta^T \Delta)^{-1/2} \Delta^T z} h^{m+k}(\|u\|^2 + \|z\|^2)du
\]
when the matrix \( \Delta^T \Delta \) is diagonal.

Example 4.1 Let \( h^N(u) = (2\pi)^{-N/2} \exp\{-u/2\} \) be the \( N \)-dimensional normal generator, where \( N = m + k \). Then, \( h^k(u) = (2\pi)^{-k/2} \exp\{-u/2\} \) and \( h^m(u) = h^m(u) \) for all \( a > 0 \). Thus (4.2) is reduced to (2.1), i.e., \( Z^* \sim FSN_{k,m}(\Delta) \).

Example 4.2 Let \( h^N(u) = c(N, \nu)\alpha^{\nu/2}/\{\alpha + u\}^{(N+\nu)/2} \), with \( c(N, \nu) = \frac{\Gamma((N+\nu)/2)}{\Gamma(\nu/2)^{N/2}} \), be the generator of a \( N \)-dimensional (generalized) Student-t distribution (see Arellano-Valle and Bol- farine, 1995), where \( \nu \) are the degrees of freedom and \( \alpha \) is a scale parameter. Denote this distribution by \( t_N(\alpha, \nu) \), and by \( t_N(\mu, \Sigma; \alpha, \nu) \) its respective location-scale extension, where \( N = m + k \). Then, \( h^k(u) = c(k, \nu)\alpha^{\nu/2}/\{\alpha + u\}^{(k+\nu)/2} \) and \( h^m(u) = c(m, \nu(m))\alpha^{\nu(m)/2}/\{\alpha + u\}^{(m+\nu(m))/2} \), where \( \nu(m) = \nu + N - m = \nu + k \) and \( \alpha_{\nu} = \alpha + a, a > 0 \). Hence, from (4.2) we have a fundamental skew-t distribution defined by the following density
\[
f_{z^*}(z) = 2^m c(k, \nu)\alpha^{\nu/2}/\{\alpha + q(z)\}^{(k+\nu)/2} T_m(\Delta^T z I_m - \Delta^T \Delta; \alpha + q(z), \nu + k),
\]
where \( q(z) = \|z\|^2 \) and \( T_m(\cdot|I_m - \Delta^T \Delta; \alpha + a, \nu + k) \) denotes the cdf of a \( t_m(0, I_m - \Delta^T \Delta; \alpha + a, \nu + k) \) distribution. We denote this fundamental skew-t distribution by \( Z^* \sim FSt_{k,m}(\Delta, \alpha, \nu) \).

The next result extends the normal marginal stochastic representation given in Proposition 2.1. Its proof is analogous to the normal case, which is a direct consequence of Proposition 3.3, and establishes that under the C-class, the marginal and conditional stochastic representations of \( Z^* \) are equivalent.

Proposition 4.1 Let \( Z^* \sim FSS_k(\Delta, h) \) the fundamental skew-spherical distribution defined by (4.2), and let \( Z = \Delta X + (I_k - \Delta^T \Delta)^{1/2} Y \), where \( \begin{pmatrix} X \\ Y \end{pmatrix} \sim S_{m+n}(h) \). Then,
\[
Z^* \overset{d}{=} \mathbb{I}(X > 0) \overset{d}{=} 2 \mathbb{I}(X > 0) + (I_k - \Delta^T \Delta)^{1/2} Y.
\]
Moreover,
\[
E(Z^*) = \tau_{1, h} \Delta I_m \quad \text{and} \quad V(Z^*) = \tau_{2, h} (I_k - \gamma_{1, h}^2 \Delta^T \Delta),
\]
where the moments \( \tau_{r, h} = E(|X|^r), r = 1, 2, \) with \( X \sim S_1(h) \), are assumed to be finite.
In the particular normal case, the extension for a fundamental skew-elliptical distribution \(FSEL_{k,m}(\mu, \Sigma; \Delta, h)\) can be obtained by considering the random vector \(W^* = \mu + \Sigma^{1/2}Z^*\), with \(Z^* \sim FSS_{k,m}(\Delta, h)\), whose density is

\[
f_{W^*}(w) = 2^m |\Sigma|^{-1/2} h^k(q(w)) H^m_m(\Delta^T \Sigma^{-1/2}(w - \mu) | I_m - \Delta^T \Delta),
\]

where \(q(w) = (w^T - \mu)^T \Sigma^{-1}(w - \mu)\).

Another special case of the \(C\)-class, is obtained by assuming that \(X \sim S_n(h_1)\) and \(Y \sim S_n(h_2)\) and are independent. Thus, \(f_{X,Y}(x, y) = h_1^m(\|x\|^2) h_2^m(\|y\|^2)\), so that, by Proposition 3.2,

\[
f_{Z^*}(z) = \frac{2^m}{|B|} \int_{\mathbb{R}^m} h_1^m(\|x\|^2) h_2^m(\|B^{-1}z - B^{-1}Ax\|^2)dx.
\]

This case is in general more complicated. However, if \(h_1\) and \(h_2\) are normal generators, then this case yields also the \(FSN\) distribution considered in Section 2.

**Acknowledgments** The first author was partially supported by the Project *Complementary Research Lines* 800004, Fondo Nacional de Ciencia y Tecnología de Chile.

**References**


Arnold, B. C. and Beaver, R. J. (2002), Skewed multivariate models related to hidden truncation and/or selective reporting. *Test* 58, 7-54.


Azzalini, A. (1986), Further results on a class of distributions which includes the normal ones. *Statistica* 44, 199-208.


