Available online at www.sciencedirect.com
ScienceDirect
Journal of the Korean Statistical Society 37 (2008) 385–392
JOURNAL OF THE KOREAN STATISTICAL SOCIETY
www.elsevier.com/locate/jkss

# Singular extended skew-elliptical distributions

José A. Díaz-García[a,*], Graciela González-Farías[b]

[a] Universidad Autónoma Agraria Antonio Narro, Department of Statistics and Computation, 25315 Buenavista, Saltillo, Coahuila, Mexico
[b] Department of Probability and Statistics, Centro de Investigación en Matemáticas, Callejón de Jalisco s/n, Mineral de Valenciana, 36240 Guanajuato, Guanajuato, Mexico

Received 19 January 2008; accepted 15 April 2008
Available online 21 May 2008

**Abstract**

Singular vector and matrix extended skew-elliptical distributions are studied in this work. Based on the vectorial case, two alternatives for singular matrix variate extended skew-elliptical distribution are also proposed. In addition, the distributions of a general linear transformation for extended skew-elliptical vectors and matrices are derived along with the corresponding density functions. These results are applied in the distribution of the residuals for a general linear model with extended skew-elliptical errors.


*AMS 2000 subject classifications:* primary 62E15; secondary 60E05

*Keywords:* Distribution of the residuals; Extended skew-distribution; Hausdorff measure; Linear transformation; Singular distribution; Singular matrix variate distribution

## 1. Introduction

In the last two decades, several research areas of univariate and multivariate statistics have replaced the hypothesis with classical normal assumptions by elliptical distribution conditions. Two summaries of those advances are given in the books by Fang and Zhang (1990) and Gupta and Varga (1993).

Some authors have studied a new family of distributions by introducing skewness in univariate and multivariate elliptical distributions, see Aigner, Lovell, and Schmidt (1977), Azzalini and Capitanio (1999), Azzalini and Dalla Valle (1996), Azzalini and Capitanio (2003) Sahu, Dey, and Branco (2003) and Genton (2004). This set of distributions contains several standard families, including multivariate skew normal and skew t distributions. In the normal case, different approaches have generated many of the multivariate skew-normal family of distributions, see for example, Branco and Dey (2001) and González-Farías, Domínguez, and Gupta (2004).

Simultaneously, many authors have studied the singular problem for normal distribution, and the vector and matrix elliptical families. Also, some related singular distributions have been proposed: the Wishart distribution, the matrix variate T, the matrix variate beta type I and II, among many others, see Díaz-García and González-Farías (2005a,b),

* Corresponding author.
*E-mail addresses:* jadiaz@uaaan.mx (J.A. Díaz-García), farias@cimat.mx (G. González-Farías).

1226-3192/$ - see front matter 
doi:10.1016/j.jkss.2008.04.004

Díaz-García and Gutiérrez-Jáimez (1997), Díaz-García and Gutiérrez-Jáimez (2006), Díaz-García, Gutiérrez-Jáimez, and Mardia (1997), Ip, Wong, and Liu (2007) and Uhlig (1994).

In the present work, we extend the results of González-Farías et al. (2004) by deriving the density function of a singular vector extended skew-elliptical distribution. Then, in Section 4, the vectorial case is also generalised, by proposing two expressions for a singular matrix variate extended skew-elliptical distribution. Finally, in Section 5, we consider the distribution of a general linear transformation for extended skew-elliptical vectors and matrices. These results are applied in the distribution of residuals for a general multivariate linear model when errors have a singular matrix variate extended skew-elliptical distribution.

## 2. Notations and preliminaries

Let $h : \Re \to [0, \infty)$ be a function such that $\int_0^\infty u^{Nm/2-1} h(u)\mathrm{d}u < \infty$. We say that a random matrix $\mathbf{Y} \in \Re^{N\times m}$ has an elliptical distribution with location parameter matrix $\boldsymbol{\mu} \in \Re^{N\times m}$ and scale parameter matrix $\mathbf{\Theta} \otimes \mathbf{\Xi} \in \Re^{Nm\times Nm}$, if its density function is given by

$$f_{\mathbf{Y}}(\mathbf{Y}) = |\mathbf{\Xi}|^{-N/2}|\mathbf{\Theta}|^{-m/2} h^{(N\times m)}\left[\mathrm{tr}\left(\mathbf{\Xi}^{-1}(\mathbf{Y}-\boldsymbol{\mu})^{\mathrm{T}}\mathbf{\Theta}^{-1}(\mathbf{Y}-\boldsymbol{\mu})\right)\right], \tag{1}$$

where $\mathbf{\Xi} > 0$ and $\mathbf{\Theta} > 0$ for $\mathbf{\Xi} \in \Re^{m\times m}$ and $\mathbf{\Theta} \in \Re^{N\times N}$, and $\otimes$ is the usual Kronecker product. The function $h^{(N\times m)}$ is called the density generator and the elliptical distribution of $\mathbf{Y}$ is denoted by $\mathbf{Y} \sim \mathcal{E}_{N\times m}(\boldsymbol{\mu}, \mathbf{\Theta}\otimes\mathbf{\Xi}, h^{(N\times m)})$. If the rank $r$ of $\mathbf{\Xi}$ is less than $m$ and/or the rank $k$ of $\Theta$ is less than $N$, this is, if $\mathbf{\Xi} \geq 0$ and/or $\mathbf{\Theta} \geq 0$, then, the distribution rank of $\mathbf{Y}$ is $kr$, i.e. $\mathbf{Y}$ has a singular distribution, see Cramér (1999, p. 297). In this case, we say that $\mathbf{Y}$ has a singular matrix-variate elliptical distribution, which is denoted by

$$\mathbf{Y} \sim \mathcal{E}_{N\times m}^{k,r}\left(\boldsymbol{\mu}, \mathbf{\Theta}\otimes\mathbf{\Xi}, h_{k,r}^{(N\times m)}\right).$$

The superscript (in $\mathcal{E}_{N\times m}^{k,r}$) or the subscript (in $h_{k,r}^{(N\times m)}$) will be omitted when $r = m$ and $k = N$. Subscript and superscript in $h$ indicate explicitly that this function depends on them.

So, we have:

**Lemma 2.1** (*Singular Matrix Variate Elliptical Distribution*). *Suppose that* $\mathbf{Y} \sim \mathcal{E}_{N\times m}^{k,r}\left(\mu, \mathbf{\Theta}\otimes\mathbf{\Xi}, h_{k,r}^{(N\times m)}\right)$ *and let* $\mathbf{\Xi}^-$ *and* $\mathbf{\Theta}^-$ *be some symmetric generalised inverses of* $\mathbf{\Xi}$ *and* $\mathbf{\Theta}$*, respectively. Then the density function of* $\mathbf{Y}$ *with respect to the Hausdorff measure* $(\mathrm{d}\mathbf{Y})$ *is*

$$\mathrm{d}F_{\mathbf{Y}}(\mathbf{Y}) = \frac{1}{\left(\prod_{j=1}^{k}\delta_j^{r/2}\right)\left(\prod_{i=1}^{r}\lambda_i^{k/2}\right)} h_{k,r}^{(N\times m)}\left(\mathrm{tr}\,\mathbf{\Xi}^-(\mathbf{Y}-\boldsymbol{\mu})^{\mathrm{T}}\mathbf{\Theta}^-(\mathbf{Y}-\boldsymbol{\mu})\right)(\mathrm{d}\mathbf{Y}), \tag{2}$$

*where* $\lambda_i$ *and* $\delta_j$ *are nonzero eigenvalues of* $\mathbf{\Xi}$ *and* $\mathbf{\Theta}$*, respectively, see also Díaz-García et al. (1997), Díaz-García and González-Farías (2005b),* Cramér (1999, p. 297) *and* Billingsley (1986, p. 247).

If $q = \min(r, k)$, explicit expressions for $(\mathrm{d}\mathbf{Y})$ can be given as functions of the QR, Polar, Singular value and QR modified decompositions, see Díaz-García and González-Farías (2005a).

Now, Khatri (1968) shows that the density function (2) is not unique, see also Díaz-García and González-Farías (2005b); however, it is important to note that once a distribution expression (2) is found, the results do not depend on the selected density, see Rao (1973).

As it was highlighted, interest in skew elliptical distributions comes from both theoretical and applied directions. On the theoretical side it enjoys of a number of formal properties which resemble those of the elliptical distributions given, for example, in Gupta and Varga (1993). From the applied viewpoint, these densities are unimodal empirical distributions with the presence of skewness and possible heavy tails. In the non-singular case, as we shall see, the term skew-elliptical comes from the parametric class of multivariate probability distributions determined by the vector $\mathbf{Y} = [\mathbf{Z}|\mathbf{W} > \mathbf{0}]$, where $\mathbf{Z} \sim \mathcal{E}_q(\tau_1, \Upsilon_1, h^{(q)})$ and $\mathbf{W} \sim \mathcal{E}_p(\tau_2, \Upsilon_2, h^{(p)})$. This definition for $p = 1$, under the elliptical model, is given in Branco and Dey (2001), and the general normal case, for arbitrary $p$, is studied by González-Farías et al. (2004).

## 3. Singular vector-variate skew-elliptical distribution

In this section we propose an expression for vector singular extended skew-elliptical density, extending the normal case derived by González-Farías et al. (2004).

Assume that

$$\mathbf{E} = \begin{pmatrix} \mathbf{E}_1 \\ \mathbf{E}_2 \end{pmatrix} \sim \mathcal{E}_{p+q}^{r+k}\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Delta} \end{pmatrix}, h_{r+k}^{(p+q)}\right),$$

where $\mathbf{E}_1 : p \times 1$, $\boldsymbol{\Sigma} \geq 0$ has rank $r \leq p$, $\mathbf{E}_2 : q \times 1$, $\boldsymbol{\Delta} \geq 0$ has rank $k \leq q$ and $\mathrm{Cov}(\mathbf{E}_1, \mathbf{E}_2) = 0 : p \times q$. Note that $\mathbf{E}_1$ and $\mathbf{E}_2$ are dependent; which is the opposite situation in the normal case.

Let

$$\mathbf{U} = \begin{pmatrix} \mathbf{I}_p & \mathbf{0} \\ \mathbf{D} & \mathbf{I}_q \end{pmatrix}\begin{pmatrix} \mathbf{E}_1 \\ \mathbf{E}_2 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\mu} \\ -\boldsymbol{\nu} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu} + \mathbf{E}_1 \\ -\boldsymbol{\nu} + \mathbf{D}\mathbf{E}_1 + \mathbf{E}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{W} \\ -\mathbf{Z} \end{pmatrix}, \tag{3}$$

where $\mathbf{D} : q \times p$ is an arbitrary matrix of constants, and, $\boldsymbol{\mu} : p \times 1$ and $\boldsymbol{\nu} : q \times 1$, are vectors of constants. Then

$$\mathbf{U} = \begin{pmatrix} \mathbf{W} \\ \mathbf{Z} \end{pmatrix} \sim \mathcal{E}_{p+q}^{r+k_1}\left(\begin{pmatrix} \boldsymbol{\mu} \\ -\boldsymbol{\nu} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}} \\ D\boldsymbol{\Sigma} & \boldsymbol{\Delta} + \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}} \end{pmatrix}, h_{r+k_1}^{(p+q)}\right),$$

where $k_1$ is the rank of $\boldsymbol{\Delta} + \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}}$. If $G(\cdot)$ is the distribution function of $g(\cdot)$, then

$$\mathrm{d}G_{\mathbf{W}|\{\mathbf{Z}\geq\mathbf{0}\}}(\mathbf{w}|\mathbf{Z} \geq \mathbf{0}) = \frac{\mathrm{d}G_{\mathbf{W}}(\mathbf{w})}{P(\mathbf{Z} \geq \mathbf{0})} P(\mathbf{Z} \geq \mathbf{0}|\mathbf{W} = \mathbf{w}), \tag{4}$$

with

$$\mathbf{W} \sim \mathcal{E}_p^r\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}, h_r^{(p)}\right) \quad \text{and} \quad \mathbf{Z} \sim \mathcal{E}_q^{k_1}\left(-\boldsymbol{\nu}, \boldsymbol{\Delta} + D\boldsymbol{\Sigma}D^{\mathrm{T}}, h_{k_1}^{(q)}\right).$$

For the random $s$-dimensional vector we will denote this fact by $g_{\mathbf{V}}^{(s)}\left(\mathbf{v}; r, \boldsymbol{\mu}, \boldsymbol{\Sigma}, h_r^{(s)}\right)$ or $\mathrm{d}G_{\mathbf{V}}^{(s)}\left(\mathbf{v}; r, \boldsymbol{\mu}, \boldsymbol{\Sigma}, h_r^{(s)}\right)$; where $r$ is the rank of the distribution; i.e. the rank of the matrix $\boldsymbol{\Sigma}$, see Cramér (1999, p. 297). Then, given that $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-}\boldsymbol{\Sigma}$, there is

$$\mathbf{Z}|\mathbf{W} = \mathbf{w} \sim \mathcal{E}_q^k\left(-\boldsymbol{\nu} + \mathbf{D}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-}(\mathbf{w} - \boldsymbol{\mu}), \boldsymbol{\Delta}, h_{\delta(\mathbf{w}),k_1}^{(q)}\right),$$

with $\delta(\mathbf{w}) = (\mathbf{w} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-}(\mathbf{w} - \boldsymbol{\mu})$ (see Theorem 2.6.4, pp. 62–65 of Gupta and Varga (1993)), where

$$h_{\delta(\mathbf{w}),k_1}^{(q)}(\tau) = \frac{\Gamma(k_1/2)}{\pi^{k_1/2}} \frac{h(\delta(\mathbf{w}) + \tau, p+q)}{\int_{\mathfrak{R}^+} v^{k_1/2-1} h(\alpha + v, p+q)\mathrm{d}v}, \quad \alpha > 0,$$

and $h(\cdot,\cdot)$ is a decreasing function, $h : \mathfrak{R}^+ \to \mathfrak{R}^+$, such that

$$\int_{\mathfrak{R}^+} h(a,b)a^{b/2-1}\mathrm{d}a < \infty.$$

Then,

$$P(\mathbf{Z} \geq \mathbf{0}) = F_{\mathbf{Z}}^{(q)}\left(0; k_1, \boldsymbol{\nu}, \boldsymbol{\Delta} + \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}}, h_{k_1}^{(q)}\right),$$

and

$$P(\mathbf{Z} \geq \mathbf{0}|\mathbf{W} = \mathbf{w}) = F_{\{\mathbf{Z}\geq\mathbf{0}\}|\mathbf{W}=\mathbf{w}}^{(q)}\left(D\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-}(\mathbf{w} - \boldsymbol{\mu}); k, \boldsymbol{\nu}, \boldsymbol{\Delta}, h_{\delta(\mathbf{w}),k}^{(q)}\right).$$

Thus, the density (4) can be expressed as

$$\mathrm{d}G_{\mathbf{W}|\{\mathbf{Z}\geq\mathbf{0}\}}(\mathbf{w}|\mathbf{Z} \geq \mathbf{0}) = \frac{F_{\mathbf{Z}\geq\mathbf{0}|\mathbf{W}=\mathbf{w}}^{(q)}\left(\mathbf{D}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-}(\mathbf{y} - \boldsymbol{\mu}); k, \nu, \boldsymbol{\Delta}, h_{\delta(\mathbf{w}),k}^{(q)}\right)}{F_{\mathbf{Z}}^{(q)}\left(\mathbf{0}; k_1, \boldsymbol{\nu}, \boldsymbol{\Delta} + \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}}, h_{k_1}^{(q)}\right)} \mathrm{d}G_{\mathbf{W}}^{(p)}\left(\mathbf{w}; r, \boldsymbol{\mu}, \boldsymbol{\Sigma}, h_r^{(p)}\right).$$

In conclusion:

**Definition 3.1** (*Singular Vector-Variate Extended Skew-Elliptical Distribution*). A random vector $\mathbf{Y}$ has a $p$-dimensional singular extended skew-elliptical distribution, with rank $r$ and parameters $q$, $k_1$, $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, $k$, $\mathbf{D}$, $\boldsymbol{\nu}$, $\boldsymbol{\Delta}$ previously given in Lemma 2.1, if its density function is given by

$$dG_{\mathbf{Y}}^{(p)}\left(\mathbf{y}; r, q, k_1, \boldsymbol{\mu}, \boldsymbol{\Sigma}, k, \mathbf{D}, \boldsymbol{\nu}, \boldsymbol{\Delta}, h_r^{(p)}\right)$$
$$= \frac{F_{\mathbf{Y}}^{(q)}\left(\mathbf{D}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-}(\mathbf{y}-\boldsymbol{\mu}); k, \boldsymbol{\nu}, \boldsymbol{\Delta}, h_{\delta(\mathbf{w}),k}^{(q)}\right)}{F_{\mathbf{Y}}^{(q)}\left(\mathbf{0}; k_1, \boldsymbol{\nu}, \boldsymbol{\Delta}+\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}}, h_{k_1}^{(q)}\right)} dG_{\mathbf{Y}}^{(p)}\left(\mathbf{y}; r, \boldsymbol{\mu}, \boldsymbol{\Sigma}, h_r^{(p)}\right), \tag{5}$$

and it is denoted by

$$\mathbf{Y} \sim \mathcal{SESE}_r^{(p)}\left(q, k_1, \boldsymbol{\mu}, \boldsymbol{\Sigma}, k, \mathbf{D}, \boldsymbol{\nu}, \boldsymbol{\Delta}, h_r^{(p)}\right).$$

Some important particular cases of this family are the following: (a)If $\boldsymbol{\Delta} > 0$, then $k = q = k_1$, and the parameters $k$ and $k_1$ are excluded in the density (5), (b) if $\boldsymbol{\Sigma} > 0$, then $r = p$ and the parameter $r$ is excluded in Definition 3.1, (c) if $\boldsymbol{\Delta} > 0$ and $\boldsymbol{\Sigma} > 0$ then non-singular extended skew-elliptical distribution is obtained; in that case the parameters $r$, $k$ and $k_1$ in Definition 3.1 are excluded, see González-Farías et al. (2004). Finally, note that the distribution $\mathcal{SESE}$ is not unique, since the singular elliptical distribution exhibits this characteristic.

## 4. Singular matrix variate extended skew-elliptical distribution

In this section we study singular extended skew-elliptical distribution for the matrix case. We shall see that the matrix distribution can be obtained as an extension of the vector distribution described in Section 3. Note that the matrix extension of the vectorial case is not unique, in fact, we can follow essentially four techniques very similar to the classical extension of the vectorial elliptical distribution, see Fang and Zhang (1990, Lemma 3.3.2). Even more, an alternative generalisation for the matrix version can be proposed, and this is given at the end of the section.

First, observe that $\mathbf{Y} \sim \mathcal{E}_{N\times m}(\boldsymbol{\mu}, \boldsymbol{\Theta}\otimes\boldsymbol{\Xi}, h)$ is equivalent to

$$\operatorname{vec}\mathbf{Y} \sim \mathcal{E}_{Nm}(\operatorname{vec}\boldsymbol{\mu}, \boldsymbol{\Theta}\otimes\boldsymbol{\Xi}, h),$$

see Muirhead (1982, p. 79) and Gupta and Varga (1993, pp. 26–27). Then, by assuming

$$\begin{pmatrix}\operatorname{vec}\mathbf{E}_1\\ \operatorname{vec}\mathbf{E}_2\end{pmatrix} \sim \mathcal{E}_{pmqn}^{r_{\boldsymbol{\Sigma}}r_{\boldsymbol{\Theta}}r_{\boldsymbol{\Delta}}r_{\boldsymbol{\Xi}}}\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}, \begin{pmatrix}\boldsymbol{\Theta}\otimes\boldsymbol{\Sigma} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Xi}\otimes\boldsymbol{\Delta}\end{pmatrix}, h_{r_{\boldsymbol{\Sigma}}r_{\boldsymbol{\Theta}}r_{\boldsymbol{\Delta}}r_{\boldsymbol{\Xi}}}^{(pmqn)}\right), \tag{6}$$

where $\mathbf{E}_1 : p\times m$ and $\mathbf{E}_2 : q\times n$ are matrices; $\boldsymbol{\Sigma} : p\times p$ has rank $r_{\boldsymbol{\Sigma}} \le p$, $\boldsymbol{\Sigma} \ge 0$; $\boldsymbol{\Theta} : m\times m$ has rank $r_{\boldsymbol{\Theta}} \le m$, $\boldsymbol{\Theta} \ge 0$; $\boldsymbol{\Delta} : q\times q$ has rank $r_{\boldsymbol{\Delta}} \le q$, $\boldsymbol{\Delta} \ge 0$ and $\boldsymbol{\Xi} : n\times n$ has rank $r_{\boldsymbol{\Xi}} \le n$, $\boldsymbol{\Xi} \ge 0$. Then the matrix version of the model (3) is given by

$$\operatorname{vec}\mathbf{U} = \begin{pmatrix}\mathbf{I} & \mathbf{0}\\ (\mathbf{D}_2^{\mathrm{T}}\otimes\mathbf{D}_1) & \mathbf{I}\end{pmatrix}\begin{pmatrix}\operatorname{vec}\mathbf{E}_1\\ \operatorname{vec}\mathbf{E}_2\end{pmatrix} + \begin{pmatrix}\operatorname{vec}\boldsymbol{\mu}\\ -\operatorname{vec}\boldsymbol{\nu}\end{pmatrix}, \tag{7}$$

where $\mathbf{D}_1 : q\times p$; $\mathbf{D}_2 : m\times n$; $\boldsymbol{\mu} : p\times m$ and $\boldsymbol{\nu} : q\times n$ are arbitrary matrices of constants. Explicitly,

$$\operatorname{vec}\mathbf{U} = \begin{pmatrix}\operatorname{vec}\mathbf{W}\\ \operatorname{vec}\mathbf{Z}\end{pmatrix} = \begin{pmatrix}\operatorname{vec}\boldsymbol{\mu} + \operatorname{vec}\mathbf{E}_1\\ -\operatorname{vec}\boldsymbol{\nu} + (\mathbf{D}_2^{\mathrm{T}}\otimes\mathbf{D}_1)\operatorname{vec}\mathbf{E}_1 + \operatorname{vec}\mathbf{E}_2\end{pmatrix}.$$

Proceeding as in Section 3 (after the model (3)), we have the following:

**Definition 4.1** (*Singular Matrix-Variate Extended Skew-Elliptical Distribution I*). It is said that a random matrix $\mathbf{Y}$ has a singular matrix variate extended skew-elliptical $p\times m$-dimensional distribution, of rank $r_{\boldsymbol{\Sigma}}r_{\boldsymbol{\Theta}}$ and parameters $q, n, k_1, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Theta}, k = r_{\boldsymbol{\Delta}}r_{\boldsymbol{\Xi}}, \mathbf{D}_1, \mathbf{D}_2, \boldsymbol{\nu}, \boldsymbol{\Delta}, \boldsymbol{\Xi}$, previously defined, if its density function is given by

$$dG_{\operatorname{vec}\mathbf{Y}}^{(pm)}\left(\operatorname{vec}\mathbf{Y}; r_{\boldsymbol{\Sigma}}r_{\boldsymbol{\Theta}}, q, n, k_1, \operatorname{vec}\boldsymbol{\mu}, \boldsymbol{\Theta}\otimes\boldsymbol{\Sigma}, k, \mathbf{D}_2^{\mathrm{T}}\otimes\mathbf{D}_1, \operatorname{vec}\boldsymbol{\nu}, \boldsymbol{\Xi}\otimes\boldsymbol{\Delta}, h_{r_{\boldsymbol{\Sigma}}r_{\boldsymbol{\Theta}}}^{(pm)}\right)$$

$$= \frac{F_{\operatorname{vec}\mathbf{Y}}^{(qn)}\left(\mathbf{D}_2^{\mathrm{T}}\mathbf{\Theta}\mathbf{\Theta}^- \otimes \mathbf{D}_1\mathbf{\Sigma}\mathbf{\Sigma}^-\operatorname{vec}(\mathbf{Y}-\boldsymbol{\mu}); k, \operatorname{vec}\boldsymbol{\nu}, \mathbf{\Xi}\otimes\mathbf{\Delta}, h_{\delta(\mathbf{W}),k}^{(qn)}\right)}{F_{\operatorname{vec}\mathbf{Y}}^{(qn)}\left(\mathbf{0}; k_1, \operatorname{vec}\boldsymbol{\nu}, \mathbf{\Xi}\otimes\mathbf{\Delta}+\mathbf{D}_2^{\mathrm{T}}\mathbf{\Theta}\mathbf{D}_2\otimes\mathbf{D}_1\mathbf{\Sigma}\mathbf{D}_1^{\mathrm{T}}, h_{k_1}^{(qn)}\right)}$$
$$\times \, dG_{\operatorname{vec}\mathbf{Y}}^{(pm)}\left(\operatorname{vec}\mathbf{Y}; r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}, \operatorname{vec}\boldsymbol{\mu}, \mathbf{\Theta}\otimes\mathbf{\Sigma}, h_{r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}}^{(pm)}\right),$$

where $\delta(\mathbf{W}) = \operatorname{vec}^{\mathrm{T}}(\mathbf{W}-\boldsymbol{\mu})(\mathbf{\Theta}\otimes\mathbf{\Sigma})^-\operatorname{vec}(\mathbf{W}-\boldsymbol{\mu})$. Under matrix notation,

$$dG_{\mathbf{Y}}^{(p\times m)}\left(\mathbf{Y}; r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}, q, n, k_1, \boldsymbol{\mu}, \mathbf{\Theta}\otimes\mathbf{\Sigma}, k, \mathbf{D}_2^{\mathrm{T}}\otimes\mathbf{D}_1, \boldsymbol{\nu}, \mathbf{\Xi}\otimes\mathbf{\Delta}, h_{r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}}^{(pm)}\right)$$
$$= \frac{F_Y^{(q\times n)}\left(\mathbf{D}_1\mathbf{\Sigma}\mathbf{\Sigma}^-(\mathbf{Y}-\boldsymbol{\mu})\mathbf{\Theta}^-\mathbf{\Theta}\mathbf{D}_2; k, \boldsymbol{\nu}, \mathbf{\Xi}\otimes\mathbf{\Delta}, h_{\delta(\mathbf{W}),k}^{(q\times n)}\right)}{F_Y^{(q\times n)}\left(\mathbf{0}; k_1, \boldsymbol{\nu}, \mathbf{\Xi}\otimes\mathbf{\Delta}+\mathbf{D}_2^{\mathrm{T}}\mathbf{\Theta}\mathbf{D}_2\otimes\mathbf{D}_1\mathbf{\Sigma}\mathbf{D}_1^{\mathrm{T}}, h_{k_1}^{(q\times n)}\right)} dG_Y^{(p\times m)}\left(Y; r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}, \boldsymbol{\mu}, \mathbf{\Theta}\otimes\mathbf{\Sigma}, h_{r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}}^{(p\times m)}\right),$$

where $\delta(\mathbf{W}) = \operatorname{tr}\mathbf{\Sigma}^-(\mathbf{W}-\boldsymbol{\mu})^{\mathrm{T}}\mathbf{\Theta}^-(\mathbf{W}-\boldsymbol{\mu})$; $k_1$ is the rank of $(\mathbf{\Xi}\otimes\mathbf{\Delta}+\mathbf{D}_2^{\mathrm{T}}\mathbf{\Theta}\mathbf{D}_2\otimes\mathbf{D}_1\mathbf{\Sigma}\mathbf{D}_1^{\mathrm{T}})$. We denote this fact by

$$\mathbf{Y} \sim \mathcal{SESE}_{r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}}^{(p\times m)}\left(q, n, k_1, \boldsymbol{\mu}, \mathbf{\Theta}\otimes\mathbf{\Sigma}, k, \mathbf{D}_2^{\mathrm{T}}\otimes\mathbf{D}_1, \boldsymbol{\nu}, \mathbf{\Xi}\otimes\mathbf{\Delta}, h_{r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}}^{(p\times m)}\right).$$

Now observe that $\operatorname{Cov}(\operatorname{vec}\mathbf{E}_1) = \mathbf{\Theta}\otimes\mathbf{\Sigma}$; this structure of the covariance matrix through the Kronecker product is a consequence of linear transformation acting on a matrix. For example, in the same context of Definition 4.1, if

$$\mathbf{V} \sim \mathcal{E}_{r_{\mathbf{\Sigma}}\times r_{\mathbf{\Theta}}}\left(\mathbf{0}, \mathbf{I}_{r_{\mathbf{\Theta}}}\otimes\mathbf{I}_{r_{\mathbf{\Sigma}}}, h^{(r_{\mathbf{\Sigma}}\times r_{\mathbf{\Theta}})}\right),$$

where $\mathbf{0}$ is a matrix of zeros of order $r_{\mathbf{\Sigma}}\times r_{\mathbf{\Theta}}$, then $\mathbf{E}_1 = \mathbf{MVN}$, and

$$\mathbf{E}_1 \sim \mathcal{E}_{p\times m}^{r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}}\left(\mathbf{0}, \mathbf{\Theta}\otimes\mathbf{\Sigma}, h_{r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}}^{(N\times m)}\right),$$

with $\mathbf{\Sigma} = \mathbf{MM}^{\mathrm{T}}$ and $\mathbf{\Theta} = \mathbf{N}^{\mathrm{T}}\mathbf{N}$.

The disadvantages of this approach – which uses the transformation of a matrix – are the restrictions on elements of the covariance matrix, $\mathbf{\Theta}\otimes\mathbf{\Sigma}$, see Press (1982, p. 253).

An alternative approach considers vectorization of the matrix to perform the linear procedure. It avoids functions of Kronecker products in the linear transformation. For the example, we know that:

$$\operatorname{vec}\mathbf{V} \sim \mathcal{E}_{r_{\mathbf{\Sigma}}\times r_{\mathbf{\Theta}}}\left(\operatorname{vec}\mathbf{0}, \mathbf{I}_{r_{\mathbf{\Theta}}}\otimes\mathbf{I}_{r_{\mathbf{\Sigma}}}, h^{(r_{\mathbf{\Sigma}}\times r_{\mathbf{\Theta}})}\right) \equiv \mathcal{E}_{r_{\mathbf{\Sigma}}\times r_{\mathbf{\Theta}}}\left(\operatorname{vec}\mathbf{0}, \mathbf{I}_{r_{\mathbf{\Theta}}r_{\mathbf{\Sigma}}}, h^{(r_{\mathbf{\Sigma}}\times r_{\mathbf{\Theta}})}\right).$$

Then we define $\operatorname{vec}\mathbf{E}_1 = \mathbf{A}\operatorname{vec}\mathbf{V}$, with $\mathbf{A} : pm \times r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}$ such that $\mathbf{\Lambda} = \mathbf{AA}^{\mathrm{T}}$. So

$$\operatorname{vec}\mathbf{E}_1 \sim \mathcal{E}_{pm}^{r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}}\left(\operatorname{vec}\mathbf{0}, \mathbf{\Lambda}, h_{r_{\mathbf{\Sigma}}r_{\mathbf{\Theta}}}^{(N\times m)}\right).$$

Using this observation instead of (6) we have

$$\begin{pmatrix}\operatorname{vec}\mathbf{E}_1\\ \operatorname{vec}\mathbf{E}_2\end{pmatrix} \sim \mathcal{E}_{pmqn}^{r_{\mathbf{\Lambda}}r_{\mathbf{\Omega}}}\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{\Lambda} & \mathbf{0}\\ \mathbf{0} & \mathbf{\Omega}\end{pmatrix}, h_{r_{\mathbf{\Lambda}}r_{\mathbf{\Omega}}}^{(pmqn)}\right),$$

where $\mathbf{\Lambda} : pm\times pm$ of rank $r_{\mathbf{\Lambda}} \le pm$, $\mathbf{\Lambda} \ge 0$ and $\mathbf{\Omega} : qn\times qn$ of rank $r_{\mathbf{\Omega}} \le qn$, $\mathbf{\Omega} \ge 0$. Now, an alternative definition of model (7) is

$$\operatorname{vec}U = \begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathcal{D} & \mathbf{I}\end{pmatrix}\begin{pmatrix}\operatorname{vec}\mathbf{E}_1\\ \operatorname{vec}\mathbf{E}_2\end{pmatrix} + \begin{pmatrix}\operatorname{vec}\boldsymbol{\mu}\\ -\operatorname{vec}\boldsymbol{\nu}\end{pmatrix},$$

where $\mathcal{D} : nq\times mp$; $\boldsymbol{\mu} : p\times m$ and $\boldsymbol{\nu} : q\times n$ are arbitrary matrices of constants. Explicitly,

$$\operatorname{vec}\mathbf{U} = \begin{pmatrix}\operatorname{vec}\mathbf{W}\\ \operatorname{vec}\mathbf{Z}\end{pmatrix} = \begin{pmatrix}\operatorname{vec}\boldsymbol{\mu}+\operatorname{vec}\mathbf{E}_1\\ -\operatorname{vec}\boldsymbol{\nu}+\mathcal{D}\operatorname{vec}\mathbf{E}_1+\operatorname{vec}\mathbf{E}_2\end{pmatrix}.$$

Then, we propose the following generalisation of Definitions 3.1 and 4.1, which holds for the vector and matrix case.

**Definition 4.2** (*Singular Matrix Variate Extended Skew-Elliptical Distribution II*). A random matrix $\mathbf{Y}$ has a $pm$-dimensional singular matrix variate extended skew-elliptical distribution, with rank $r_{\mathbf{\Lambda}}$ and parameters $q$, $n$, $k_1$, $\boldsymbol{\mu}$, $\mathbf{\Lambda}$, $k$, $\mathcal{D}$, $\boldsymbol{\nu}$, $\mathbf{\Omega}$, previously defined, if its density function is given by

$$\begin{aligned}&\mathrm{d}G^{(pm)}_{\operatorname{vec}\mathbf{Y}}\left(\operatorname{vec}\mathbf{Y}; r_{\mathbf{\Lambda}}, q, k_1, \boldsymbol{\mu}, \mathbf{\Lambda}, k, \mathcal{D}, \boldsymbol{\nu}, \mathbf{\Omega}, h^{(pm)}_{r_{\mathbf{\Lambda}}}\right)\\ &\quad= \frac{F^{(qn)}_{\operatorname{vec}\mathbf{Y}}\left(\mathcal{D}\mathbf{\Lambda}\mathbf{\Lambda}^{-}\operatorname{vec}(\mathbf{Y}-\boldsymbol{\mu}); k, \operatorname{vec}\boldsymbol{\nu}, \mathbf{\Omega}, h^{(qn)}_{\delta(\operatorname{vec}\mathbf{W}),k}\right)}{F^{(qn)}_{\operatorname{vec}\mathbf{Y}}\left(\mathbf{0}; k_1, \operatorname{vec}\boldsymbol{\nu}, \mathbf{\Omega}+\mathcal{D}\mathbf{\Lambda}\mathcal{D}^{\mathrm{T}}, h^{(qn)}_{k_1}\right)}\mathrm{d}G^{(pm)}_{\operatorname{vec}\mathbf{Y}}\left(\operatorname{vec}\mathbf{Y}; r_{\mathbf{\Lambda}}, \operatorname{vec}\boldsymbol{\mu}, \mathbf{\Lambda}, h^{(pm)}_{r_{\mathbf{\Lambda}}}\right),\end{aligned}$$

where $k_1$ is the rank of $\mathbf{\Omega}+\mathcal{D}\mathbf{\Lambda}\mathcal{D}^{\mathrm{T}}$ and $\delta(\operatorname{vec}\mathbf{W}) = \operatorname{vec}^{\mathrm{T}}(\mathbf{W}-\boldsymbol{\mu})\mathbf{\Lambda}^{-}\operatorname{vec}(\mathbf{W}-\boldsymbol{\mu})$. We denote this by

$$\mathbf{Y} \sim \mathcal{SESE}2^{(pm)}_{r_{\mathbf{\Lambda}}}\left(q, k_1, \operatorname{vec}\boldsymbol{\mu}, \mathbf{\Lambda}, k, \mathcal{D}, \operatorname{vec}\boldsymbol{\nu}, \mathbf{\Delta}, h^{(pm)}_{r_{\mathbf{\Lambda}}}\right).$$

## 5. General linear transformation

The first result of this section finds the distribution of the general linear transformation $\mathbf{AY}+\mathbf{b}$, when $\mathbf{Y} \sim \mathcal{SESE}^{(p)}_r\left(q, k_1, \boldsymbol{\mu}, \mathbf{\Sigma}, k, \mathbf{D}, \boldsymbol{\nu}, \mathbf{\Delta}, h^{(p)}_r\right)$, $\mathbf{b}$ is a vector of constants and $\mathbf{A}$ is any matrix of constants. This problem has been studied by González-Farías et al. (2004) in the skew-normal distribution case and under different conditions on the rank of the matrix $\mathbf{A}$. Specifically, they considered $\mathbf{A}$, as a non-singular matrix; $\mathbf{Y}$, with a non-singular distribution; and $\mathbf{A}$, of rank $m \le n$.

At the end of the section we apply these results finding the distribution of residuals for a multivariate linear model.

**Theorem 5.1.** *Assume that* $\mathbf{Y} \sim \mathcal{SESE}^{(p)}_r\left(q, k_1, \boldsymbol{\mu}, \mathbf{\Sigma}, k, \mathbf{D}, \boldsymbol{\nu}, \mathbf{\Delta}, h^{(p)}_r\right)$. *Let* $\mathbf{A}$ *be an* $s \times p$ *matrix of constants of rank* $s_1 \le \min(s, p)$ *and let* $\mathbf{b}$ *be a constant* $s \times 1$ *vector. And also consider* $\mathbf{a}_j \in \mathcal{I}m(\mathbf{A\Sigma A}^{\mathrm{T}})$ *for all* $j = 1, \dots, q$, *where* $\mathbf{a}_j$ *are the columns of the matrix* $\mathbf{A\Sigma D}^{\mathrm{T}}$ *and* $\mathcal{I}m(N)$ *denotes the image of the matrix* $N$. *Then,*

$$\mathbf{AY}+\mathbf{b} \sim \mathcal{SESE}^{(s)}_{s_2}\left(q, k_1, \mathbf{A}\boldsymbol{\mu}+\mathbf{b}, \mathbf{\Sigma}_{\mathbf{A}}, k_2, D_{\mathbf{A}}, \boldsymbol{\nu}, \mathbf{\Delta}_{\mathbf{A}}, h^{(s)}_{s_2}\right),$$

*where*

$$\begin{aligned}&\mathbf{\Sigma}_{\mathbf{A}} = \mathbf{A\Sigma A}^{\mathrm{T}},\\ &D_{\mathbf{A}} = D\mathbf{\Sigma A}^{\mathrm{T}}\mathbf{\Sigma}^{-}_{\mathbf{A}},\\ &\mathbf{\Delta}_{\mathbf{A}} = \mathbf{\Delta} + D(\mathbf{\Sigma} - \mathbf{\Sigma A}^{\mathrm{T}}\mathbf{\Sigma}^{-}_{\mathbf{A}}\mathbf{A\Sigma})\mathbf{D}^{\mathrm{T}},\\ &s_2 \text{ is the rank of } (\mathbf{A\Sigma A}^{\mathrm{T}}),\\ &k_1 \text{ is the rank of } (\mathbf{\Delta}_{\mathbf{A}} + \mathbf{D}_{\mathbf{A}}\mathbf{\Sigma}_{\mathbf{A}}\mathbf{D}^{\mathrm{T}}_{\mathbf{A}})(= \text{ to rank of } (\mathbf{\Delta}+\mathbf{D\Sigma D}^{\mathrm{T}})),\\ &k_2 \text{ is the rank of } \mathbf{\Delta}_{\mathbf{A}}.\end{aligned}$$

**Proof.** Define $\mathbf{V} = \mathbf{AU}+\mathbf{b}_1$, with

$$B = \begin{pmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{pmatrix}, \qquad \mathbf{U} = \begin{pmatrix}\mathbf{W}\\ \mathbf{Z}\end{pmatrix}, \quad \text{and} \quad \mathbf{b}_1 = \begin{pmatrix}\mathbf{b}\\ \mathbf{0}\end{pmatrix},$$

where

$$E(\mathbf{V}) = E\begin{pmatrix}\mathbf{AW}+\mathbf{b}\\ \mathbf{Z}\end{pmatrix} = \begin{pmatrix}\mathbf{A}\boldsymbol{\mu}+\mathbf{b}\\ -\boldsymbol{\nu}\end{pmatrix},$$

and

$$\operatorname{Cov}(\mathbf{V}) = \mathbf{B}\operatorname{Cov}(\mathbf{V})\mathbf{B}^{\mathrm{T}} = \begin{pmatrix}\mathbf{A\Sigma A}^{\mathrm{T}} & \mathbf{A\Sigma D}^{\mathrm{T}}\\ \mathbf{D\Sigma A}^{\mathrm{T}} & \mathbf{\Delta}+\mathbf{D\Sigma D}^{\mathrm{T}}\end{pmatrix}.$$

Then

$$\mathbf{V} \sim \mathcal{E}_{s+q}^{s_2+k_1}\left(\begin{pmatrix}\mathbf{A}\boldsymbol{\mu}+\mathbf{b}\\ -\boldsymbol{\nu}\end{pmatrix}, \begin{pmatrix}\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}} & \mathbf{A}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}}\\ \mathbf{D}\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}} & \boldsymbol{\Delta}+\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}}\end{pmatrix}, h_{s_2+k_1}^{(s+q)}\right),$$

where $s_2$ is the rank of $(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}})$ and, as before, $k_1$ is the rank of $(\boldsymbol{\Delta}+\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}})$. But observe that

$$\begin{pmatrix}\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}} & \mathbf{A}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}}\\ \mathbf{D}\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}} & \boldsymbol{\Delta}+\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}}\end{pmatrix} = \begin{pmatrix}\boldsymbol{\Sigma}_{\mathbf{A}} & \boldsymbol{\Sigma}_{\mathbf{A}}\mathbf{D}_{\mathbf{A}}^{\mathrm{T}}\\ \mathbf{D}_{\mathbf{A}}\boldsymbol{\Sigma}_{\mathbf{A}} & \boldsymbol{\Delta}_{\mathbf{A}}+\mathbf{D}_{\mathbf{A}}\boldsymbol{\Sigma}_{\mathbf{A}}\mathbf{D}_{\mathbf{A}}^{\mathrm{T}}\end{pmatrix},$$

where $\boldsymbol{\Sigma}_{\mathbf{A}} = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}}$ and $\mathbf{D}_{\mathbf{A}} = \mathbf{D}\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{A}}^{-}$, for which

$$\boldsymbol{\Sigma}_{\mathbf{A}}\mathbf{D}_{\mathbf{A}}^{\mathrm{T}} = \boldsymbol{\Sigma}_{\mathbf{A}}\boldsymbol{\Sigma}_{\mathbf{A}}^{-}\mathbf{A}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}} = \mathbf{A}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}}.$$

The last equation is valid when $\mathbf{a}_j \in \mathcal{I}m(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}})$ for all $j = 1,\dots,q$, where the $\mathbf{a}_j$ are the columns of the matrix matrix $\mathbf{A}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}}$, noting that $\boldsymbol{\Sigma}_{\mathbf{A}}\boldsymbol{\Sigma}_{\mathbf{A}}^{-}$ is the projector of the image of $\boldsymbol{\Sigma}_{\mathbf{A}}$. Now, observing that, if $\boldsymbol{\Delta}_{\mathbf{A}} = \boldsymbol{\Delta}\boldsymbol{\Delta}+\mathbf{D}(\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{A}}^{-}\mathbf{A}\boldsymbol{\Sigma})\mathbf{D}^{\mathrm{T}}$, then

$$\mathbf{D}_{\mathbf{A}}\boldsymbol{\Sigma}_{\mathbf{A}}\mathbf{D}_{\mathbf{A}}^{\mathrm{T}} = \mathbf{D}\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{A}}^{-}\boldsymbol{\Sigma}_{\mathbf{A}}\boldsymbol{\Sigma}_{\mathbf{A}}^{-}\mathbf{A}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}} = \mathbf{D}\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{A}}^{-}\mathbf{A}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}},$$

we have $\boldsymbol{\Delta}_{\mathbf{A}}+\mathbf{D}_{\mathbf{A}}\boldsymbol{\Sigma}_{\mathbf{A}}\mathbf{D}_{\mathbf{A}}^{\mathrm{T}} = \boldsymbol{\Delta}+\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^{\mathrm{T}}$. Finally, if $\mathbf{Y} \stackrel{d}{=} \mathbf{W}|\{\mathbf{Z} \geq \mathbf{0}\}$, then $A\mathbf{Y}+\mathbf{b} \stackrel{d}{=} A\mathbf{W}+\mathbf{b}|\{\mathbf{Z} \geq \mathbf{0}\}$, where $\stackrel{d}{=}$ denotes equality distributed.The expected result is reached by applying a similar procedure in the proof of (4). ■

Observe that if $\boldsymbol{\Sigma} > 0$, $\boldsymbol{\Delta} > 0$ and $s_1 = s \leq p$, then $s_2 = s$, $k_1 = q$ and $k_2 = q$, and, in the notation of González-Farías et al. (2004),

$$A\mathbf{Y}+\mathbf{b} \sim \mathcal{ESE}_{s,q}\left(\mathbf{A}\boldsymbol{\mu}+\mathbf{b}, \boldsymbol{\Sigma}_{\mathbf{A}}, D_{\mathbf{A}}, \boldsymbol{\nu}, \boldsymbol{\Delta}_{\mathbf{A}}, h\right).$$

Similar results to Theorem 5.1 can be derived in the matrix case by using Definitions 4.1 and 4.2.

**Corollary 5.1.** *Consider the general multivariate linear model* $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta}+\boldsymbol{\xi}$ *where* $\mathbf{Y} : N \times m$, $\mathbf{X} : N \times l$, *of rank* $\tau \leq l \leq N$, $\boldsymbol{\beta} : l \times m$ *and*

$$\boldsymbol{\xi} \sim \mathcal{SESE}_{r_{\Sigma}N}^{(N\times m)}\left(q, n, k_1, \mathbf{0}, I_N \otimes \boldsymbol{\Sigma}, k, \mathbf{D}_2^{\mathrm{T}} \otimes \mathbf{D}_1, \boldsymbol{\nu}, \boldsymbol{\Xi} \otimes \boldsymbol{\Delta}, h_{r_{\Sigma}N}^{(N\times m)}\right).$$

*If* $\mathbf{R} : N \times m$ *denotes the residual matrix, then*

$$\mathbf{R} \sim \mathcal{SESE}_{s_2N}^{(N\times m)}\left(q, n, k_1, \mathbf{0}, (\mathbf{I}_N \otimes \boldsymbol{\Sigma})_{\mathbf{A}}, k, (\mathbf{D}_2^{\mathrm{T}} \otimes \mathbf{D}_1)_{\mathbf{A}}, \boldsymbol{\nu}, (\boldsymbol{\Xi} \otimes \boldsymbol{\Delta})_{\mathbf{A}}, h_{s_2N}^{(N\times m)}\right),$$

*where* $\mathbf{A} = (\mathbf{I} \otimes \mathbf{P})$, $\mathbf{P} = (\mathbf{I}-\mathbf{X}\mathbf{X}^{+})$ *with* $\mathbf{C}^{+}$ *is the Moore-Penrose inverse of the matrix* $\mathbf{C}$, *and*

$$(\mathbf{I}_N \otimes \boldsymbol{\Sigma})_{\mathbf{A}} = (\mathbf{I}_N \otimes \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}),$$

$s_2$ *is the rank of* $(\mathbf{I}_N \otimes \boldsymbol{\Sigma})_{\mathbf{A}}$,

$$(\mathbf{D}_2^{\mathrm{T}} \otimes \mathbf{D}_1)_{\mathbf{A}} = (\mathbf{D}_2^{\mathrm{T}} \otimes \mathbf{D}_1\boldsymbol{\Sigma}\mathbf{P}(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P})^{-}) \quad \textit{and}$$

$$(\boldsymbol{\Xi} \otimes \boldsymbol{\Delta})_{\mathbf{A}} = \boldsymbol{\Xi} \otimes \boldsymbol{\Delta} + (\mathbf{D}_2^{\mathrm{T}} \otimes \mathbf{D}_1)(\mathbf{I}_N \otimes \boldsymbol{\Sigma} - \mathbf{I}_N \otimes \boldsymbol{\Sigma}\mathbf{P}(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P})^{-}\mathbf{P}\boldsymbol{\Sigma})(\mathbf{D}_2 \otimes \mathbf{D}_1^{\mathrm{T}}).$$

**Proof.** Recall that $\mathbf{R} = \mathbf{Y}-\widehat{\mathbf{Y}} = \mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}} = \mathbf{Y}-\mathbf{X}\mathbf{X}^{+}\mathbf{Y} = (\mathbf{I}-\mathbf{X}\mathbf{X}^{+})\mathbf{Y} = \mathbf{P}\mathbf{Y}$, where $\mathbf{P} = (\mathbf{I}-\mathbf{X}\mathbf{X}^{+})$ and $\widehat{\boldsymbol{\beta}}$ is any solution of the system of normal matrix equations $(\mathbf{X}^{\mathrm{T}}\mathbf{X})\widehat{\boldsymbol{\beta}} = \mathbf{X}^{\mathrm{T}}\mathbf{Y}$, see Rao (1973) or Muirhead (1982).

Now observe that, the linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta}+\boldsymbol{\xi}$ is a linear transformation of the matrix $\boldsymbol{\xi}$, and $\operatorname{vec}\mathbf{Y} = (\mathbf{I} \otimes \mathbf{X})\operatorname{vec}\boldsymbol{\beta}+\operatorname{vec}\boldsymbol{\xi}$, thus by Theorem 5.1 we obtain

$$\mathbf{Y} \sim \mathcal{SESE}_{r_{\Sigma}N}^{(N\times m)}\left(q, n, k_1, \mathbf{X}\boldsymbol{\beta}, \mathbf{I}_N \otimes \boldsymbol{\Sigma}, k, \mathbf{D}_2^{\mathrm{T}} \otimes \mathbf{D}_1, \boldsymbol{\nu}, \boldsymbol{\Xi} \otimes \boldsymbol{\Delta}, h_{r_{\Sigma}N}^{(N\times m)}\right).$$

Now, applying Theorem 5.1 and observing that $\operatorname{vec}\mathbf{R} = (\mathbf{I} \otimes \mathbf{P})\operatorname{vec}\mathbf{Y}$, then the proof is complete. ■

## 6. Conclusions

It is easy to check that all univariate and multivariate nonsingular extended skew elliptically contoured distributions can be obtained as particular cases of the results given in this work. In the same way, distributions of any kind of general linear transformation $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$ of a nonsingular extended skew elliptical distribution with all their variants (dimension of $\mathbf{A}$ and/or ranks of $\mathbf{A}$ and $\mathbf{X}$), also can be derived as corollaries of Section 5. Many interesting applications are generated by this result. Unfortunately, at present, some expressions can only be seen as theoretical results, because even in the univariate case there are important problems for classical estimation of parameters, see Azzalini (2005). In fact, some theoretical results for singular distributions of different kind of residuals (normalized, standardized, internally or externally studentized residual) cannot be established because the jacobians with respect to the Hausdorff measure are unknown; for example, no expressions for the simplest jacobian related with a linear transformation is given in literature, see Díaz-García (2007). In addition, from a Bayesian point of view we propose the a-priori parameter distributions of the corresponding matrix variate extended skew elliptically distribution and by using Definition 4.2 we avoid the parameter restrictions of Definition 4.1, see Press (1982, p. 253).

## Acknowledgments

The authors wish to express their gratitude to the editor and the referees for their helpful comments and suggestions. This research work was partially supported by CONACYT-México, Research Grant No. 45974-F and IDI-Spain, grant MTM2005-09209. Also, thanks to Francisco José Caro Lopera (CIMAT-México) for his careful reading and excellent comments.

## References

Aigner, D. J., Lovell, C. A. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function model. *Journal of Econometrics*, *12*, 21–37.

Azzalini, A. (2005). The multivariate skew normal distribution and related multivariate families. *Scandinavian Journal of Statistics*, *32*, 159–188.

Azzalini, A., & Capitanio, A. (1999). Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society B*, *61*, 579–602.

Azzalini, A., & Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew $t$-distribution. *Journal of the Royal Statistical Society B*, *65*, 367–389.

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

Billingsley, P. (1986). *Probability and measure* (2nd ed.). New York: John Wiley & Sons.

Branco, M., & Dey, D. K. (2001). A general class of multivariate skew elliptical distributions. *Journal of Multivariate Analysis*, *79*, 99–115.

Cramér, H. (1999). *Mathematical methods of statistics* (Nineteenth printing). Princeton University Press.

Díaz-García, J. A. (2007). A note about measures and Jacobians of singular random matrices. *Journal of Multivariate Analysis*, *98*, 960–969.

Díaz-García, J. A., & González-Farías, G. (2005a). Singular random matrix decompositions: Jacobians. *Journal of Multivariate Analysis*, *93*, 196–212.

Díaz-García, J. A., & González-Farías, G. (2005b). Singular random matrix decompositions: Distributions. *Journal of Multivariate Analysis*, *94*, 109–122.

Díaz-García, J. A., & Gutiérrez-Jáimez, R. (1997). Proof of conjectures of H. Uhlig on the singular multivariate beta and the Jacobian of a certain matrix transformation. *The Annals of Statistics*, *25*, 2018–2023.

Díaz-García, J. A., & Gutiérrez-Jáimez, R. (2006). Distribution of the generalised inverse of a random matrix and its applications. *Journal of Statistical Planning and Inference*, *136*, 183–192.

Díaz-García, J. A. , R., Gutiérrez-Jáimez, K. V., & Mardia, (1997). Wishart and Pseudo-Wishart distributions and some applications to shape theory. *Journal of Multivariate Analysis*, *63*, 73–87.

Fang, K. T., & Zhang, Y. T. (1990). *Generalized multivariate analysis*. Beijing: Science Press, Springer-Verlag.

Genton, M. G. (2004). *Skew-elliptical distributions and their applications: A journey beyond normality*. Boca Raton, FL: Chapman & Hall/CRC.

González-Farías, G. A., Domínguez, M., & Gupta, A. (2004). The closed skew normal distribution. In Marc G. Genton (Ed.), *Skew - elliptical distributions and their applications: A journey beyond normality* (pp. 25–42). USA: Chapman & Hall/CRC. Chapter 2.

Gupta, A. K., & Varga, T. (1993). *Elliptically contoured models in statistics*. Dordrecht: Kluwer Academic Publishers.

Ip, W. Ch., Wong, H., & Liu, J. S. (2007). Inverse Wishart distributions based on singular elliptically contoured distribution. *Linear Algebra and its Applications*, *420*, 424–432.

Khatri, C. G. (1968). Some results for the singular normal multivariate regression models. *Sankyā A*, *30*, 267–280.

Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. New York: John Wiley & Sons.

Press, S. J. (1982). In Robert E. Krieger (Ed.), *Applied multivariate analysis: Using Bayesian and frequentist methods of inference* (2nd ed.). Malabar, FL: Publishing Company.

Rao, C. R. (1973). *Linear startistical inference and its applications* (2nd ed.). New York: John Wiley & Sons.

Sahu, S. K., Dey, D. K., & Branco, M. D. (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *Canadian Journal of Statistics*, *31*, 129–150.

Uhlig, H. (1994). On singular Wishart and singular multivartiate beta distributions. *The Annals of Statistics*, *22*, 395–405.