SINGULAR MATRIX VARIATE SKEW-ELLIPTICAL DISTRIBUTION AND THE DISTRIBUTION OF GENERAL LINEAR TRANSFORMATION

José A. Díaz-García and Graciela González-Farías

Singular matrix variate skew-elliptical distribution and the distribution of general linear transformation

José A. Díaz-García [1]

Universidad Autónoma Agraria Antonio Narro
Department of Statistics and Computation
25315 Buenavista, Saltillo
Coahuila, México

Graciela González-Farías

Centro de Investigación en Matemáticas
Department of Probability and Statistics
Callejón de Jalisco s/n, Mineral de Valenciana
36240 Guanajuato
Guanajuato, México

Abstract

In the case of a singular random matrix, an expression is proposed for the skew-elliptical distribution. Under this general framework we derive the distribution of a linear transformation of a matrix without any rank restrictions showing all its properties, and extending in this way, the results of singular elliptical random matrices to the family of skew-elliptical distributions. In particular we are able to establish the distribution of the residuals in a general multivariate linear model, defining the singular matrix variate skew-elliptical distribution.

Key words: Skew-distribution, Hausdorff measure, singular distribution, linear transformation, matrix-variate singular distribution, residual distribution
PACS: 62E15, 60E05.

Email addresses: `jadiaz@uaaan.mx` (José A. Díaz-García), `farias@cimat.mx` (Graciela González-Farías).
[1] This article was written during his stay as Visiting Professor at the Dept. of Probability and Statistics of the Centro de Investigación en Matemáticas, Guanajuato, México.

1 Introduction

In the last two decades, many of the statistical techniques, particulary in the multivariate framework, have substituted the assumption of normality to an assumption of elliptical distribution. These techniques have been grouped as Generalized Multivariate Analysis, see Gupta & Varga (1993) and Fang and Zhang (1990). Two of the most important impacts of this development are the family of elliptic distribution which in some cases is well known for distributions such as normal distribution, Pearson Type II, t and Kotz distributions; and the family of elliptic distributions, similar to the test statistic distribution under the assumption of normality, many of the test statistics are consistent throughout.

In the late part of the 90's, special attention was devoted to the study of the different versions of the family of skew-elliptic distributions. The case of the univariate, its generalization and other extensions to the multivariate case have been studied by many researchers, among many: see Aigner *et al.* (1977), Azzalini and Dalla Valle, A. (1996), Sahu *et al.* (2003) and Genton (2004).

In the normal cases, these studies have created the family of multivariate skew-normal, with different approaches, see Branco and Dey (2001), González-Farías et al. (2004), just to mention a few.

In parallel, studies issues have been addressed on the multivariate singular distributions (vector and matrix cases), see Uhlig (1994), Díaz-García and Gutiérrez-Jáimez (1997), Díaz-García et al. (1997), Díaz-García and Gutiérrez-Jáimez (2004), Díaz-García and González-Farías (2004c) and Díaz-García and González-Farías (2004b). Nevertheless, some of the problems had been already studied for the normal case before, as we can see from Khatri (1968) and Rao (1973).

An interesting issue for the statistical modelling can be formulated as follows, given a vector (or a random matrix) $X : n \times 1$, look for its distribution in a linear application, i.e. the distribution of the $Y = AX + b$, where $A : m \times n$ and $b : m \times 1$ are not random. Its significance as we know, is in the general linear model which is characterized by a form similar to linear applications. This problem has been studied for the case of skew-normal distribution allowing more general conditions on the rank of the matrix A, in González-Farías et al. (2004), that is, in the case of A being a non-singular matrix, X has a non singular distribution and the rank of $A = m \leq n$. However, it is also interesting to find the distribution of Y when X has a singular distribution, and for any relation between m and n or when A is actually singular. As it is noted, the residuals in a general linear model (multivariate or univariate) have a singular distribution, moreover, they are obtained through a linear transformation of

the model's error distribution. This is critical in the sensitivity analysis of a linear model and in general, all the test statistics utilized to detect the influential points when using the technique of eliminating one or several observations, are a function of some of the residual classes. Consequently, the distribution of those test statistics is a function of the distribution of the residuals, which is singular, see Chatterjee and Hadi (1988) and Díaz-García and González-Farías (2004a). It is then of relevance to know the singular distribution particularly in the general multivariate linear model, at least when the variance-covariance matrix is singular and therefore, the error matrix has also a singular distribution. This issue has been studied in the case of normality in Khatri (1968) from a classic approach, and in Díaz-García and Gutiérrez-Jáimez (2004) from a bayesian point of view.

The present work contains an expression for the density function of an elliptic singular random matrix, highlighting the fact that such density, for a given elliptic distribution is not unique, see section 2. In section 3, there is an expression for a singular extended skew-elliptical distribution (SESE) for the vectorial case, whose result is extended to the matrix case in two versions. Finally, section 4 includes a distribution of a linear general transformation of a random vector with SESE distribution. The conclusion, determines the distribution of the residuals in a multivariate general linear model, when assuming the errors have a matrix distribution SESE.

2 Notation and preliminary results

Let $\mathcal{L}_{m,N}(q)$ be the linear space of all $N \times m$ real matrices of rank $q \leq \min(N, m)$; $\mathcal{L}_{m,N}^+(q)$ be the linear space of all $N \times m$ real matrices of rank $q \leq \min(N, m)$ with q distinct singular values. The set of matrices $H_1 \in \mathcal{L}_{m,N}(m)$ such that $H_1^T H_1 = I_m$ is the Stiefel manifold denoted by $\mathcal{V}_{m,N}$. In particular, $\mathcal{V}_{m,m}$ is the group of orthogonal matrices $\mathcal{O}(m)$.

Definition 1 (Matrix-variate Singular Elliptical Distribution) *. Let $Y \in \mathcal{L}_{m,N}^+(q)$, such that $Y \sim \mathcal{E}_{N \times m}(\mu, \Theta \otimes \Xi, h)$, with $\Xi : m \times m$ of rank $r < m$ or $\Theta : N \times N$ of rank $k < N$. This distribution will be called a matrix-variate singular elliptically contoured distribution and will be denoted by*

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

omitting the supra-index when $r = m$ and $k = N$. In addition, its density

3

function is given by

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

$$\left.\begin{array}{l} Q_2^T(Y-\mu)M_1^T = 0 \\ Q_1^T(Y-\mu)M_2^T = 0 \\ Q_2^T(Y-\mu)M_2^T = 0 \end{array}\right\} \; a. \; s. \tag{2}$$

for some function $h_{k,r}^{(N\times m)}$, where the supra-index denotes the dimension of the matrix Y and the index denotes the rank of the distribution. Also A^- denotes a symmetric generalized inverse, λ_i and δ_j are the nonzero eigenvalues of Ξ and Θ respectively. Let $Q = (Q_1|Q_2) \in \mathcal{O}(N)$ and $M = (M_1^T|M_2^T) \in \mathcal{O}(m)$ be matrices associated with the spectral decomposition of matrices Ξ and Θ respectively with $Q_1 \in \mathcal{V}_{k,N}$, $Q_2 \in \mathcal{V}_{N-k,N}$, $M_1^T \in \mathcal{V}_{r,m}$ and $M_2^T \in \mathcal{V}_{m-r,m}$, see Díaz-García and Gutiérrez-Jáimez (2003) and Díaz-García and González-Farías (2004c).

Alternatively, this density can be written as (for the Normal distribution case, see Khatri (1968))

$$dF_Y(Y) = \frac{1}{\left(\prod_{i=1}^{r}\lambda_i^{k/2}\right)\left(\prod_{j=1}^{k}\delta_j^{r/2}\right)}h_{k,r}^{(N\times m)}\left(\operatorname{tr}\Xi^-(Y-\mu)^T\Theta^-(Y-\mu)\right)(dY), \tag{3}$$

where (dY) is the Hausdorff measure, which coincides with that of Lebesgue measure when it is defined on the subspace $\mathcal{M}$ given by the hyperplanes (2), see Díaz-García et al. (1997), Cramér (1999, p. 297) and Billingsley (1986, p. 247). Observe that $\mathcal{M}$ is an affine subspace. If $q = \min(r,k)$, explicit expressions of (dY) can be given as function of QR, Polar, SV and QR modified decompositions, see Díaz-García and González-Farías (2004b).

Remark 2 *It is important to note that the density function (3) is not unique, because Ξ^- and Θ^- are not unique and the explicit form of (dY) is not unique either, see Khatri (1968) and Díaz-García and González-Farías (2004c). However, once the density expression is found (3), the results do not depend on the selected density, see Rao (1973).*

3 Singular skew-elliptical distribution

In this section an expression for singular skew-elliptical density in the vectorial case is proposed, and based on that we obtain, two alternatives for a matrix-variate singular skew-elliptical distribution.

Assuming that

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

where $\mathbf{E}_1 : p \times 1$, $\Sigma \geq 0$ of rank $r \leq p$, $\mathbf{E}_2 : q \times 1$, $\Delta \geq 0$ with rank $k \leq q$, $\mathrm{Cov}(\mathbf{E}_1, \mathbf{E}_2) = 0 : p \times q$, noting $\mathbf{E}_1$ and $\mathbf{E}_2$ are not independent, like in the case of a normal.

Let $\mathbf{U} = A\mathbf{E} + \rho$ be defined as

$$\mathbf{U} = \begin{pmatrix} I_p & 0 \\ D & I_q \end{pmatrix} \begin{pmatrix} \mathbf{E}_1 \\ \mathbf{E}_2 \end{pmatrix} + \begin{pmatrix} \mu \\ -\nu \end{pmatrix} = \begin{pmatrix} \mu + \mathbf{E}_1 \\ -\nu + D\mathbf{E}_1 + \mathbf{E}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{W} \\ -\mathbf{Z} \end{pmatrix} \qquad (4)$$

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

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

k_1 is the rank of $\Delta + D\Sigma D^T$ where, if $G(\cdot)$, is the distribution function of $g(\cdot)$

$$dG_{\mathbf{W}|\{\mathbf{Z} \geq 0\}}(\mathbf{w}|\mathbf{Z} \geq 0) = \frac{dG_{\mathbf{W}}(\mathbf{w})}{P(\mathbf{Z} \geq 0)} P(\mathbf{Z} \geq 0 | \mathbf{W} = \mathbf{w}) \qquad (5)$$

with
$$\mathbf{W} \sim \mathcal{E}_p^r \left(\mu, \Sigma, h_r^{(p)} \right) \quad \text{and} \quad \mathbf{Z} \sim \mathcal{E}_q^{k_1} \left(-\nu, \Delta + D\Sigma D^T, h_{k_1}^{(q)} \right),$$
whose density for a random vector s-dimensional, will be generically denoted by, $g_{\mathbf{V}}^{(s)} \left(\mathbf{v}; r, \mu, \Sigma, h_r^{(s)} \right)$ o $dG_{\mathbf{V}}^{(s)} \left(\mathbf{v}; r, \mu, \Sigma, h_r^{(s)} \right)$, being r the rank of the distribution (which is defined as the rank of the matrix Σ, see Cramér (1999, p. 297)). Then, given $\Sigma = \Sigma\Sigma^-\Sigma$, there is

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

with $\delta(\mathbf{w}) = (\mathbf{w} - \mu)^T \Sigma^-(\mathbf{w} - \mu)$, see Theorem 2.6.4, pp. 62-65 in Gupta &

Varga (1993); where

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

with $h(\cdot,\cdot)$ a decreasing function, $h : \mathbb{R}^+ \to \mathbb{R}^+$, such that

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

Then,

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

and

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

This way, the density (5) can be expressed as

$$dG_{\mathbf{W}|\{\mathbf{z}\geq\mathbf{0}\}}(\mathbf{w}|\mathbf{Z} \geq \mathbf{0}) =$$

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

in conclusion,

Definition 3 (Singular extended skew-elliptical distribution) *A random vector* $\mathbf{Y}$ *has a singular extended skew-elliptical distribution p-dimensional, with rank r and parameters* $q, k_1, \mu, \Sigma, k, D, \nu, \Delta$, *as defined, if its density function is given by*

$$dG_{\mathbf{Y}}^{(p)}\left(\mathbf{y}; r, q, k_1, \mu, \Sigma, k, D, \nu, \Delta, h_r^{(p)}\right) =$$

$$\frac{F_{\mathbf{Y}}^{(q)}\left(D\Sigma\Sigma^-(\mathbf{y} - \mu); k, \nu, \Delta, h_{\delta(\mathbf{w}),\ k}^{(q)}\right)}{F_{\mathbf{Y}}^{(q)}\left(0; k_1, \nu, \Delta + D\Sigma D^T, h_{k_1}^{(q)}\right)}\ dG_{\mathbf{Y}}^{(p)}\left(\mathbf{y}; r, \mu, \Sigma, h_r^{(p)}\right) \qquad (6)$$

indicating this by

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

Specific cases of this family are present when: a). $\Delta > 0$, then $k = q = k_1$ where the parameters k and k_1 are excluded in the density (6), b). Si $\Sigma > 0$, then $r = p$ and the parameter r is excluded in the definition 3, c). if $\Delta > 0$

and $\Sigma > 0$ is obtained the non-singular extended skew-elliptical distribution, in such case the parameters r, k, k_1 in the definition 3 are excluded, see González-Farías et al. (2004). Finally, observe that as a consequence of the non uniqueness of the elliptic singular distribution, the distribution $\mathcal{SESE}$ is not unique either, see Remark 2.

4   Matrix-variate singular skew-elliptical distribution

In this section we study, the singular skew-elliptical distribution for the matrix case. As it will be noted, the matrix distribution will be obtained as an extension of the vector distribution described in section 3. This section ends up with a proposal to utilize the extension as a generalization for the matrix version.

First of all, we know that $Y \sim \mathcal{E}_{N \times m}(\mu, \Theta \otimes \Xi, h)$ is be equivalent to $\text{vec}\, Y \sim \mathcal{E}_{Nm}(\text{vec}\, \mu, \Theta \otimes \Xi, h)$, see Muirhead (1982, p. 79) and Gupta & Varga (1993, pp. 26-27). Then assuming

$$\begin{pmatrix} \text{vec}\, E_1 \\ \text{vec}\, E_2 \end{pmatrix} \sim \mathcal{E}_{pmqn}^{r_\Sigma r_\Theta r_\Delta r_\Xi} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Theta \otimes \Sigma & 0 \\ 0 & \Xi \otimes \Delta \end{pmatrix}, h_{r_\Sigma r_\Theta r_\Delta r_\Xi}^{(pmqn)} \right) \tag{7}$$

where $E_1 : p \times m$ and $E_2 : q \times n$ are matrices; $\Sigma : p \times p$ of rank $r_\Sigma \leq p$, $\Sigma \geq 0$; $\Theta : m \times m$ of rank $r_\Theta \leq m$, $\Theta \geq 0$; $\Delta : q \times q$ of rank $r_\Delta \leq q$, $\Delta \geq 0$ and $\Xi : n \times n$ of rank $r_\Xi \leq n$, $\Xi \geq 0$. Then the matrix version of the model (4) is given by

$$\text{vec}\, U = \begin{pmatrix} I & 0 \\ (D_2^T \otimes D_1) & I \end{pmatrix} \begin{pmatrix} \text{vec}\, E_1 \\ \text{vec}\, E_2 \end{pmatrix} + \begin{pmatrix} \text{vec}\, \mu \\ -\text{vec}\, \nu \end{pmatrix} \tag{8}$$

where $D_1 : q \times p$; $D_2 : m \times n$; $\mu : p \times m$ and $\nu : q \times n$ are arbitrary matrices of constants. Explicitly, there is

$$\text{vec}\, U = \begin{pmatrix} \text{vec}\, W \\ \text{vec}\, Z \end{pmatrix} = \begin{pmatrix} \text{vec}\, \mu + \text{vec}\, E_1 \\ -\text{vec}\, \nu + (D_2^T \otimes D_1) \text{vec}\, E_1 + \text{vec}\, E_2 \end{pmatrix}.$$

Doing a parallel development to the one presented in section 3 after the model (4), there is the following:

Definition 4 (Matrix-variate singular extended skew-elliptical) *It is said that a random matrix Y has a matrix-variate singular extended skew-elliptical*

distribution $p \times m$- dimensional, of rank $r_\Sigma r_\Theta$ and parameters q, n, k_1, μ, Σ, Θ, $k = r_\Delta r_\Xi$, D_1, D_2, ν, Δ, Ξ, defined, if its density function is given by

$$dG_{\text{vec} Y}^{(pm)}\left(\text{vec } Y; r_\Sigma r_\Theta, q, n, k_1, \text{vec } \mu, \Theta \otimes \Sigma, k, D_2^T \otimes D_1, \text{vec } \nu, \Xi \otimes \Delta, h_{r_\Sigma r_\Theta}^{(pm)}\right) =$$

$$\frac{F_{\text{vec} Y}^{(qn)}\left(D_2^T \Theta \Theta^- \otimes D_1 \Sigma \Sigma^- \text{vec}(Y - \mu); k, \text{vec } \nu, \Xi \otimes \Delta, h_{\delta(W),\ k}^{(qn)}\right)}{F_{\text{vec} Y}^{(qn)}\left(0; k_1, \text{vec } \nu, \Xi \otimes \Delta + D_2^T \Theta D_2 \otimes D_1 \Sigma D_1^T, h_{k_1}^{(qn)}\right)}$$

$$dG_{\text{vec} Y}^{(pm)}\left(\text{vec } Y; r_\Sigma r_\Theta, \text{vec } \mu, \Theta \otimes \Sigma, h_{r_\Sigma r_\Theta}^{(pm)}\right)$$

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

$$dG_Y^{(p \times m)}\left(Y; r_\Sigma r_\Theta, q, n, k_1, \mu, \Theta \otimes \Sigma, k, D_2^T \otimes D_1, \nu, \Xi \otimes \Delta, h_{r_\Sigma r_\Theta}^{(pm)}\right) =$$

$$\frac{F_Y^{(q \times n)}\left(D_1 \Sigma \Sigma^-(Y - \mu)\Theta^- \Theta D_2; k, \nu, \Xi \otimes \Delta, h_{\delta(W),\ k}^{(q \times n)}\right)}{F_Y^{(q \times n)}\left(0; k_1, \nu, \Xi \otimes \Delta + D_2^T \Theta D_2 \otimes D_1 \Sigma D_1^T, h_{k_1}^{(q \times n)}\right)}$$

$$dG_Y^{(p \times m)}\left(Y; r_\Sigma r_\Theta, \mu, \Theta \otimes \Sigma, h_{r_\Sigma r_\Theta}^{(p \times m)}\right)$$

where $\delta(W) = \text{tr }\Sigma^-(W - \mu)^T \Theta^-(W - \mu)$; k_1 is the rank of $(\Xi \otimes \Delta + D_2^T \Theta D_2 \otimes D_1 \Sigma D_1^T)$. And this fact will be indicated by

$$Y \sim \mathcal{SESE}_{r_\Sigma r_\Theta}^{(p \times m)}\left(q, n, k_1, \mu, \Theta \otimes \Sigma, k, D_2^T \otimes D_1, \nu, \Xi \otimes \Delta, h_{r_\Sigma r_\Theta}^{(p \times m)}\right)$$

Now, the structure of the covariance matrix, of the matrix E_1 (as example) through the Kronecker product, is a consequence of the linear transformations over a matrix. For example, by the same context of the definition 4, if

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

where 0 is a matrix of zeros of order $r_\Sigma \times r_\Theta$. Then $E_1 = MVN$, with $\Sigma = MM^T$ and $\Theta = N^T N$ is such that

$$E_1 \sim \mathcal{E}_{p \times m}^{r_\Sigma r_\Theta}\left(0, \Theta \otimes \Sigma, h_{r_\Sigma r_\Theta}^{(N \times m)}\right)$$

The disadvantage of this approach - that make the transformations over a matrix - is that the elements of the covariances matrix, $\Theta \otimes \Sigma$, have certain restrictions. From a Bayesian point of view this is of the utmost importance, because to propose an apriori distribution of this parameter, the restrictions have to be taken into account, and this increases the difficulty in the estimation. see Press (1982, p. 253).

An alternative approach is to make the linear transformation over the vectorization of the matrix without considering a structure in the linear transformation as function of the Kronecker product. For the example, we have to start from the fact that:

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

Then define $\operatorname{vec} E_1 = A \operatorname{vec} V$, with $A : pm \times r_\Sigma r_\Theta$ such that $\Lambda = AA'$, then

$$\operatorname{vec} E_1 \sim \mathcal{E}_{pm}^{r_\Sigma r_\Theta} \left(\operatorname{vec} 0, \Lambda, h_{r_\Sigma r_\Theta}^{(N \times m)} \right).$$

Taking into consideration this observation, alternatively to (7) we have

$$\begin{pmatrix} \operatorname{vec} E_1 \\ \operatorname{vec} E_2 \end{pmatrix} \sim \mathcal{E}_{pmqn}^{r_\Lambda r_\Omega} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \Lambda & 0 \\ 0 & \Omega \end{pmatrix}, h_{r_\Lambda r_\Omega}^{(pmqn)} \right),$$

where $\Lambda : pm \times pm$ of rank $r_\Lambda \leq pm$, $\Lambda \geq 0$ and $\Omega : qn \times qn$ of rank $r_\Omega \leq qn$, $\Omega \geq 0$. An alternative model (8) is defined as

$$\operatorname{vec} U = \begin{pmatrix} I & 0 \\ \mathcal{D} & I \end{pmatrix} \begin{pmatrix} \operatorname{vec} E_1 \\ \operatorname{vec} E_2 \end{pmatrix} + \begin{pmatrix} \operatorname{vec} \mu \\ -\operatorname{vec} \nu \end{pmatrix}$$

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

$$\operatorname{vec} U = \begin{pmatrix} \operatorname{vec} W \\ \operatorname{vec} Z \end{pmatrix} = \begin{pmatrix} \operatorname{vec} \mu + \operatorname{vec} E_1 \\ -\operatorname{vec} \nu + \mathcal{D} \operatorname{vec} E_1 + \operatorname{vec} E_2 \end{pmatrix}.$$

Consequently, we obtained the following general definition 4

Definition 5 (Matrix variate singular extended skew-elliptical II) *A random matrix Y has a matrix variate singular extended skew-elliptical distribution pm- dimensional, with rank r_Λ and parameters q, n, k_1, μ, Λ, k, $\mathcal{D}$, ν, Ω, defined, if its density function is given by*

$$dG_{\operatorname{vec} Y}^{(pm)} \left(\operatorname{vec} Y ; r_\Lambda, q, k_1, \mu, \Lambda, k, \mathcal{D}, \nu, \Omega, h_{r_\Lambda}^{(pm)} \right) =$$

$$\frac{F_{\operatorname{vec} Y}^{(qn)} \left(\mathcal{D} \Lambda \Lambda^- \operatorname{vec}(Y - \mu); k, \operatorname{vec} \nu, \Omega, h_{\delta(\operatorname{vec} W),\, k}^{(qn)} \right)}{F_{\operatorname{vec} Y}^{(qn)} \left(0; k_1, \operatorname{vec} \nu, \Omega + \mathcal{D} \Lambda \mathcal{D}^T, h_{k_1}^{(qn)} \right)}$$

$$dG_{\operatorname{vec} Y}^{(pm)} \left(\operatorname{vec} Y ; r_\Lambda, \operatorname{vec} \mu, \Lambda, h_{r_\Lambda}^{(pm)} \right)$$

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

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

5 General linear transformation

Now, our goal is to find the distribution of the general linear transformation, $A\mathbf{Y} + \mathbf{b}$ when $\mathbf{Y} \sim \mathcal{SESE}_r^{(p)}\left(q, k_1, \mu, \Sigma, k, D, \nu, \Delta, h_r^{(p)}\right)$, in which $\mathbf{b}$ is a constant vector and A is any constant matrix, and from this illustrate its application deriving the distribution of the residuals in the case of a multivariate linear model.

Theorem 6 *Assuming that* $\mathbf{Y} \sim \mathcal{SESE}_r^{(p)}\left(q, k_1, \mu, \Sigma, k, D, \nu, \Delta, h_r^{(p)}\right)$ *and let A be a matrix of constants $s \times p$ of rank $s_1 \leq \min(s, p)$ and $\mathbf{b}$ a constant vector $s \times 1$, then if $\mathbf{a}_j \in \mathcal{I}m(A\Sigma A^T)$ for all $j = 1, \ldots, q$, where $\mathbf{a}_j$ are the columns of the matrix $A\Sigma D^T$ and $\mathcal{I}m(N)$ denote the image of the matrix N, we get,*

$$A\mathbf{Y} + \mathbf{b} \sim \mathcal{SESE}_{s_2}^{(s)}\left(q, k_1, \mathbf{A}\mu + \mathbf{b}, \Sigma_A, k_2, D_A, \nu, \Delta_A, h_{s_2}^{(s)}\right)$$

where

$$\Sigma_A = A\Sigma A^T$$

$$D_A = D\Sigma A^T \Sigma_A^-$$

$$\Delta_A = \Delta + D(\Sigma - \Sigma A^T \Sigma_A^- A\Sigma)D^T$$

s_2 *is the rank of* $(A\Sigma A^T)$

k_1 *is the rank of* $(\Delta_A + D_A \Sigma_A D_A^T)(= \text{ to rank of } (\Delta + D\Sigma D^T))$

k_2 *is the rank of* Δ_A

Proof. Define $\mathbf{V} = A\mathbf{U} + \mathbf{b}_1$, with

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

where

$$E(\mathbf{V}) = \begin{pmatrix} A\mu + \mathbf{b} \\ -\nu \end{pmatrix} \text{ and } \text{Cov}(\mathbf{V}) = B\,\text{Cov}(\mathbf{V})B^T = \begin{pmatrix} A\Sigma A^T & A\Sigma D^T \\ D\Sigma A^T & \Delta + D\Sigma D^T \end{pmatrix}.$$

Then

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

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

$$\begin{pmatrix} A\Sigma A^T & A\Sigma D^T \\ D\Sigma A^T & \Delta + D\Sigma D^T \end{pmatrix} = \begin{pmatrix} \Sigma_A & \Sigma_A D_A^T \\ D_A \Sigma_A & \Delta_A + D_A \Sigma_A D_A^T \end{pmatrix}.$$

Where $\Sigma_A = A\Sigma A^T$ and $D_A = D\Sigma A^T \Sigma_A^-$, of which

$$\Sigma_A D_A^T = \Sigma_A \Sigma_A^- A\Sigma D^T = A\Sigma D^T$$

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

$$D_A \Sigma_A D_A^T = D\Sigma A^T \Sigma_A^- \Sigma_A \Sigma_A^- A\Sigma D^T = D\Sigma A^T \Sigma_A^- A\Sigma D^T.$$

From here we get $\Delta_A + D_A \Sigma_A D_A^T = \Delta + D\Sigma D^T$. Finally $\mathbf{Y} \stackrel{d}{=} \mathbf{W}|\{\mathbf{Z} \geq \mathbf{0}\}$, then $A\mathbf{Y} + \mathbf{b} \stackrel{d}{=} A\mathbf{W} + \mathbf{b}|\{\mathbf{Z} \geq \mathbf{0}\}$, where $\stackrel{d}{=}$ indicates equal in distribution. Proceeding in a parallel form to (5) the expected result is obtained. ∎

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

$$A\mathbf{Y} + \mathbf{b} \sim \mathcal{ESE}_{s,\, q}\left(\mathbf{A}\mu + \mathbf{b}, \Sigma_A, D_A, \nu, \Delta_A, h\right).$$

Similar results to those presented in theorem 6 can be shown for the matrix case, based on the definitions 4 and 5.

Corollary 7 *Consider the general multivariate linear model* $Y = X\beta + \xi$ *where* $Y : N \times m$, $X : N \times l$, *of rank* $\tau \leq l \leq N$, $\beta : l \times m$ *and*

$$\xi \sim \mathcal{SESE}_{r_\Sigma N}^{(N \times m)}\left(q, n, k_1, 0, I_N \otimes \Sigma, k, D_2^T \otimes D_1, \nu, \Xi \otimes \Delta, h_{r_\Sigma N}^{(N \times m)}\right).$$

If $R : N \times m$ *denotes the residual matrix. Then*

$$R \sim \mathcal{SESE}_{s_2 N}^{(N \times m)}\left(q, n, k_1, 0, (I_N \otimes \Sigma)_A, k, (D_2^T \otimes D_1)_A, \nu, (\Xi \otimes \Delta)_A, h_{s_2 N}^{(N \times m)}\right),$$

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

$(I_N \otimes \Sigma)_A = (I_N \otimes P\Sigma P)$,

s_2 *is the rank of* $(I_N \otimes \Sigma)_A$,

$(D_2^T \otimes D_1)_A = (D_2^T \otimes D_1 \Sigma P(P\Sigma P)^-)$ *and*

$(\Xi \otimes \Delta)_A = \Xi \otimes \Delta + (D_2^T \otimes D_1)(I_N \otimes \Sigma - I_N \otimes \Sigma P(P\Sigma P)^- P\Sigma)(D_2 \otimes D_1^T)$.

Proof. Keeping in mind that $R = Y - \widehat{Y} = Y - X\widehat{\beta} = Y - XX^+Y = (I - XX^+)Y = PY$, where $P = (I - XX^+)$ and $\widehat{\beta}$ is any solution of the system of normal matrix equations $(X^TX)\widehat{\beta} = X^TY$, see Rao (1973) or Muirhead (1982). Now the model $Y = X\beta + \xi$ is a linear transformation of the matrix ξ, by the Theorem 6, considering the vectorization of the linear model $\operatorname{vec} Y = (I \otimes X)\operatorname{vec}\beta + \operatorname{vec}\xi$, there is

$$Y \sim \mathcal{SESE}_{r\Sigma N}^{(N\times m)}\left(q, n, k_1, X\beta, I_N \otimes \Sigma, k, D_2^T \otimes D_1, \nu, \Xi \otimes \Delta, h_{r\Sigma N}^{(N\times m)}\right).$$

To conclude, observe that now $R = PY$ is a linear transformation of the matrix Y, then applying the Theorem 6, observing that $\operatorname{vec} R = (I \otimes P)\operatorname{vec} Y$, the expected result is obtained. ∎

Many interesting applications maybe generated from this particular result especially in the sensitivity analysis as we mentioned before. Also it is clear from (6) that now we are free to use any linear transformation (no rank restrictions), and still be able to complete established its distribution providing a very general characterization for the family of matrix-variate extended skew-elliptical distributions.

Other important consequence comes from the fact that if the estimator for β is given as a linear applications, for example $\hat{\beta} = RY$, we are able to established the distribution for $\hat{\beta}$ even when the matrix R is singular or Y follows a matrix-variate singular extended skew-elliptical distribution.

6   Acknowledgment

References

D. J. Aigner, C. A. K. Lovell, and P. Scmidt, Formulation and estimation of stochastic frontier production function model, J. Econometric 12 (1977) 21-37.

B. C. Arnold and R. J. Beaver, Elliptical Models Subject to Hidden Truncation or Selective Sampling, In: Skew - Elliptical Distributions and Their Applications: A Journey Beyond Normality, Chapter 6, Edited by Marc G. Genton, Chapman & Hall/CRC, USA, (2004) 101-112.

A. Azzalini, and A. Capitanio, Statistical applications of the multivariate skew normal distribution, J. R. Statist. Soc. B, 61 (1999) 579-602.

A. Azzalini, and A. Capitanio, Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution, J. R. Statist. Soc. B, 65 Part 2, (2003) 367-389.

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

P. Billingsley, Probability and Measure, 2nd Edition, John Wiley & Sons, New York, 1986.

M. Branco, and D. K. Dey, A general calss of multivariate skew elliptical distributions, J. Multivariate Anal. 79 (2001) 99-115.

S. Chatterjee, and A. S. Hadi, Sensitivity Analysis in Linear Regression, John Wiley & Sons. New York, 1988.

H. Cramér, Mathematical Methods of Statistics, Nineteenth printing, Princeton University Press, 1999.

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

J. A. Díaz-García and R. Gutiérrez-Jáimez, Proof of conjectures of H. Uhlig on the singular multivariate Beta and the Jacobian of a certain matrix transformation, Ann. of Statist. 25 (1997) 2018-2023.

J. A. Díaz-García and R. Gutiérrez-Jáimez, Wishart and Pseudo-Wishart distributions and some applications to shape theory II, Comunicación Técnica No. I-03-01 (PE/CIMAT) (2003), http://www.cimat.mx/biblioteca/RepTec.

J. A. Díaz-García and G. González-Farías, A Note on the Cook's Distance, J. Statist. Plann.Infer. 120 (2004a) 112-136.

J. A. Díaz-García and G. González-Farías, Singular Random Matrix decompositions: Jacobians, J. Multivariate Anal. (2004a). In press. Also see, http://www.scincedirect.com.

J. A. Díaz-García and G. González-Farías, Singular Random Matrix decompositions: Distributions, J. Multivariate Anal. (2004b). In press. Also see, http://www.scincedirect.com.

J. A. Díaz-García and R. Gutiérrez-Jáimez, Distribution of the generalised inverse of a random matrix and its applications, J. Statistical Plann. Inference. (2004). In press. Also see, http://www.scincedirect.com.

K. T. Fang, and Y. T. Zhang, "Generalized Multivariate Analysis," Science Press, Springer-Verlang, Beijing, 1990.

M. G. Genton, Skew - Elliptical Distributions and Their Applications: A Journey Beyond Normality, Editor, Chapman & Hall/CRC, USA, 2004.

G. González-Farías and A. Domínguez M. and A. Gupta, The closed skew normal distribution, In: Skew - Elliptical Distributions and Their Applications: A Journey Beyond Normality, Chapter 2, Edited by Marc G. Genton, Chapman & Hall/CRC, USA, (2004) 25-42.

A. K. Gupta & T. Varga, Elliptically Contoured Models in Statistics. Kluwer Academic Publishers, Dordrecht, 1993.

C. G. Khatri, Some results for the singular normal multivariate regression

models, Sankyā A 30 (1968) 267-280.

R. J. Muirhead, Aspects of multivariate statistical theory. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1982.

S. J. Press, Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference, Second Edition, Robert E. Krieger Publishing Company, Malabar, Florida, 1982.

C. R. Rao, Linear Statistical Inference and its Applications, 2nd Editon, John Wiley & Sons, New York, 1973.

S. K. Sahu, D. K. Dey, and M. D. Branco, A New Class of Multivariate Skew Distributions with Applications to Bayesian Regression Models, Canadian J. Statist. 31 (2003) 129-150.

H. Uhlig, On singular wishart and singular multivartiate beta distributions, Ann. Statistic. 22 (1994) 395-405.