JIRSS (2024)
Vol. 23, No. 02, pp 35-67
DOI: 10.22034/jirss.2025.2042155.1078

Multimatricvariate and Multimatrix Variate Distributions Based on Elliptically Contoured Laws under Real Normed Division Algebras

José A. Díaz-García[1], Francisco J. Caro-Lopera[2]

[1] Universidad Autónoma de Chihuahua, Facultad de Zootecnia y Ecología, Periférico Francisco R. Almada Km 1, Zootecnia, 33820 Chihuahua, Chihuahua, México.
[2] University of Medellin, Faculty of Basic Sciences, Carrera 87 No.30-65, Medellin, Colombia.

Received: 28/09/2024, Accepted: 18/02/2025, Published online: 08/06/2025

Abstract. This paper proposes famillies of multimatricvariate and multimatrix variate distributions based on elliptically contoured laws in the context of real normed division algebras. The work allows to answer the following inference problems about random matrix variate distributions: 1) Modeling of two or more probabilistically dependent random variables in all possible combinations whether univariate, vector and matrix simultaneously. 2) Expected marginal distributions under independence and joint estimation of models under likelihood functions of dependent samples. 3) Definition of a likelihood function for dependent samples in the mentioned random dimensions and under real normed division algebras. The corresponding real distributions are alternative approaches to the existing univariate and vector variate copulas, with the additional advantages previously listed. An application for quaternionic algebra is illustrated by a computable dependent sample joint distribution for landmark data emerged from shape theory.

Keywords. Multimatrix Variate, Real Normed Division Algebras, Matrix Variate, Multimatricvariate, Random Matrices, Matrix Variate Elliptical Distributions.

MSC: 60E05, 62E15, 15A23, 15B52.

Corresponding Author: José A. Díaz-García (jadiaz@cimat.mx)
Francisco J. Caro-Lopera (fjcaro@udemedellin.edu.co).

1 Introduction

A postmodern world of increasingly heuristic and interdisciplinary problems requires integrative solutions that do not always have the answer in the current frameworks of mathematics and statistics. Chaos theory as a paradigmatic element of phenomenal interrelation links more and more random variables. They are not only univariate and vector variables, but matrix variates, due to their great versatility in the description of variability and intrinsic multiple correlation. Likewise, the domain of the variables began to move from the real field to the remaining real normed division algebras, such as complex, quaternionic and octonionic. There the emerging models of multiple natural, exact and engineering sciences have their own scope and they are not restricted by the real field. Given multiple statistical models for marginal distributions in univariate and vector cases, the additional problem of parametric estimation arises from a set of samples provided by the scientific expert user of statistics. Then comes another challenge for statistics, which has historically been resolved by optimising likelihood functions as a joint law for sample distribution. The theory of copulas emerged as a possible solution to the problem of explaining a joint phenomenon of known interrelation. Countless link functions then appear for two variables, however, the intention of a dependent relationship is diluted when using the typical likelihood functions for estimating copula parameters from independent samples. The role of likelihood in the history of statistics is of such transcendence that it is a popular tool for estimating parameters. While copulas and similar theories represent the solution to the problem of dependence, the concept of likelihood function, defined as the product of the marginal densities, seems immutable and universal. Recently, the authors study the role of classical likelihood functions in statistics, showing the differences in application in databases declared dependently probabilistic. Time series in particular are the source of the greatest discrepancy, given the historical effort in proposing models for variance and volatility. But maintaining the postulate of an estimate via likelihood functions on independent data is an obvious contradiction in the face of a temporal process that intrinsically is founded on dependency. Recently, Díaz-García *et al.* (2022), proposed a new way of reframing likelihood functions. In a real data series, the discrepancy of likelihood function over independent time samples versus likelihood function over dependent samples was observed. The study indicated that decisions diffused about the mean estimated from the database using independent likelihood were in the tails of the dependent expected likelihood distribution. The finding was possible thanks to the definition of the so-termed multivector variate distribution, a natural way of defining a likelihood function of dependent vector samples, further parameterised by a large class of elliptically contoured distributions that are isolated from the choices of popular models. Until then, the mutivector could be an approach to vector copulas, but the implemented theory is designed to address the matrix case, which is still an open problem for copula theory. Then, the matrix version appeared in the form of the so-termed classes of multimatricvariate distributions, which provided the likelihood function or joint matrix distributions with a family of elliptically contoured distributions with a variety of kurtosis and symmetries, see Díaz-García and Caro-Lopera (2022). The choice of distributions that are invariant under the class of elliptically contoured laws within the family of multimatricvariate and multimatrix variate distributions gives robustness to the analysis since it avoids the adjusting the link function, if we speak in parallel to the procedure. in vector copulas.

Multimatricvariate distributions specialise in distributions based on the determinant. Still to be defined a class of multiple distributions of random matrices that depended on the trace, the other function that has historically governed the theory of matrix distributions. Recently, the so-termeded multimatrix distributions appear in Díaz-García and Caro-Lopera (2024). Like the multimatricvariate distributions, they maintain the philosophy of probabilistic dependence, computability and a wide class of underlying distributions.

When we surpass the level of random vectors, where the most popular statistical techniques remain, such as copulas, we find immense difficulty for the calculation. We must observe the domain of the matrices, their transformations and Jacobians, and therefore their integration. It is generally difficult because it involves averaging over orthogonal groups, cones and hypercubes. The central, isotropic, non-central and non-isotropic cases also emerge. These integrations comprises the geometric filtering that comes for their definitions and factorisations.

At this point, we find that the new multimatricvariate and multimatrix variate distributions on the real field satisfy the first 2 of the 3 conditions that we establish for a robust point of view of joint distributions, namely: 1) modeling of several dependent matrices capable of being combined with all univariate and/or vector variables unde a elliptically contoured models; 2) with strictly verifiable marginals under the trivial case of independence and testable from the dependent sample likelihood; and 3) founded on the same algebraic principles that allow them to be applied indistinctly from the real normed division algebras.

We now address the third desirable characteristic: its versatility in the four unique real normed division algebras.

Until a couple of decades ago it would be unthinkable to unify the theory of real random matrices with complex ones. History shows us that the theory of distributions of the central cases overflowed into an immense effort to characterize the indispensable elements of random matrices, namely: After its establishment, works on the complex case began to appear very slowly and completely conceptually distant from their real analogues. The so-termed statistical theory under real normed division algebras is to statistics what the expected theory of unification of forces is to physics. Its simplification is simple, reducing to modifying a beta parameter that goes from 1 (real) to 2 (complex) to 4 (quaternionic) to 8 (octonionic), see Díaz-García (2014) were the transitions to complex, quaternionic and octonion are reached by changing the support group from orthogonal to unitary, compact symplectic or exceptional type.

If the real multimatricvariate distributions (Díaz-García and Caro-Lopera, 2022) and multimatrix variate distributions (Díaz-García and Caro-Lopera, 2024) articles are parallel compared with the results of the present work, an apparent repetition shall appear by its simplicity; Its notation makes the beta parameter open one of the four algebras by simply changing the value, although understanding the profound difference that underpins them requires an extensive and difficult literature that began in other areas far from statistics. The non commutativity for quaternions and non associative for octonions promotes a deep research for some applications in those algebras.

We place the above discussion into the setting of two main problems that we shall address in this article.

The interest in multimatricvariate and multimatrix variate distributions has been motivated by the following two situations:

1. In different areas of knowledge (such as Finance and Hydrology, among others), people are interested in simultaneously modeling two random variables, say X and Y, which are suspected of not being probabilistically independent. On the one hand, the marginal distributions of each variable are known, whether they are $f_X(x)$ and $g_Y(y)$. Typically this problem has been approached assuming that the random variables X and Y are independent and, as a function of joint density of the two-dimensional vector $(X, Y)'$, the product of the marginals, $r_{X,Y}(x, y) = f_X(x)g_Y(y)$, has been considered. Thus, for example, in this case the likelihood function, given the two-dimensional sample $(x_1, y_1), \cdots, (x_k, y_k)$, denoted as $L(\theta; (x_1, y_1), \cdots, (x_k, y_k))$ is defined as

$$\begin{aligned} L(\theta; (x_1, y_1), \cdots, (x_n, y_k)) &= \prod_{j=1}^{k} r_{X_j,Y_j}(x_j, y_j) \\ &= \prod_{j=1}^{k} f_{X_j}(x_j) g_{Y_j}(y_j), \end{aligned}$$

   For some parameter vector $\theta \in \Re^p$ which is part of the density $r_{X,Y}(x, y)$.

   Alternatively, based on variable changes on a set of independent random variables, the random vector $(X, Y)'$ was generated, where now the variables X and Y are not independent and their density function is known joint $t_{X,Y}(x, y) \neq f_X(x)g_Y(y)$, such that

$$f_X(x) = \int_{\Re} t_{X,Y}(x, y)(dy) \quad \text{and} \quad g_Y(y) = \int_{\Re} t_{X,Y}(x, y)(dx).$$

   Under this approach we have that

$$L(\theta; (x_1, y_1), \cdots, (x_k, y_k)) = \prod_{j=1}^{k} t_{X_j,Y_j}(x_j, y_j),$$

See Libby and Novick (1982), Chen and Novick (1984), Olkin and Liu (2003), Nadarajah (2007, 2013) and Sarabia *et al.* (2014) among many others.

This situation also occurs in other multivariate problems. Then, parallel solutions were proposed in the vector and matrix cases, giving foothold to the study of bi-matrix variate distributions in the real and complex cases. In the last case joint distributions of random matrices, say $\mathbf{X}$ and $\mathbf{Y}$ dependent with joint density function, termed bimatrix variate distribution, are proposed such that the marginal densities of $\mathbf{X}$ and $\mathbf{Y}$ are the usual assumptions, see Olkin and Rubin (1964), Díaz-García, and Gutiérrez-Jáimez (2010a,b, 2011), Bekker *et al.* (2011), and Ehlers (2011) and references therein.

2. In another case, we are interested in defining the likelihood function by a joint function of the sample, but which is not defined as the product of the marginals, that is, the sample is not independent. In the univariate problem, one answer considers the elliptically contoured distribution of the vector $(X_1, \ldots, X_k)'$ as a likelihood function, noting that in reality the elliptically contoured distribution actually defines a distribution family, see Fang *et al.* (1990), Fang and Zhang (1990), Gupta *et al.* (2013) and the reference therein.

   Based on the family of matrix variate elliptically contoured distributions, the multimatrix variate and multimatricvariate distributions were proposed as a generalisation of the bi-matrix variate distributions, which are defined as the joint distribution of the dependent random matrices $\mathbf{X}_1, \ldots, \mathbf{X}_k$, see Díaz-García *et al.* (2022), and Díaz-García and Caro-Lopera (2022, 2024). Thus, the multimatrix variate and multimatricvariate distributions can be used as likelihood functions for a sample of dependent random matrices with certain (usual) marginal distributions. Thus, the likelihood function of the sample $\mathbf{X}_1, \ldots, \mathbf{X}_k$ is defined as

$$L(\mathbf{\Theta}; \mathbf{X}_1, \ldots, \mathbf{X}_k) = f_{\mathbf{X}_1,\ldots,\mathbf{X}_k}(\mathbf{X}_1, \ldots, \mathbf{X}_k).$$

Under the theory of multimatrix matrix or multimatricvariate distributions, each matrix $\mathbf{X}_j$, $j = 1, \ldots, k$ into the density $f_{\mathbf{X}_1,\ldots,\mathbf{X}_k}(\mathbf{X}_1, \ldots, \mathbf{X}_k)$ can follow a different marginal distribution. This answers the following problem under an independent or dependent samples: Suppose that we have a matrix random sample as follows:

$\mathbf{X}_1$	$\mathbf{X}_{11}$	$\mathbf{X}_{12}$	$\cdots$	$\mathbf{X}_{1r}$
$\mathbf{X}_2$	$\mathbf{X}_{21}$	$\mathbf{X}_{22}$	$\cdots$	$\mathbf{X}_{2r}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\mathbf{X}_k$	$\mathbf{X}_{k1}$	$\mathbf{X}_{k2}$	$\cdots$	$\mathbf{X}_{kr}$

And assume that the matrix $\mathbf{\Theta}$ contains the parameters of interest. Then the likelihood function $L(\mathbf{\Theta};\cdot)$ can be defined as:

$$\begin{cases} \displaystyle\prod_{j=1}^{r} f_{\mathbf{X}_{1j},\dots,\mathbf{X}_{kj}}(\mathbf{X}_{1j},\dots,\mathbf{X}_{kj}), & \text{independence case,} \\ f_{\mathbf{X}_{11},\dots,\mathbf{X}_{1r},\dots,\mathbf{X}_{k1},\dots,\mathbf{X}_{kr}}(\mathbf{X}_{11},\dots,\mathbf{X}_{1r},\dots,\mathbf{X}_{k1},\dots,\mathbf{X}_{kr}), & \text{dependence case.} \end{cases}$$

In the bi-matrix variate case, these problems have been studied in the real and complex cases, each giving rise to a series of non correlated publications. Fortunately, in terms of the theory of the real normed division algebras, a unification of the real and complex cases is possible. And an extension to the quaternionic and octonionic algebras is also feasible. It is worth mentioning that the octornionic case is still under research. At this time, they are valid for 2×2 octornionic matrices and in general it can only be conjectured that they may be valid.

In the present work, the multimatrix variate and multimatricvariate distributions are studied for matrix arguments which elements belong to the real normed division algebras. A brief description of the notation and some Jacobians for real normed division algebras is presented in Section 2. In addition, two more Jacobians are obtained and the definition of the matrix variate elliptically contoured distribution for real normed division algebras is presented. The main results on multimatrix variate and multimatricvariate distributions for real normed division algebras are obtained in Section 3. Some properties and extensions of multimatrix variate and multimatricvariate distribution with more than two different types of distributions in their arguments are studied in Section 4. An example in the quaternionic case is full derived in Section 5.

2 Notation and Preliminary Results

A detailed discussion of real normed division algebras may be found in Baez (2002). For convenience, we shall introduce some notations, although in general we adhere to standard notations.

A **vector space** is always a finite-dimensional module over the field of real numbers. An **algebra** $\mathfrak{F}$ is a vector space that is equipped with a bilinear map $m : \mathfrak{F}\times\mathfrak{F} \to \mathfrak{F}$ termed *multiplication* and a nonzero element $1 \in \mathfrak{F}$ termed the *unit* such that $m(1, a) = m(a, 1) = 1$. As usual, we abbreviate $m(a, b) = ab$ as ab. We do not assume $\mathfrak{F}$ associative. Given an algebra, we freely think of real numbers as elements of this algebra via the map $\omega \mapsto \omega 1$.

An algebra $\mathfrak{F}$ is a **division algebra** if given $a, b \in \mathfrak{F}$ with $ab = 0$, then either $a = 0$ or $b = 0$. Equivalently, $\mathfrak{F}$ is a division algebra if the operation of left and right multiplications by any nonzero element is invertible. A **normed division algebra** is an algebra $\mathfrak{F}$ that is also a normed vector space with $||ab|| = ||a||||b||$. This implies that $\mathfrak{F}$ is a division algebra and that $||1|| = 1$.

There are exactly four normed division algebras: real numbers ($\mathfrak{R}$), complex numbers ($\mathfrak{C}$), quaternions ($\mathfrak{H}$) and octonions ($\mathfrak{O}$), see Baez (2002). Taking into account that

$\mathfrak{R}$, $\mathfrak{C}$, $\mathfrak{H}$ and $\mathfrak{O}$ are the only normed division algebras; moreover, they are the only alternative division algebras, and all division algebras have a real dimension of $1, 2, 4$ or 8, which is denoted by β, see Baez (2002, Theorems 1, 2 and 3). In other branches of mathematics, the parameter $\alpha = 2/\beta$ is used, see Edelman and Rao (2005).

Let $\mathcal{L}^{\beta}_{m,n}$ be the linear space of all $n \times m$ matrices of rank $m \leq n$ over $\mathfrak{F}$ with m distinct positive singular values, where $\mathfrak{F}$ denotes a *real finite-dimensional normed division algebra*. Let $\mathfrak{F}^{n\times m}$ be the set of all $n \times m$ matrices over $\mathfrak{F}$. The dimension of $\mathfrak{F}^{n\times m}$ over $\mathfrak{R}$ is βmn. Let $\mathbf{A} \in \mathfrak{F}^{n\times m}$, then $\mathbf{A}^H = \overline{\mathbf{A}}^T$ denotes the usual conjugate transpose.

The set of matrices $\mathbf{H}_1 \in \mathfrak{F}^{n\times m}$ such that $\mathbf{H}_1^H\mathbf{H}_1 = \mathbf{I}_m$ is a manifold denoted $\mathcal{V}^{\beta}_{m,n}$, is termed the *Stiefel manifold* ($\mathbf{H}_1$ is also known as *semi-orthogonal* ($\beta = 1$), *semi-unitary* ($\beta = 2$), *semi-symplectic* ($\beta = 4$) and *semi-exceptional type* ($\beta = 8$) matrices, see Dray and Manogue (1999)). The dimension of $\mathcal{V}^{\beta}_{m,n}$ over $\mathfrak{R}$ is $[\beta mn - m(m-1)\beta/2 - m]$. In particular, $\mathcal{V}^{\beta}_{m,m}$ with dimension over $\mathfrak{R}$, $[m(m+1)\beta/2 - m]$, is the maximal compact subgroup $\mathfrak{U}^{\beta}(m)$ of $\mathcal{L}^{\beta}_{m,m}$ and consists of all matrices $\mathbf{H} \in \mathfrak{F}^{m\times m}$ such that $\mathbf{H}^H\mathbf{H} = \mathbf{I}_m$. Therefore, $\mathfrak{U}^{\beta}(m)$ is the *real orthogonal group* $\mathcal{O}(m)$ ($\beta = 1$), the *unitary group* $\mathcal{U}(m)$ ($\beta = 2$), *compact symplectic group* $\mathcal{S}p(m)$ ($\beta = 4$) or *exceptional type matrices* $\mathcal{O}o(m)$ ($\beta = 8$), for $\mathfrak{F} = \mathfrak{R}, \mathfrak{C}, \mathfrak{H}$ or $\mathfrak{O}$, respectively.

We denote by $\mathfrak{S}^{\beta}_m$ the real vector space of all $\mathbf{S} \in \mathfrak{F}^{m\times m}$ such that $\mathbf{S} = \mathbf{S}^H$. Let $\mathfrak{P}^{\beta}_m$ be the *cone of positive definite matrices* $\mathbf{S} \in \mathfrak{F}^{m\times m}$; then $\mathfrak{P}^{\beta}_m$ is an open subset of $\mathfrak{S}^{\beta}_m$. Over $\mathfrak{R}$, $\mathfrak{S}^{\beta}_m$ consist of *symmetric* matrices; over $\mathfrak{C}$, *Hermitian* matrices; over $\mathfrak{H}$, *quaternionic Hermitian* matrices (also termed *self-dual matrices*) and over $\mathfrak{O}$, *octonionic Hermitian* matrices. Generically, the elements of $\mathfrak{S}^{\beta}_m$ are termed **Hermitian matrices**, irrespective of the nature of $\mathfrak{F}$. The dimension of $\mathfrak{S}^{\beta}_m$ over $\mathfrak{R}$ is $[m(m-1)\beta + 2m]/2$.

Let $\mathfrak{D}^{\beta}_m$ be the *diagonal subgroup* of $\mathcal{L}^{\beta}_{m,m}$ consisting of all $\mathbf{D} \in \mathfrak{F}^{m\times m}$, $\mathbf{D} = \operatorname{diag}(d_1, \ldots, d_m)$.

For any matrix $\mathbf{X} \in \mathfrak{F}^{n\times m}$, $d\mathbf{X}$ denotes the *matrix of differentials* (dx_{ij}). Finally, we define the *measure* or volume element $(d\mathbf{X})$ when $\mathbf{X} \in \mathfrak{F}^{m\times n}, \mathfrak{S}^{\beta}_m$, $\mathfrak{D}^{\beta}_m$ or $\mathcal{V}^{\beta}_{m,n}$, see Dumitriu (2002).

If $\mathbf{X} \in \mathfrak{F}^{n\times m}$ then $(d\mathbf{X})$ (the Lebesgue measure in $\mathfrak{F}^{n\times m}$) denotes the exterior product of the βmn functionally independent variables

$$(d\mathbf{X}) = \bigwedge_{i=1}^{n}\bigwedge_{j=1}^{m} dx_{ij} \quad \text{where} \quad dx_{ij} = \bigwedge_{r=1}^{\beta} dx_{ij}^{(r)}.$$

If $\mathbf{S} \in \mathfrak{S}^{\beta}_m$ (or $\mathbf{S} \in \mathfrak{T}^{\beta}_L(m)$) then $(d\mathbf{S})$ (the Lebesgue measure in $\mathfrak{S}^{\beta}_m$ or in $\mathfrak{T}^{\beta}_L(m)$) denotes the exterior product of the $m(m+1)\beta/2$ functionally independent variables (or denotes the exterior product of the $m(m-1)\beta/2+m$ functionally independent variables, if $s_{ii} \in \mathfrak{R}$

for all $i = 1, \dots, m$)

$$(d\mathbf{S}) = \begin{cases} \displaystyle\bigwedge_{i\leq j}^{m} \bigwedge_{r=1}^{\beta} ds_{ij}^{(r)}, \\ \displaystyle\bigwedge_{i=1}^{m} ds_{ii} \bigwedge_{i<j}^{m} \bigwedge_{r=1}^{\beta} ds_{ij}^{(r)}, \quad \text{if } s_{ii} \in \Re. \end{cases}$$

The context generally establishes the conditions on the elements of $\mathbf{S}$, that is, if $s_{ij} \in \Re$, $\in \mathfrak{C}$, $\in \mathfrak{H}$ or $\in \mathfrak{O}$. It is considered that

$$(d\mathbf{S}) = \bigwedge_{i\leq j}^{m} \bigwedge_{r=1}^{\beta} ds_{ij}^{(r)} \equiv \bigwedge_{i=1}^{m} ds_{ii} \bigwedge_{i<j}^{m} \bigwedge_{r=1}^{\beta} ds_{ij}^{(r)}.$$

Observe, too, that for the Lebesgue measure $(d\mathbf{S})$ defined thus, it is required that $\mathbf{S} \in \mathfrak{P}_m^\beta$, that is, $\mathbf{S}$ must be a non singular Hermitian matrix (Hermitian positive definite matrix).

If $\mathbf{\Lambda} \in \mathfrak{D}_m^\beta$ then $(d\mathbf{\Lambda})$ (the Legesgue measure in $\mathfrak{D}_m^\beta$) denotes the exterior product of the βm functionally independent variables

$$(d\mathbf{\Lambda}) = \bigwedge_{i=1}^{n} \bigwedge_{r=1}^{\beta} d\lambda_i^{(r)}.$$

If $\mathbf{H}_1 \in \mathcal{V}_{m,n}^\beta$ then

$$(\mathbf{H}_1^H d\mathbf{H}_1) = \bigwedge_{i=1}^{n} \bigwedge_{j=i+1}^{m} \mathbf{h}_j^H d\mathbf{h}_i.$$

where $\mathbf{H} = (\mathbf{H}_1|\mathbf{H}_2) = (\mathbf{h}_1, \dots, \mathbf{h}_m|\mathbf{h}_{m+1}, \dots, \mathbf{h}_n) \in \mathfrak{U}^\beta(n)$. It can be proved that this differential form does not depend on the choice of the $\mathbf{H}_2$ matrix. When $m = 1$; $\mathcal{V}_{1,n}^\beta$ defines the unit sphere in $\mathfrak{F}^n$. This is, of course, an $(n-1)\beta$- dimensional surface in $\mathfrak{F}^n$. When $m = n$ and denoting $\mathbf{H}_1$ by $\mathbf{H}$, $(\mathbf{H}^H d\mathbf{H})$ is termed the *Haar measure* on $\mathfrak{U}^\beta(m)$.

The surface area or volume of the Stiefel manifold $\mathcal{V}_{m,n}^\beta$ is

$$\text{Vol}(\mathcal{V}_{m,n}^\beta) = \int_{\mathbf{H}_1 \in \mathcal{V}_{m,n}^\beta} (\mathbf{H}_1^H d\mathbf{H}_1) = \frac{2^m \pi^{mn\beta/2}}{\Gamma_m^\beta[n\beta/2]}, \tag{2.1}$$

where $\Gamma_m^\beta[a]$ denotes the multivariate Gamma function for the space $\mathfrak{S}_m^\beta$, and is defined by

$$\begin{aligned} \Gamma_m^\beta[a] &= \int_{\mathbf{A} \in \mathfrak{P}_m^\beta} \text{etr}\{-\mathbf{A}\} |\mathbf{A}|^{a-(m-1)\beta/2-1} (d\mathbf{A}) \\ &= \pi^{m(m-1)\beta/4} \prod_{i=1}^{m} \Gamma[a - (i-1)\beta/2], \end{aligned}$$

where $\text{etr}(\cdot) = \exp(\text{tr}(\cdot))$, $|\cdot|$ denotes the determinant and $\text{Re}(a) > (m-1)\beta/2$, see Gross and Richards (1987). If $\mathbf{A} \in \mathcal{L}^{\beta}_{m,n}$ then by $\text{vec}(\mathbf{A})$ we mean the $mn \times 1$ vector formed by stacking the columns of $\mathbf{A}$ under each other; that is, if $\mathbf{A} = [\mathbf{a}_1\mathbf{a}_2 \dots \mathbf{a}_m]$, where $\mathbf{a}_j \in \mathcal{L}^{\beta}_{1,n}$ for $j = 1, 2, \dots, m$

$$\text{vec}(\mathbf{A}) = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}.$$

Below are summarised some Jacobians in terms of the β parameter. For a detailed discussion of this and related issues see Dumitriu (2002), Edelman and Rao (2005), Forrester (2009) and Kabe (1984).

Proposition 2.1. *Let* $\mathbf{A} \in \mathcal{L}^{\beta}_{n,n}$, $\mathbf{B} \in \mathcal{L}^{\beta}_{m,m}$ *and* $\mathbf{C} \in \mathcal{L}^{\beta}_{m,n}$ *be matrices of constants,* $\mathbf{Y}$ *and* $\mathbf{X} \in \mathcal{L}^{\beta}_{m,n}$ *a matrices of functionally independent variables such that* $\mathbf{Y} = \mathbf{AXB} + \mathbf{C}$. *Then*

$$(d\mathbf{Y}) = |\mathbf{A}^H\mathbf{A}|^{\beta m/2}|\mathbf{B}^H\mathbf{B}|^{\beta n/2}(d\mathbf{X}). \tag{2.2}$$

Proposition 2.2. *Let* $\mathbf{S} \in \mathfrak{P}^{\beta}_{m}$. *If* $\mathbf{Y} = \mathbf{ASA}^H$, $\mathbf{A} \in \mathfrak{P}^{\beta}_{m}$.

$$(d\mathbf{Y}) = |\mathbf{A}^H\mathbf{A}|^{\beta(m-1)/2+1}(d\mathbf{X}). \tag{2.3}$$

Proposition 2.3 (Singular value decomposition, *SVD*)**.** *Assume that* $\mathbf{X} \in \mathcal{L}^{\beta}_{m,n}$, *such that* $\mathbf{X} = \mathbf{V}_1\mathbf{D}\mathbf{W}^H$ *with* $\mathbf{V}_1 \in \mathcal{V}^{\beta}_{m,n}$, $\mathbf{W} \in \mathfrak{U}^{\beta}(m)$ *and* $\mathbf{D} = \text{diag}(d_1, \cdots, d_m) \in \mathfrak{D}^1_m$, $d_1 > \cdots > d_m > 0$. *Then* $(d\mathbf{X})$ *is*

$$2^{-m}\pi^{\tau}\prod_{i=1}^{m} d_i^{\beta(n-m+1)-1}\prod_{i<j}^{m}(d_i^2 - d_j^2)^{\beta}(d\mathbf{D})(\mathbf{V}_1^H d\mathbf{V}_1)(\mathbf{W}^H d\mathbf{W}), \tag{2.4}$$

where

$$\tau = \begin{cases} 0, & \beta = 1, \\ -m, & \beta = 2, \\ -2m, & \beta = 4, \\ -4m, & \beta = 8. \end{cases}$$

As a consequence of this result, we have the following statement.

Proposition 2.4. *Let* $\mathbf{X} \in \mathcal{L}^{\beta}_{m,n}$, *and* $\mathbf{S} = \mathbf{X}^H\mathbf{X} \in \mathfrak{P}^{\beta}_{m}$. *Then*

$$(d\mathbf{X}) = 2^{-m}|\mathbf{S}|^{\beta(n-m+1)/2-1}(d\mathbf{S})(\mathbf{V}_1^H d\mathbf{V}_1), \tag{2.5}$$

with $\mathbf{V}_1 \in \mathcal{V}^{\beta}_{m,n}$.

Proposition 2.5. *Let* $\mathbf{S} \in \mathfrak{P}^{\beta}_{m}$. *Then ignoring the sign, if* $\mathbf{Y} = \mathbf{S}^{-1}$

$$(d\mathbf{Y}) = |\mathbf{S}|^{-\beta(m-1)-2}(d\mathbf{S}). \tag{2.6}$$

Theorem 2.1. *Assume that* $\mathbf{X} \in \mathcal{L}_{m,n}^{\beta}$ *and* $\mathbf{Y} \in \mathcal{L}_{m,n}^{\beta}$ *are matrices of functionally independent variables.*

i) *Define* $\mathbf{Y} = \mathbf{X}(\mathbf{I}_m - \mathbf{X}^H\mathbf{X})^{-1/2}$. *Then*

$$(d\mathbf{Y}) = \left|\mathbf{I}_m - \mathbf{X}^H\mathbf{X}\right|^{-\beta(n+m+1)/2-1}(d\mathbf{X}). \tag{2.7}$$

ii) *If* $\mathbf{X} = \mathbf{Y}(\mathbf{I}_m + \mathbf{Y}^H\mathbf{Y})^{-1/2}$, *we have*

$$(d\mathbf{X}) = \left|\mathbf{I}_m + \mathbf{Y}^H\mathbf{Y}\right|^{-\beta(n+m+1)/2-1}(d\mathbf{Y}), \tag{2.8}$$

where $n \geq m$.

Proof. **i**) Define $\mathbf{A} = \mathbf{Y}^H\mathbf{Y} = (\mathbf{I}_m - \mathbf{X}^H\mathbf{X})^{-1/2}\mathbf{X}^H\mathbf{X}(\mathbf{I}_m - \mathbf{X}^H\mathbf{X})^{-1/2}$ and $\mathbf{B} = \mathbf{X}^H\mathbf{X}$. And observe that

$$\begin{aligned}
\mathbf{A} &= (\mathbf{I}_m - \mathbf{X}^H\mathbf{X})^{-1/2}\mathbf{X}^H\mathbf{X}(\mathbf{I}_m - \mathbf{X}^H\mathbf{X})^{-1/2} \\
&= (\mathbf{I}_m - \mathbf{B})^{-1/2}\mathbf{B}(\mathbf{I}_m - \mathbf{B})^{-1/2} \\
&= \left[\mathbf{B}^{-1}(\mathbf{I}_m - \mathbf{B})\right]^{-1/2}\left[\mathbf{B}^{-1}(\mathbf{I}_m - \mathbf{B})\right]^{-1/2} \\
&= \left(\mathbf{B}^{-1} - \mathbf{I}_m\right)^{-1/2}\left(\mathbf{B}^{-1} - \mathbf{I}_m\right)^{-1/2} \\
&= \left(\mathbf{B}^{-1} - \mathbf{I}_m\right)^{-1} = (\mathbf{I}_m - \mathbf{B})^{-1}\mathbf{B} = (\mathbf{I}_m - \mathbf{B})^{-1} - \mathbf{I}_m.
\end{aligned}$$

Then by (2.5), for $\mathbf{H}_1, \mathbf{G}_1 \in \mathcal{V}_{m,n}$, and writing these for $(d\mathbf{A})$ and $(d\mathbf{B})$ we obtain

$$(d\mathbf{A}) = 2^m|\mathbf{A}|^{-\beta(n-m+1)/2+1}(d\mathbf{Y})(\mathbf{H}_1^H d\mathbf{H}_1)^{-1}, \tag{2.9}$$

$$(d\mathbf{B}) = 2^m|\mathbf{B}|^{-\beta(n-m+1)/2+1}(d\mathbf{X})(\mathbf{G}_1^H d\mathbf{G}_1)^{-1}. \tag{2.10}$$

Since $\mathbf{A} = (\mathbf{I}_m - \mathbf{B})^{-1} - \mathbf{I}_m$, from (2.6), we have that

$$(d\mathbf{A}) = |\mathbf{I}_m - \mathbf{B}|^{-\beta(m-1)-2}(d\mathbf{B}). \tag{2.11}$$

Substituting (2.10) and (2.9) into (2.11) we have

$$\begin{aligned}
&2^m|\mathbf{A}|^{-\beta(n-m+1)/2+1}(d\mathbf{Y})(\mathbf{H}_1^H d\mathbf{H}_1)^{-1} \\
&\qquad = 2^m|\mathbf{I}_m - \mathbf{B}|^{-(m+1)}|\mathbf{B}|^{-\beta(n-m+1)/2+1}(d\mathbf{X})(\mathbf{G}_1^H d\mathbf{G}_1)^{-1}.
\end{aligned}$$

Then, by the uniqueness of the nonnormalised measure on Stiefel manifold, $(\mathbf{H}_1^H d\mathbf{H}_1) = (\mathbf{G}_1^H d\mathbf{G}_1)$. Thus,

$$(d\mathbf{Y}) = |\mathbf{A}|^{\beta(n-m+1)/2+1}|\mathbf{I}_m - \mathbf{B}|^{-\beta(m-1)-2}|\mathbf{B}|^{-\beta(n-m+1)/2+1}(d\mathbf{X}).$$

And using

$$|\mathbf{A}| = |(\mathbf{I}_m - \mathbf{X}^H\mathbf{X})^{-1/2}\mathbf{X}^H\mathbf{X}(\mathbf{I}_m - \mathbf{X}^H\mathbf{X})^{-1/2}| = |(\mathbf{I}_m - \mathbf{X}^H\mathbf{X})|^{-1}|\mathbf{X}^H\mathbf{X}|,$$

and $|\mathbf{B}| = |\mathbf{X}^H\mathbf{X}|$, the required result is obtained.

ii). The proof is similar to the preceding exposition given in **i**). □

Corollary 2.1. Assume that $\mathbf{X} \in \mathcal{L}^{\beta}_{m,n}$ and $\mathbf{Y} \in \mathcal{L}^{\beta}_{m,n}$ are matrices of functionally independent variables.

i) Let $\mathbf{Y} = (1 - \operatorname{tr} \mathbf{X}^H\mathbf{X})^{-1/2}\mathbf{X}$. Then

$$(d\mathbf{Y}) = (1 - \operatorname{tr} \mathbf{X}^H\mathbf{X})^{-(\beta nm/2+1)}(d\mathbf{X}). \tag{2.12}$$

ii) If $\mathbf{X} = (1 + \operatorname{tr} \mathbf{Y}^H\mathbf{Y})^{-1/2}\mathbf{Y}$, we have

$$(d\mathbf{X}) = (1 + \operatorname{tr} \mathbf{Y}^H\mathbf{Y})^{-(\beta nm/2+1)}(d\mathbf{Y}). \tag{2.13}$$

Proof. **i**). Observing that $\mathbf{Y} = (1 - \operatorname{tr} \mathbf{X}^H\mathbf{X})^{-1/2}\mathbf{X}$ can be write as $\mathbf{y} = (1 - \mathbf{x}^H\mathbf{x})^{-1/2}\mathbf{x}$, where $\mathbf{x} = \operatorname{vec} \mathbf{X} \in \mathcal{L}^{\beta}_{1,nm}$ and $\mathbf{y} = \operatorname{vec} \mathbf{Y} \in \mathcal{L}^{\beta}_{1,nm}$. Its proof is obtained as a particular case to that given for the theorem. **ii**). Its proof is analogous to one given to **i**). □

Definition 2.1. It is said that the random matrix $\mathbf{Y} \in \mathcal{L}^{\beta}_{m,n}$ has a *matrix variate elliptical distribution*, denoted as $\mathbf{Y} \sim \mathcal{E}^{\beta}_{n\times m}(\boldsymbol{\mu}, \boldsymbol{\Theta}, \boldsymbol{\Sigma}, h)$, if its density is

$$\frac{1}{|\boldsymbol{\Sigma}|^{\beta n/2}|\boldsymbol{\Theta}|^{\beta m/2}} h\left\{\beta \operatorname{tr}\left[\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \boldsymbol{\mu})^H\boldsymbol{\Theta}^{-1}(\mathbf{Y} - \boldsymbol{\mu})\right]\right\}(d\mathbf{Y}). \tag{2.14}$$

where

$$\int_{\mathfrak{P}^{\beta}_1} u^{mn\beta/2-1}h(\beta u)du < \infty, \tag{2.15}$$

and where $\boldsymbol{\Theta} \in \mathfrak{P}^{\beta}_n$, $\boldsymbol{\Sigma} \in \mathfrak{P}^{\beta}_m$ and $\boldsymbol{\mu} \in \mathcal{L}^{\beta}_{m,n}$ are constant matrices.

Observe that this class of matrix variate elliptical distribution includes normal, contaminated normal, Pearson type II and VII, Kotz, stable, Jensen-Logistic, power exponential and Bessel distributions, among others; these distributions have tails that are more or less weighted, and/or present a greater or smaller degree of kurtosis than the normal distribution. see Fang *et al.* (1990), Fang and Zhang (1990) and Gupta *et al.* (2013).

Finally, note that for $a \in \mathcal{L}^{\beta}_{1,1}$ constant, making the change of variable $v = u/a$ and $(du) = a^{\beta}(dv)$ in Fang *et al.* (1990, Equation 2.21, p. 26) we have

$$\int_{v\in\mathfrak{P}^{\beta}_1} v^{nm\beta/2-1}h(\beta a v)(dv) = \frac{a^{-nm\beta/2}\Gamma^{\beta}_1[nm\beta/2]}{\pi^{nm\beta/2}}. \tag{2.16}$$

Finally, from Díaz-García, and Gutiérrez-Jáimez (2011)

$$\int_{\mathfrak{P}^{\beta}_m} |\mathbf{V}|^{\beta(n-m+1)/2-1} h\left(\beta \operatorname{tr} \boldsymbol{\Sigma}^{-1}\mathbf{V}\right)(d\mathbf{V}) = \frac{|\boldsymbol{\Sigma}|^{\beta n/2}\Gamma^{\beta}_m[\beta n/2]}{\pi^{\beta nm/2}}, \tag{2.17}$$

where $\mathbf{V} \in \mathfrak{P}^{\beta}_m$ and $\boldsymbol{\Sigma} \in \mathfrak{P}^{\beta}_m$.

3 Multimatrix Variate and Multimatricvariate Distributions

As the Reviewers suggest, the genesis of the terms *matrix variate* and *matricvariate* should be elucidated. From an applied point of view, both distributions are epistemologically disjoint, namely, they have non coincident study objects, then a phenomenon cannot be described by these laws at the same time. Only a distribution can be chosen within the family of matrix variate distributions or within the family of matricvariate distributions, since each class of laws model completely different situations. *It will be the responsibility of the expert in the area of application* to decide between a distribution of the matrix variate type or the matricvariate type. Moreover, from a language perspective, classical statistical and probability literature usually refer as synonymous the following expressions for a random matrix $\mathbf{X}$: a) *random matix*, in the field of *Statistics* (Muirhead (2005) and Fang and Zhang (1990)), being also the nomenclature used in *Random Matrix Theory* (Edelman and Rao (2005), Dumitriu (2002), Forrester (2009), and references therein); b) *Matrix Variate (Matrix-Variate or Matrixvariate)*, see Fang and Zhang (1990) and Gupta *et al.* (2013), among many others references; c) *Matrix Multivariate*, see Goodall and Mardia (1993); and d) *Matricvariate* proposed by Dickey (1967) and used by the author in several articles, see for example Díaz-García, and Gutiérrez-Jáimez (2012). Although these four terms have been used equivalently, only the first three options are truly synonyms.

Although the term *matricvariate* defines a random matrix properly, this expression was the result of the following quality of the t-distribution: it is know that $\mathbf{t} \in \mathcal{L}_m^1$ with $t-$multivariate distribution can be expressed in two forms, see Kotz and Nadarajah (2004, p. 2, 4); specifically:

$$\mathbf{t} = \begin{cases} R^{-1}\mathbf{y} + \boldsymbol{\mu}, & \text{with } \dfrac{\alpha R^2}{\sigma^2} \sim \chi^2(\alpha) \text{ and } \mathbf{y} \sim \mathcal{N}_m^1(\mathbf{0}, \boldsymbol{\Sigma}), \\ \mathbf{V}^{-1/2}\mathbf{y} + \boldsymbol{\mu}, & \text{with } \mathbf{V} \sim \mathcal{W}_m^1(\alpha + m - 1, \boldsymbol{\Sigma}) \text{ and } \mathbf{y} \sim \mathcal{N}_m^1(\mathbf{0}, \alpha\mathbf{I}_p), \end{cases}$$

where $\mathbf{V}^{1/2}$ is the non-negative definite square root of $\mathbf{V}$, such that $(\mathbf{V}^{1/2})^2 = \mathbf{V}$ and $\boldsymbol{\mu} \in \mathcal{L}_m^1$ is a constant vector. Now, consider the sample $\mathbf{t}_1, \dots, \mathbf{t}_n$ of a multivariate population with a $\mathbf{t}$ distribution, arranged in the matrix $\mathbf{T} = (\mathbf{t}_1 \cdots \mathbf{t}_n) \in \mathcal{L}_{n,m}^1$, then

$$\mathbf{T} = \begin{cases} \begin{pmatrix} R^{-1}\mathbf{y}_1^T + \boldsymbol{\mu}_1^T \\ \vdots \\ R^{-1}\mathbf{y}_n^T + \boldsymbol{\mu}_n^T \end{pmatrix}^T = R^{-1}\mathbf{Y} + \mathbf{M}; & \begin{cases} \dfrac{\alpha R^2}{\sigma^2} \sim \chi^2(\alpha) \\ \text{and} \\ \mathbf{Y} \sim \mathcal{N}_{m\times n}^1(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I}_n), \end{cases} \\ \text{or} \\ \begin{pmatrix} \mathbf{y}_1^T\mathbf{V}^{-1/2} + \boldsymbol{\mu}_1^T \\ \vdots \\ \mathbf{y}_n^T\mathbf{V}^{-1/2} + \boldsymbol{\mu}_n^T \end{pmatrix}^T = \mathbf{V}^{-1/2}\mathbf{Y} + \mathbf{M}; & \begin{cases} \mathbf{V} \sim \mathcal{W}_m^1(\alpha + m - 1, \boldsymbol{\Sigma}) \\ \text{and} \\ \mathbf{Y} \sim \mathcal{N}_{m\times n}^1(\mathbf{0}, \alpha(\mathbf{I}_m \otimes \mathbf{I}_n)), \end{cases} \end{cases}$$

where $\mathbf{M} = (\boldsymbol{\mu}_1 \cdots \boldsymbol{\mu}_n) \in \mathcal{L}_{n,m}^1$, and $\mathbf{Y} = (\mathbf{y}_1 \cdots \mathbf{y}_n) \in \mathcal{L}_{n,m}^1$. But matrix $\mathbf{T}$ does not have the same distribution under the above two representations, even when their columns have the same multivariate t-distribution. In the first representation, it is said that

$\mathbf{T}$ has a matrix variate $t-$distribution and under the second one, it is said that has a matricvariate $T-$distribution, see Dickey (1967). Moreover, observe that the matricvariate $T-$distribution cannot be obtained from the matrix-variate $t-$distribution, nor vice versa. This idea of differentiation between the *matrix variate* and *matricvariate* distributions was established in the context of the *Elliptically Contoured Matrix Distributions*, see Fang and Zhang (1990, Section 3.3). Essentially, using the notation of Fang and Zhang (1990, Chapter 3), the matrix variate distributions coincide with the *EVS (Elliptically-Vector-Spherical)* family distributions and the matricvariate distributions are the *ESS (Elliptically-Spherical)* family distributions. Particular examples of matrix variate distributions (*EVS* distributions) and matricvariate distributions (*ESS* distributions) as Kotz and Pearson II matrix distributions are studied in Gupta *et al.* (2013, Section 2.7.3) and Fang and Li (1999), respectively.

Now, first, observe that any new particular distribution indexed by the kernel $h(\cdot)$ is part of its density function, then it defines a family of distributions in terms of each possible choice of $h(\cdot)$.

Assume that $\mathbf{X} \sim \mathcal{E}^{\beta}_{n\times m}(\mathbf{0}, \mathbf{I}_n, \mathbf{I}_m; h)$, such that $n_0 + n_1 + \cdots + n_k = n$, and $\mathbf{X} = \left(\mathbf{X}_0^H, \mathbf{X}_1^H, \ldots, \mathbf{X}_k^H\right)^H$. Then (2.14) can be written as

$$dF_{\mathbf{X}_0,\mathbf{X}_1,\ldots,\mathbf{X}_k}(\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_k) = h[\beta \operatorname{tr}(\mathbf{X}_0^H\mathbf{X}_0 + \mathbf{X}_1^H\mathbf{X}_1 + \cdots + \mathbf{X}_k^H\mathbf{X}_k)] \bigwedge_{i=0}^{k} (d\mathbf{X}_i), \tag{3.1}$$

where $\mathbf{X}_i \in \mathcal{L}^{\beta}_{m,n_i}$, $i = 0, 1, \ldots, k$. Take into account that only under a matrix variate normal distribution the random matrices are independent, see Fang and Zhang (1990), Gupta *et al.* (2013) and Fang *et al.* (1990). In general, random matrices $\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_k$ are probabilistically dependent.

Theorem 3.1. *Suppose that* $\mathbf{X} \sim \mathcal{E}^{\beta}_{n\times m}(\mathbf{0}, \mathbf{I}_n, \mathbf{I}_m; h)$, *with* $\mathbf{X}_i \in \mathcal{L}^{\beta}_{m,n_i}$, *(recall that* $n_i \geq m$*),* $i = 0, 1, \ldots, k$.

i) *Define* $V = \operatorname{tr}\mathbf{X}_0^H\mathbf{X}_0$. *Then, the joint density* $dF_{V,\mathbf{X}_1,\ldots,\mathbf{X}_k}(v, \mathbf{X}_1, \ldots, \mathbf{X}_k)$ *is given by*

$$\frac{\pi^{n_0 m\beta/2}}{\Gamma_1^{\beta}[n_0 m\beta/2]} h\left[\beta\left(v + \operatorname{tr}\sum_{i=1}^{k}\mathbf{X}_i^H\mathbf{X}_i\right)\right] v^{n_0 m\beta/2-1}(dv) \bigwedge_{i=1}^{k}(d\mathbf{X}_i), \tag{3.2}$$

where $V \in \mathfrak{P}_1^{\beta}$. *This distribution shall be termed* multimatrix variate generalised Gamma - Elliptical distribution.

ii) *Let* $\mathbf{V} = \mathbf{X}_0^H\mathbf{X}_0$. *Hence, the joint density* $dF_{\mathbf{V},\mathbf{X}_1,\ldots,\mathbf{X}_k}(\mathbf{V}, \mathbf{X}_1, \ldots, \mathbf{X}_k)$ *is given by*

$$\frac{\pi^{\beta n_0 m/2}|\mathbf{V}|^{\beta(n_0-m+1)/2-1}}{\Gamma_m^{\beta}[\beta n_0/2]} h\left[\beta \operatorname{tr}\left(\mathbf{V} + \sum_{i=1}^{k}\mathbf{X}_i^H\mathbf{X}_i\right)\right](d\mathbf{V}) \bigwedge_{i=1}^{k}(d\mathbf{X}_i), \tag{3.3}$$

where $\mathbf{V} \in \mathfrak{P}_m^\beta$*. This distribution shall be called* multimatricvariate generalised Whishart - Elliptical distribution.

Proof. We have that the joint density function of $\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_k$ is

$$h[\beta \operatorname{tr}(\mathbf{X}_0^H \mathbf{X}_0 + \mathbf{X}_1^H \mathbf{X}_1 + \cdots + \mathbf{X}_k^H \mathbf{X}_k)] \bigwedge_{i=0}^{k} (d\mathbf{X}_i). \tag{3.4}$$

i) Let $V = \operatorname{tr} \mathbf{X}_0^H \mathbf{X}_0$, hence by (2.5)

$$(d\mathbf{X}_0) = 2^{-1} v^{n_0 m\beta/2-1} (dv) \wedge (\mathbf{h}_1^H d\mathbf{h}_1),$$

where $\mathbf{h}_1 \in \mathcal{V}_{1,n_0m}^\beta$. Thus, the multimatrix variate joint density function

$dF_{V,\mathbf{h}_1,\mathbf{X}_1,\dots,\mathbf{X}_k}(v, \mathbf{h}_1, \mathbf{X}_1, \dots, \mathbf{X}_k)$ is

$$\frac{v^{n_0 m\beta/2-1}}{2} h\left[\beta\left(v + \operatorname{tr}\sum_{i=1}^{k} \mathbf{X}_i^H \mathbf{X}_i\right)\right] (dv) \wedge (\mathbf{h}_1^H d\mathbf{h}_1) \bigwedge_{i=1}^{k} (d\mathbf{X}_i).$$

By integration over $\mathbf{h}_1 \in \mathcal{V}_{1,n_0m}^\beta$ using (2.1), the desired result is obtained.

ii) Define $\mathbf{V} = \mathbf{X}_0^H \mathbf{X}_0$. Then by (2.5), we have that

$$(d\mathbf{X}_0) = 2^{-m} |\mathbf{V}|^{\beta(n_0-m+1)/2-1} (d\mathbf{V})(\mathbf{H}_1^H d\mathbf{H}_1),$$

where $\mathbf{H}_1 \in \mathcal{V}_{m,n_0}^\beta$. The desired result is obtained by making the change of variable in (3.4) and integrating over $\mathbf{H}_1 \in \mathcal{V}_{m,n_0}^\beta$, using (2.1).

□

Proceeding as Díaz-García *et al.* (2022, Equation 4.2), defining $V_i = \operatorname{tr} \mathbf{X}_i^H \mathbf{X}_i$, $i = 0, 1, \dots, k$, and as in Díaz-García and Caro-Lopera (2022, Equation (1), p. 216) the following general result is obtained.

Theorem 3.2. *Suppose that* $\mathbf{X} = \left(\mathbf{X}_0^H, \dots, \mathbf{X}_k^H\right)^H$ *has a matrix variate spherical distribution, with* $\mathbf{X}_i \in \mathcal{L}_{m,n_i}^\beta$, $i = 0, 1, \dots, k$. *Then:*

i) *If we define* $V_i = \operatorname{tr} \mathbf{X}_i^H \mathbf{X}_i$, $i = 0, 1, \dots, k$, *then*

$$dF_{V_0,\dots,V_k}(v_0, \dots, v_k) = \pi^{nm\beta/2} \prod_{i=0}^{k} \frac{v_i^{n_i m\beta/2-1}}{\Gamma_1^\beta[n_i m\beta/2]} h\left(\beta \sum_{i=0}^{k} v_i\right) \bigwedge_{i=0}^{k} (dv_i), \tag{3.5}$$

where $n = n_0 + \cdots + n_k$, $V_i \in \mathfrak{P}_1^\beta$, $i = 0, 1, \dots, k$, *which distribution shall be termed* multivariate generalised Gamma distribution.

ii) *The joint density* $dF_{\mathbf{V}_0,\mathbf{V}_1,\dots,\mathbf{V}_k}(\mathbf{V}_0,\mathbf{V}_1,\dots,\mathbf{V}_k)$*, when* $\mathbf{V}_i = \mathbf{X}_i^H\mathbf{X}_i$*,* $i = 0,1,\dots,k$*, is given by*

$$\pi^{\beta nm/2}\prod_{i=0}^{k}\frac{|\mathbf{V}_i|^{\beta(n_i-m+1)/2-1}}{\Gamma_m^{\beta}[\beta n_i/2]}h\left(\beta\operatorname{tr}\sum_{i=0}^{k}\mathbf{V}_i\right)\bigwedge_{i=0}^{k}(d\mathbf{V}_i)\,, \tag{3.6}$$

where $\mathbf{V}_i \in \mathfrak{P}_m^{\beta}$*,* $i = 0,1,\dots,k$*. This distribution shall be termed* multimatricvariate generalised Whishart distribution.

Particular cases of these two distributions have been studied in the literature, separately, under a normal distribution in the real, complex and quaternionic cases, see Nadarajah (2007), Fang and Zhang (1990), Libby and Novick (1982), Díaz-García and Caro-Lopera (2022) and Li and Xue (2009), among many other authors.

Remark 1. In (3.1), consider the change of variables $\mathbf{X}_i = \mathbf{\Theta}^{-1/2}\mathbf{Y}_i\mathbf{\Sigma}_i^{-1/2}$, $i = 0,1,\dots,k$, where $(\mathbf{\Theta}^{1/2})^2 = \mathbf{\Theta} \in \mathfrak{P}_n^{\beta}$, $(\mathbf{\Sigma}^{1/2})^2 = \mathbf{\Sigma}_i \in \mathfrak{P}_{m_i}^{\beta}$ and $\mathbf{Y}_i \in \mathcal{L}_{m,n}^{\beta}$. Taking into account that $(d\mathbf{X}_i) = |\mathbf{\Theta}|^{-\beta m/2}|\mathbf{\Sigma}_i|^{-\beta n_i/2}(d\mathbf{X}_i)$, $i = 0,1,\dots,k$, we obtain

$$dF_{\mathbf{Y}_0,\mathbf{Y}_1,\dots,\mathbf{Y}_k}(\mathbf{Y}_0,\mathbf{Y}_1,\dots,\mathbf{Y}_k) = \frac{1}{|\mathbf{\Theta}|^{\beta km/2}\displaystyle\prod_{i=0}^{k}|\mathbf{\Sigma}_i|^{\beta n_i/2}}$$

$$\times h[\beta\operatorname{tr}(\mathbf{\Sigma}_0^{-1}\mathbf{Y}_0^H\mathbf{\Theta}^{-1}\mathbf{Y}_0+\mathbf{\Sigma}_1^{-1}\mathbf{Y}_1^H\mathbf{\Theta}^{-1}\mathbf{Y}_1+\cdots+\mathbf{\Sigma}_k^{-1}\mathbf{Y}_k^H\mathbf{\Theta}^{-1}\mathbf{Y}_k)]\bigwedge_{i=0}^{k}(d\mathbf{Y}_i).$$

Now, if $\mathbf{V}_i = \mathbf{Y}_i^H\mathbf{\Theta}^{-1}\mathbf{Y}_i = \mathbf{S}_i^H\mathbf{S}_i$, with $\mathbf{S}_i = \mathbf{\Theta}^{-1/2}\mathbf{Y}_i$, $i = 0,1,\dots,k$, then

$$(d\mathbf{S}_i) = 2^{-m}|\mathbf{V}_i|^{\beta(n_i-m+1)/2-1}(d\mathbf{V}_i)\left(\mathbf{H}_{1_i}^H d\mathbf{H}_{1_i}\right),$$

where $\mathbf{H}_{1_i} \in \mathcal{V}_{n_i,m}^{\beta}$, $i = 0,1,\dots,k$, and

$$(d\mathbf{S}_i) = |\mathbf{\Theta}|^{-\beta m/2}(d\mathbf{Y}_i).$$

Hence the joint density $dF_{\mathbf{V}_0,\mathbf{V}_1,\dots,\mathbf{V}_k}(\mathbf{V}_0,\mathbf{V}_1,\dots,\mathbf{V}_k)$ is given by

$$\pi^{\beta nm/2}\prod_{i=0}^{k}\frac{|\mathbf{V}_i|^{\beta(n_i-m+1)/2-1}}{|\mathbf{\Sigma}_i|^{\beta n_i/2}\Gamma_m^{\beta}[\beta n_i/2]}h\left(\beta\operatorname{tr}\sum_{i=0}^{k}\mathbf{\Sigma}_i^{-1}\mathbf{V}_i\right)\bigwedge_{i=0}^{k}(d\mathbf{V}_i)\,.$$

This result was obtained by Díaz-García and Caro-Lopera (2022) in the real algebra when $\mathbf{\Theta} = \mathbf{I}_n$. Moreover, from an applied point of view, assume that $\mathbf{X} \in \mathcal{L}_{m,n}^{\beta}$ denotes the observation matrix of m variables on n individuals. Generally, it is assumed that the vectors of observation between individuals are uncorrelated, which implies that $\mathbf{\Theta} = \mathbf{I}_n$, see Muirhead (2005).

Theorem 3.3. *Assume that* $\mathbf{X} \sim \mathcal{E}_{n\times m}^{\beta}(\mathbf{0},\mathbf{I}_n,\mathbf{I}_m;h)$*, with* $\mathbf{X}_i \in \mathcal{L}_{m,n_i}^{\beta}$*,* $i = 0,1,\dots,k$*.*

i) *Define* $V = \operatorname{tr} \mathbf{X}_0^H \mathbf{X}_0$ *and* $\mathbf{T}_i = V^{-1/2}\mathbf{X}_i$, $i = 1, \dots, k$. *The joint density* $dF_{V,\mathbf{T}_1,\dots,\mathbf{T}_k}(v, \mathbf{T}_1, \dots, \mathbf{T}_k)$ *is given by*

$$\frac{\pi^{n_0 m\beta/2}}{\Gamma_1^\beta[n_0 m\beta/2]} h\left[\beta v \left(1 + \sum_{i=1}^{k} \operatorname{tr} \mathbf{T}_i^H \mathbf{T}_i\right)\right] v^{nm\beta/2-1}(dv) \bigwedge_{i=1}^{k} (d\mathbf{T}_i)\,, \tag{3.7}$$

where $n = n_0 + n_1 + \cdots + n_k$, $V \in \mathfrak{P}_1^\beta$ *and* $\mathbf{T}_i \in \mathcal{L}_{m,n_i}^\beta$, $i = 1, \dots, k$. *This distribution shall be termed* multimatrix variate generalised Gamma - Pearson type VII distribution.

ii) *Define* $\mathbf{V} = \mathbf{X}_0^H \mathbf{X}_0$ *and* $\mathbf{T}_i = \mathbf{X}_i \mathbf{V}^{-1/2}$, $i = 1, \dots, k$.
Then, the joint density $dF_{\mathbf{V},\mathbf{T}_1,\dots,\mathbf{T}_k}(\mathbf{V}, \mathbf{T}_1, \dots, \mathbf{T}_k)$ *is given by*

$$\frac{\pi^{\beta n_0 m/2}|\mathbf{V}|^{\beta(n-m+1)/2-1}}{\Gamma_m^\beta[\beta n_0/2]} h\left[\beta \operatorname{tr} \mathbf{V}\left(\mathbf{I}_m + \sum_{i=1}^{k} \mathbf{T}_i^H \mathbf{T}_i\right)\right](d\mathbf{V}) \bigwedge_{i=1}^{k} (d\mathbf{T}_i)\,, \tag{3.8}$$

where $n = n_0 + n_1 + \cdots + n_k$, $\mathbf{V} \in \mathfrak{P}_m^\beta$ *and* $\mathbf{T}_i \in \mathcal{L}_{m,n_i}^\beta$, $i = 1, \dots, k$. *This distribution shall be called* multimatricvariate generalised Wishart-T distribution.

Proof. **i)** The density (3.7) follows from (3.2) by defining $\mathbf{T}_i = V^{-1/2}\mathbf{X}_i$, $i = 1, \dots, k$. Hence by Proposition 2.1, we have

$$\bigwedge_{i=1}^{k} (d\mathbf{X}_i) = v^{(n-n_0)m\beta/2} \bigwedge_{i=1}^{k} (d\mathbf{T}_i).$$

Then the required result follows.

ii) Now, take $\mathbf{T}_i = \mathbf{X}_i \mathbf{V}^{-1/2}$, $i = 1, \dots, k$ in (3.3), and consider

$$\bigwedge_{i=1}^{k} (d\mathbf{X}_i) = |\mathbf{V}|^{\beta(n-n_0)m/2} \bigwedge_{i=1}^{k} (d\mathbf{T}_i).$$

by Proposition 2.1; thus the claimed result is derived.

□

Corollary 3.1. Under the hypotheses of the theorem

i) The marginal density $dF_{\mathbf{T}_1,\dots,\mathbf{T}_k}(\mathbf{T}_1, \dots, \mathbf{T}_k)$ is termed *multimatrix variate Pearson type VII* and is given by

$$\frac{\Gamma_1^\beta[nm\beta/2]}{\pi^{(n-n_0)m\beta/2}\Gamma_1^\beta[n_0 m\beta/2]} \left(1 + \operatorname{tr} \sum_{i=1}^{k} \mathbf{T}_i^H \mathbf{T}_i\right)^{-nm\beta/2} \bigwedge_{i=1}^{k} (d\mathbf{T}_i)\,, \tag{3.9}$$

where $\mathbf{T}_i \in \mathcal{L}_{m,n_i}^\beta$ and $n = n_0 + n_1 + \cdots + n_k$.

ii) Similarly, the termed *multimatricvariate Pearson VII distribution* is the marginal density $dF_{\mathbf{T}_1,\dots,\mathbf{T}_k}(\mathbf{T}_1,\dots,\mathbf{T}_k)$ of $\mathbf{T}_1,\dots,\mathbf{T}_k$ and is given by

$$\frac{\Gamma_m^\beta[\beta n/2]}{\pi^{\beta(n-n_0)m/2}\Gamma_m^\beta[\beta n_0/2]}\left|\mathbf{I}_m+\sum_{i=1}^k \mathbf{T}_i^H\mathbf{T}_i\right|^{-\beta n/2}\bigwedge_{i=1}^k(d\mathbf{T}_i). \tag{3.10}$$

With $\mathbf{T}_i \in \mathcal{L}_{m,n_i}^\beta$ and $n = n_0+n_1+\cdots+n_k$.

Proof. **i)** Integrating (3.7) over $V \in \mathfrak{P}_1^\beta$ using (2.16) the density (3.9) is obtained.

ii) Analogously, integrating (3.8) over $\mathbf{V} \in \mathfrak{P}_m^\beta$ via (2.17), ithe following result is obtained

$$\int_{\mathfrak{P}_m^\beta}|\mathbf{V}|^{\beta(n-m+1)/2-1}h\left[\beta\operatorname{tr}\mathbf{V}\left(\mathbf{I}_m+\sum_{i=1}^k\mathbf{T}_i^H\mathbf{T}_i\right)\right](d\mathbf{V})$$

$$=\frac{\Gamma_m^\beta[\beta n/2]\left|\mathbf{I}_m+\sum_{i=1}^k\mathbf{T}_i^H\mathbf{T}_i\right|^{-\beta n/2}}{\pi^{\beta nm/2}},$$

and the desired result is archived.

□

Theorem 3.4. *Assume that* $\mathbf{X} = \left(\mathbf{X}_0^H,\dots,\mathbf{X}_k^H\right)^H$ *has a matrix variate spherical distribution, with* $\mathbf{X}_i \in \mathcal{L}_{m,n_i}^\beta$, $i = 0,1,\dots,k$.

i) *Define* $V = \operatorname{tr}\mathbf{X}_0^H\mathbf{X}_0$ *and* $\mathbf{R}_i = \left(V+\operatorname{tr}\mathbf{X}_i^H\mathbf{X}_i\right)^{-1/2}\mathbf{X}_i$, $i=1,\dots,k$. *Then the joint density of* $V,\mathbf{R}_1,\dots,\mathbf{R}_k$, *is given by*

$$\frac{\pi^{n_0m\beta/2}}{\Gamma_1^\beta[n_0m\beta/2]}v^{nm\beta/2-1}h\left[\beta v\left(1+\sum_{i=1}^k\frac{\operatorname{tr}\mathbf{R}_i^H\mathbf{R}_i}{\left(1-\operatorname{tr}\mathbf{R}_i^H\mathbf{R}_i\right)}\right)\right]$$

$$\times\prod_{i=i}^k\left(1-\operatorname{tr}\mathbf{R}_i^H\mathbf{R}_i\right)^{-n_im\beta/2-1}(dv)\bigwedge_{i=1}^k(d\mathbf{R}_i), \tag{3.11}$$

where $n=n_0+n_1+\cdots+n_k$, $V\in\mathfrak{P}_1^\beta$ *and* $\mathbf{R}_i\in\mathcal{L}_{m,n_i}^\beta$, *and* $\operatorname{tr}\mathbf{R}_i^H\mathbf{R}_i<1$ $i=1,\dots,k$. *This distribution shall be termed* multimatrix variate generalised Gamma-Pearson type II distribution.

ii) *Define* $\mathbf{V} = \mathbf{X}_0^H\mathbf{X}_0$ *and* $\mathbf{R}_i = \mathbf{X}_i(\mathbf{V}+\mathbf{X}_i^H\mathbf{X}_i)^{-1/2}$, $i=1,\dots,k$. *Then the joint density of* $\mathbf{V},\mathbf{R}_1,\dots,\mathbf{R}_k$, *denoted as* $dF_{\mathbf{V},\mathbf{R}_1,\dots,\mathbf{R}_k}(\mathbf{V},\mathbf{R}_1,\dots,\mathbf{R}_k)$, *can be written as*

$$\frac{\pi^{\beta n_0m/2}|\mathbf{V}|^{\beta(n-m+1)/2-1}}{\Gamma_m^\beta[\beta n_0/2]}h\left[\beta\operatorname{tr}\mathbf{V}\left(\mathbf{I}_m+\sum_{i=1}^k(\mathbf{I}_m-\mathbf{R}_i^H\mathbf{R}_i)^{-1}\mathbf{R}_i^H\mathbf{R}_i\right)\right]$$

$$\times \prod_{i=i}^{k} |\mathbf{I}_m - \mathbf{R}_i^H \mathbf{R}_i|^{\beta(n_i-m+1)/2-1} (d\mathbf{V}) \bigwedge_{i=1}^{k} (d\mathbf{R}_i), \tag{3.12}$$

where $n = n_0 + n_1 + \cdots + n_k$, $\mathbf{V} \in \mathfrak{P}_m^\beta$ *and* $\mathbf{R}_i \in \mathcal{L}_{m,n_i}^\beta$, $\mathbf{I}_m - \mathbf{R}_i^H \mathbf{R}_i \in \mathfrak{P}_m^\beta$, $i = 1, \dots, k$. *This distribution shall be termed* multimatricvariate generaliosed Wishart-Pearson type II distribution.

Proof. **i)** Consider $\mathbf{T}_i = (1 - tr\mathbf{R}_i^H \mathbf{R}_i)^{-1/2} \mathbf{R}_i$, $i = 1, \dots, k$ in (3.7), hence by Theorem 2.1,

$$\bigwedge_{i=1}^{k} (d\mathbf{T}_i) = \prod_{i=i}^{k} \left(1 - tr\mathbf{R}_i^H \mathbf{R}_i\right)^{-(\beta n_i m/2+1)} \bigwedge_{i=1}^{k} (d\mathbf{R}_i),$$

and the desired result follows.

ii) Define $\mathbf{T}_i = \mathbf{R}_i (\mathbf{I}_m - \mathbf{X}_i^H \mathbf{X}_i)^{-1/2}$, $i = 1, \dots, k$, with

$$\bigwedge_{i=1}^{k} (d\mathbf{T}_i) = \left|\mathbf{I}_m - \mathbf{R}^H \mathbf{R}\right|^{-\beta(n_i+m+1)/2-1} \bigwedge_{i=1}^{k} (d\mathbf{R}_i).$$

The result follows by making this change of variable in (3.8).

□

Similarly to Corollary 3.1, by (3.11) and (3.12) we can obtain the multimatrix variate and multimatricvariate marginal densities $dF_{\mathbf{R}_1,\dots,\mathbf{R}_k}$.

Corollary 3.2. Consider the assumptions of Theorem 3.4. Then

i) The multimatrix variate marginal density $dF_{\mathbf{R}_1,\dots,\mathbf{R}_k}(\mathbf{R}_1, \dots, \mathbf{R}_k)$ is

$$\frac{\Gamma_1^\beta[nm\beta/2]}{\pi^{(n-n_0)m\beta/2} \Gamma_1^\beta[n_0 m\beta/2]} \left[1 + \sum_{i=1}^{k} \frac{\operatorname{tr} \mathbf{R}_i^H \mathbf{R}_i}{\left(1 - \operatorname{tr} \mathbf{R}_i^H \mathbf{R}_i\right)}\right]^{-nm\beta/2}$$

$$\times \prod_{i=i}^{k} \left(1 - \operatorname{tr} \mathbf{R}_i^H \mathbf{R}_i\right)^{-n_i m\beta/2-1} \bigwedge_{i=1}^{k} (d\mathbf{R}_i), \tag{3.13}$$

which shall be termed *multimatrix variate Pearson type II distribution*.

ii) The multimaticvariate marginal density $dF_{\mathbf{R}_1,\dots,\mathbf{R}_k}(\mathbf{R}_1, \dots, \mathbf{R}_k)$ is

$$\frac{\Gamma_m^\beta[\beta n/2]}{\pi^{\beta(n-n_0)m/2} \Gamma_m^\beta[n_0/2]} \left|\mathbf{I}_m - \sum_{i=1}^{k} (\mathbf{I}_m - \mathbf{R}_i^H \mathbf{R}_i)^{-1} \mathbf{R}_i^H \mathbf{R}_i\right|^{-\beta n/2}$$

$$\times \prod_{i=1}^{k} \left|\mathbf{I}_m - \mathbf{R}_i^H \mathbf{R}_i\right|^{\beta(n_i-m+1)/2-1} \bigwedge_{i=1}^{k} (d\mathbf{R}_i), \tag{3.14}$$

which shall be termed *multimatricvariate Pearson type II distribution*, where $n = n_0 + n_1 + \cdots + n_k$, and $\mathbf{R}_i \in \mathcal{L}_{m,n_i}^\beta$, $i = 1, \dots, k$.

Proof. **i)** First, integration of (3.11), over $V \in \mathfrak{P}_1^\beta$ via (2.16), provides the result (3.13).

ii) Similarly, (3.14) follows by integration of (3.12) over $V \in \mathfrak{P}_m^\beta$ by using (2.17). □

Theorem 3.5. *Assuming the hypotheses of Theorem 3.3 and defining* $\mathbf{F}_i = \mathbf{T}_i^H \mathbf{T}_i > \mathbf{0}$, $i = 1, \dots, k$.

i) *The joint density* $dF_{V,\mathbf{F}_1,\dots,\mathbf{F}_k}(v, \mathbf{F}_1, \dots, \mathbf{F}_k)$ *is*

$$\frac{\pi^{nm\beta/2} v^{nm\beta/2-1}}{\Gamma_1^\beta[n_0 m\beta/2]} \prod_{i=1}^{k} \left(\frac{|\mathbf{F}_i|^{\beta(n_i-m+1)/2-1}}{\Gamma_m^\beta[n_i\beta/2]} \right) \times h\left[\beta v \left(1 + \sum_{i=1}^{k} \operatorname{tr} \mathbf{F}_i\right)\right] (dv) \bigwedge_{i=1}^{k} (d\mathbf{F}_i). \tag{3.15}$$

This distribution shall be termed multimatrix variate generalised Gamma-beta type II distribution.

ii) *Then the joint density* $dF_{\mathbf{V}_0,\mathbf{F}_1,\dots,\mathbf{F}_k}(\mathbf{V}, \mathbf{F}_1, \dots, \mathbf{F}_k)$ *is*

$$\frac{\pi^{\beta mn/2}}{\prod_{i=0}^{k} \Gamma_m^\beta[\beta n_i/2]} |\mathbf{V}|^{\beta(n-m+1)/2-1} \prod_{i=1}^{k} |\mathbf{F}_i|^{\beta(n_i-m+1)/2-1} \times h\left(\beta \operatorname{tr} \mathbf{V}\left(\mathbf{I}_m + \sum_{i=1}^{k} \mathbf{F}_i\right)\right) (d\mathbf{V}) \bigwedge_{i=1}^{k} (d\mathbf{F}_i). \tag{3.16}$$

This distribution can be termed multimatricvariate generalised Wishart-beta type II distribution.

Proof. Multimatrix variate and multimatricvariate density functions (3.15) and (3.16) are obtained considering the change of variable $\mathbf{F}_i = \mathbf{T}_i^H \mathbf{T}_i$, $i = 1, \dots, k$ in expressions (3.7) and (3.8), respectively. Now, by (2.4)

$$\bigwedge_{i=1}^{k} (d\mathbf{T}_i) = 2^{-mk} \prod_{i=1}^{k} |\mathbf{F}_i|^{\beta(n_i-m+1)/2-1} \bigwedge_{i=1}^{k} (d\mathbf{F}_i) \bigwedge_{i=1}^{k} \left(\mathbf{H}_{1_i}^H d\mathbf{H}_{1_i}\right),$$

where $\mathbf{H}_{1_i} \in \mathcal{V}_{n_i,m}^\beta$, $i = 1, \dots, k$. Hence, integration over $\mathbf{H}_{1_i} \in \mathcal{V}_{n_i,m}^\beta$ $i = 1, \dots, k$ by (2.1) achieves

$$\int_{\mathbf{H}_{1_1}} \cdots \int_{\mathbf{H}_{1_k}} \bigwedge_{i=1}^{k} \left(\mathbf{H}_{1_i}^H d\mathbf{H}_{1_i}\right) = \frac{2^{mk} \pi^{\beta(n-n_0)m/2}}{\prod_{i=1}^{k} \Gamma_m^\beta[\beta n_i/2]},$$

and the proof is complete. □

Corollary 3.3. The corresponding marginal densities $dF_{\mathbf{F}_1,\dots,\mathbf{F}_k}(\mathbf{F}_1,\dots,\mathbf{F}_k)$ of (3.15) are

i)

$$\frac{\Gamma_1^\beta[nm\beta/2]}{\Gamma_1^\beta[n_0m\beta/2]}\prod_{i=1}^k\left(\frac{|\mathbf{F}_i|^{\beta(n_i-m+1)/2-1}}{\Gamma_m^\beta[n_i\beta/2]}\right)\left(1+\sum_{i=1}^k\operatorname{tr}\mathbf{F}_i\right)^{-nm\beta/2}\bigwedge_{i=1}^k(d\mathbf{F}_i)\,, \tag{3.17}$$

whose distribution can be termed *multimatrix variate beta type II distribution*.

ii) And associated marginal density function $dF_{\mathbf{F}_1,\dots,\mathbf{F}_k}(\mathbf{F}_1,\dots,\mathbf{F}_k)$ of (3.16) is given by

$$\frac{\Gamma_m^\beta[n/2]}{\prod_{i=0}^k\Gamma_m^\beta[\beta n_i/2]}\prod_{i=1}^k|\mathbf{F}_i|^{\beta(n_i-m+1)/2-1}\left|\mathbf{I}_m+\sum_{i=1}^k\mathbf{F}_i\right|^{-\beta n/2}\bigwedge_{i=1}^k(d\mathbf{F}_i)\,. \tag{3.18}$$

This distribution shall be termed *multimatricvariate beta type II distribution*.

Proof. The proof of both results are obtained by integration of (3.15) and (3.16) respect to $V\in\mathfrak{P}_1^\beta$ and $\mathbf{V}\in\mathfrak{P}_m^\beta$ via (2.16) and (2.17), respectively. Or proceed exactly as in the proof of Theorem 3.5. □

The density function (3.17) contains real case proposed in Muirhead (2005, Problem 3.18, p.118).

Theorem 3.6. *Assume that* $\mathbf{B}_i=\mathbf{R}_i^H\mathbf{R}_i\in\mathfrak{P}_{\mathrm{m}}^\beta$, $i=1,\dots,k$, *in Theorem 3.4.*

i) *The joint density* $dF_{V,\mathbf{B}_1,\dots,\mathbf{B}_k}(v,\mathbf{B}_1,\dots,\mathbf{B}_k)$, *with* $\operatorname{tr}\mathbf{B}_i<1$, $i=1,\dots,k$ *is*

$$\frac{\pi^{nm\beta/2}v^{nm\beta/2-1}}{\Gamma_1^\beta[n_0m\beta/2]}\prod_{i=1}^k\left(\frac{|\mathbf{B}_i|^{\beta(ni-m+1)/2-1}}{\Gamma_m^\beta[n_i\beta/2]}\right)h\left[\beta v\left(1+\sum_{i=1}^k\frac{\operatorname{tr}\mathbf{B}_i}{(1-\operatorname{tr}\mathbf{B}_i)}\right)\right]$$
$$\times\prod_{i=i}^k(1-\operatorname{tr}\mathbf{B}_i)^{-n_im\beta/2-1}(dv)\bigwedge_{i=1}^k(d\mathbf{B}_i)\,. \tag{3.19}$$

This distribution shall be termed multimatrix variate generalised Gamma-beta type I distribution.

ii) *Then the joint density* $dF_{\mathbf{V},\mathbf{B}_1,\dots,\mathbf{B}_k}(\mathbf{V},\mathbf{B}_1,\dots,\mathbf{B}_k)$, *where* $\mathbf{I}_m-\mathbf{B}_i\in\mathfrak{P}_m^\beta$, $i=1,2,\dots,k$ *is*

$$\frac{\pi^{\beta mn/2}}{\prod_{i=0}^k\Gamma_m^\beta[\beta n_i/2]}|\mathbf{V}|^{\beta(n-m+1)/2-1}\prod_{i=1}^k\left(\frac{|\mathbf{B}_i|}{|\mathbf{I}_m-\mathbf{B}_i|}\right)^{\beta(n_i-m+1)/2-1}$$

$$\times h\left(\beta \operatorname{tr} \mathbf{V}\left(\mathbf{I}_m + \sum_{i=1}^{k}(\mathbf{I}_m - \mathbf{B}_i)^{-1}\mathbf{B}_i\right)\right)(d\mathbf{V}) \bigwedge_{i=1}^{k}(d\mathbf{B}_i)\,. \tag{3.20}$$

This distribution can be termed multimatricvariate generalised Wishart-beta type I distribution*.*

Proof. Making the change of variable $\mathbf{B}_i = \mathbf{R}_i^H\mathbf{R}_i$, $i = 1,\dots,k$ in the densities (3.11) and (3.12), respectively, and noting that

$$\bigwedge_{i=1}^{k}(d\mathbf{R}_i) = 2^{-mk}\prod_{i=1}^{k}|\mathbf{B}_i|^{\beta(n_i-m+1)/2-1}\bigwedge_{i=1}^{k}(d\mathbf{B}_i)\bigwedge_{i=1}^{k}\left(\mathbf{H}_{1_i}^H d\mathbf{H}_{1_i}\right),$$

where $\mathbf{H}_{1_i} \in \mathcal{V}_{n_i,m}^{\beta}$, $i = 1,\dots,k$; then the result is reached, integrating over $\mathbf{H}_{1_i} \in \mathcal{V}_{n_i,m}^{\beta}$ $i = 1,\dots,k$ via (2.1). In which case

$$\int_{\mathbf{H}_{1_1}}\cdots\int_{\mathbf{H}_{1_k}}\bigwedge_{i=1}^{k}\left(\mathbf{H}_{1_i}^H d\mathbf{H}_{1_i}\right) = \frac{2^{mk}\pi^{\beta(n-n_0)m/2}}{\displaystyle\prod_{i=1}^{k}\Gamma_m^{\beta}[\beta n_i/2]}.$$

□

Integrating (3.19) and (3.20) respect to v and $\mathbf{V}$ by (2.16) and (2.17), respectively; we obtain the marginal densities $dF_{\mathbf{B}_1,\dots,\mathbf{B}_k}(\mathbf{B}_1,\dots,\mathbf{B}_k)$ of the multimatrix variate and multimatricvariate beta type I distribution. Summarising:

Corollary 3.4. **i)** The density function $dF_{\mathbf{B}_1,\dots,\mathbf{B}_k}(\mathbf{B}_1,\dots,\mathbf{B}_k)$ can be written as

$$\frac{\Gamma_1^{\beta}[nm\beta/2]}{\Gamma_1^{\beta}[n_0m\beta/2]}\prod_{i=1}^{k}\left(\frac{|\mathbf{B}_i|^{\beta(n_i-m+1)/2-1}}{\Gamma_m^{\beta}[n_i\beta/2]\,(1-\operatorname{tr}\mathbf{B}_i)^{n_im\beta/2+1}}\right)$$
$$\times\left(1+\sum_{i=1}^{k}\frac{\operatorname{tr}\mathbf{B}_i}{(1-\operatorname{tr}\mathbf{B}_i)}\right)^{-nm\beta/2}\bigwedge_{i=1}^{k}(d\mathbf{B}_i)\,. \tag{3.21}$$

Whose marginal distribution shall be termed *multimatrix variate beta type I distribution*.

ii) In this case, the density function $dF_{\mathbf{B}_1,\dots,\mathbf{B}_k}(\mathbf{B}_1,\dots,\mathbf{B}_k)$ is

$$\frac{\Gamma_m^{\beta}[\beta n/2]}{\displaystyle\prod_{i=0}^{k}\Gamma_m^{\beta}[\beta n_i/2]}\prod_{i=1}^{k}\left(\frac{|\mathbf{B}_i|}{|\mathbf{I}_m-\mathbf{B}_i|}\right)^{\beta(n_i-m+1)/2-1}$$
$$\times\left|\mathbf{I}_m+\sum_{i=1}^{k}(\mathbf{I}_m-\mathbf{B}_i)^{-1}\mathbf{B}_i\right|^{-\beta n/2}\bigwedge_{i=1}^{k}(d\mathbf{B}_i)\,. \tag{3.22}$$

This marginal distribution shall be named *multimatricvariate beta type I distribution*.

Finally,

Theorem 3.7. *Assume that* $\mathbf{X} = \left(\mathbf{X}_0^H, \dots, \mathbf{X}_k^H\right)^H$ *has a matrix variate spherical distribution, with* $\mathbf{X}_i \in \mathcal{L}_{m,n_i}^{\beta}$, $n_i \geq m$, $i = 0, 1, \dots, k$. *Define* $V = \operatorname{tr} \mathbf{X}_0^H \mathbf{X}_0$ *and* $\mathbf{V}_i = \mathbf{X}_i^H \mathbf{X}_i$, $i = 1, \dots, k$. *Then, the joint density* $dF_{V,\mathbf{V}_1,\dots,\mathbf{V}_k}(v, \mathbf{V}_1, \dots, \mathbf{V}_k)$ *is given by*

$$\frac{\pi^{\beta nm/2} v^{\beta nm/2-1}}{\Gamma_1^{\beta}[\beta n_0 m/2]} \prod_{i=1}^{k} \left(\frac{|\mathbf{V}_i|^{\beta(ni-m+1)/2-1}}{\Gamma_m^{\beta}[\beta n_i/2]} \right)$$
$$h\left[\beta\left(v + \sum_{i=1}^{k} \operatorname{tr} \mathbf{V}_i\right)\right] (dv) \bigwedge_{i=1}^{k} (d\mathbf{V}_i), \tag{3.23}$$

where $V \in \mathfrak{P}_1^{\beta}$, $\mathbf{V}_i \in \mathfrak{P}_m^{\beta}$, $i = 1, \dots, k$. *This distribution shall be named* multimatrix variate generalised Gamma - generalised Wishart distribution.

Proof. Theorem 3.1 implies that

$$\frac{\pi^{n_0 m\beta/2} v^{n_0 m\beta/2-1}}{\Gamma_1^{\beta}[n_0 m\beta/2]} h\left[\beta\left(v + \operatorname{tr} \sum_{i=1}^{k} \mathbf{X}_i^H \mathbf{X}_i\right)\right] (dv) \bigwedge_{i=1}^{k} (d\mathbf{X}_i).$$

Defining $\mathbf{V}_i = \mathbf{X}_i^H \mathbf{X}_i$ with $i = 1, \dots, k$ and emulating the proof of Theorem 3.6, the result is reseached. □

4 Some Properties and Extensions

Now, the previous results easily induct multimatrix and multimatric variate distributions for two- and three-matrix arguments. Thus we can obtain two or more different classes of marginal distributions. The methodology can be extended to more than three different marginal distributions. In addition, the inverse distributions of some multimatrix and multimatric variate distributions are also obtained.

Theorem 4.1. *Assume that* $\mathbf{X} = \left(\mathbf{X}_0^H, \mathbf{X}_1^H, \mathbf{X}_2^H\right)^H$ *has a matrix variate spherical distribution, with* $\mathbf{X}_i \in \mathcal{L}_{m,n_i}^{\beta}$.

i) *Define* $V_0 = \operatorname{tr} \mathbf{X}_0^H \mathbf{X}_0$, $\mathbf{T} = V_0^{-1/2} \mathbf{X}_1$, *and* $\mathbf{R} = \mathbf{R} = V^{-1/2} \mathbf{X}_2$, *where* $V = (V_0 + \operatorname{tr} \mathbf{X}_2^H \mathbf{X}_2)$. *The joint density* $dF_{V,\mathbf{T},\mathbf{R}}(v, \mathbf{T}, \mathbf{R})$ *is given by*

$$\frac{\pi^{n_0 m\beta/2} v^{nm\beta/2-1}}{\Gamma\beta_1[n_0 m\beta/2]} h\left\{\beta v\left[1 + \left(1 - \operatorname{tr} \mathbf{R}^H \mathbf{R}\right) \operatorname{tr} \mathbf{T}^H \mathbf{T}\right]\right\}$$
$$\times \left(1 - \operatorname{tr} \mathbf{R}^H \mathbf{R}\right)^{(n_0+n_1)m\beta/2-1} (dv) \wedge (d\mathbf{T}) \wedge d\mathbf{R}). \tag{4.1}$$

where $n = n_0 + n_1 + n_2$, $V \in \mathfrak{P}_1^{\beta}$, $\mathbf{T} \in \mathcal{L}_{m,n_1}^{\beta}$, $\mathbf{R} \in \mathcal{L}_{m,n_2}^{\beta}$ *such that* $\operatorname{tr} \mathbf{R}^H \mathbf{R} \leq 1$. *This distribution shall be termed* threematrix variate generalised Gamma - Pearson type VII - Pearson type II distribution.

ii) *Let be* $\mathbf{V}_0 = \mathbf{X}_0^H\mathbf{X}_0$, $\mathbf{T} = \mathbf{X}_1\mathbf{V}_0^{-1/2}$, $\mathbf{V} = \mathbf{V}_0 + \mathbf{X}_2^H\mathbf{X}_2$, *and* $\mathbf{R} = \mathbf{X}_2\mathbf{V}^{-1/2}$. *Then the joint density* $dF_{\mathbf{V},\mathbf{T},\mathbf{R}}(\mathbf{V},\mathbf{T},\mathbf{R})$ *is given by*

$$\frac{\pi^{\beta n_0 m/2}}{\Gamma_m^\beta[\beta n_0/2]}|\mathbf{V}|^{\beta(n-m+1)/2-1}h\left\{\beta\operatorname{tr}\left[\mathbf{V}+(\mathbf{I}_m-\mathbf{R}^H\mathbf{R})\mathbf{V}^{1/2}\mathbf{T}^H\mathbf{T}\mathbf{V}^{1/2}\right]\right\}$$

$$\times|\mathbf{I}_m-\mathbf{R}^H\mathbf{R}|^{\beta(n_0+n_1-m+1)/2-1}(d\mathbf{V})\wedge(d\mathbf{T})\wedge(d\mathbf{R}), \qquad (4.2)$$

where $n = n_0 + n_1 + n_2$, $\mathbf{V} \in \mathfrak{P}_m^\beta$, $\mathbf{I}_m - \mathbf{R}^H\mathbf{R} \in \mathfrak{P}_m^\beta$, $\mathbf{T} \in \mathcal{L}_{m,n_1}^\beta$ *and* $\mathbf{R} \in \mathcal{L}_{m,n_1}^\beta$. *The distribution of* $\mathbf{V},\mathbf{T},\mathbf{R}$ *shall be termed* trimatricvariate generalised Wishart-Pearson VII-Pearson type II distribution.

Proof. **i)** From Theorem 3.1, $dF_{V_0,\mathbf{X}_1,\mathbf{X}_2}(v_0,\mathbf{X}_1,\mathbf{X}_2)$ is

$$\frac{\pi^{n_0 m\beta/2}}{\Gamma_1^\beta[n_0 m\beta/2]}h\left[\beta\left(v_0+\operatorname{tr}\mathbf{X}_1^H\mathbf{X}_1+\operatorname{tr}\mathbf{X}_2^H\mathbf{X}_2\right)\right]v_0^{n_0 m\beta/2-1}$$

$$(dv_0)\wedge(d\mathbf{X}_1)\wedge(d\mathbf{X}_2). \qquad (4.3)$$

Let $V = V_0 + \operatorname{tr}\mathbf{X}_2^H\mathbf{X}_2$, $\mathbf{T} = V_0^{-1/2}\mathbf{X}_1$, and $\mathbf{R} = V^{-1/2}\mathbf{X}_2$. Thus, $\mathbf{X}_1 = V_0^{1/2}\mathbf{T}$, $\mathbf{X}_2 = V^{1/2}\mathbf{R}$, and

$$V_0 = V - \operatorname{tr}\mathbf{X}_2^H\mathbf{X}_2 = V - V\operatorname{tr}\mathbf{R}^H\mathbf{R} = V(1-\operatorname{tr}\mathbf{R}^H\mathbf{R}).$$

Thus $\mathbf{T} = [V(1-\operatorname{tr}\mathbf{R}^H\mathbf{R})]^{-1/2}\mathbf{X}_1$ and $(dv) = (dv_0)$. Then, the volume element $(dv_0)\wedge(d\mathbf{X}_1)\wedge(d\mathbf{X}_2)$ is

$$v^{(n_1+n_2)m\beta/2}(1-\operatorname{tr}\mathbf{R}^H\mathbf{R})^{n_1 m\beta/2}(dv)\wedge(d\mathbf{T})\wedge(d\mathbf{R}). \qquad (4.4)$$

From (4.3), substituting $V = V_0 + \operatorname{tr}\mathbf{X}_2^H\mathbf{X}_2$, $\mathbf{X}_1 = [V(1-\operatorname{tr}\mathbf{R}^H\mathbf{R})]^{-1/2}\mathbf{T}$ and via (4.4), the desired result is obtained.

ii) By (3.3), $dF_{\mathbf{V}_0,\mathbf{X}_1,\mathbf{X}_2}(\mathbf{V}_0,\mathbf{X}_1,\mathbf{X}_2)$ is

$$\frac{\pi^{\beta n_0 m/2}|\mathbf{V}_0|^{\beta(n_0-m+1)/2-1}}{\Gamma_m^\beta[\beta n_0/2]}$$

$$\times h\left[\beta\operatorname{tr}\left(\mathbf{V}_0+\mathbf{X}_1^H\mathbf{X}_1+\mathbf{X}_2^H\mathbf{X}_2\right)\right](d\mathbf{V}_0)\wedge(d\mathbf{X}_1)\wedge(d\mathbf{X}_2).$$

Define $\mathbf{V} = \mathbf{V}_0 + \mathbf{X}_2^H\mathbf{X}_2$, $\mathbf{T} = \mathbf{X}_1\mathbf{V}_0^{-1/2}$, and $\mathbf{R} = \mathbf{X}_2\mathbf{V}^{-1/2}$. Hence by Proposition 2.1,

$$(d\mathbf{V}_0)\wedge(d\mathbf{X}_1)\wedge(d\mathbf{X}_2) = |\mathbf{V}_0|^{\beta n_1/2}|\mathbf{V}_0+\mathbf{X}_2^H\mathbf{X}_2|^{\beta n_2/2}(d\mathbf{V})\wedge(d\mathbf{T})\wedge(d\mathbf{R}).$$

But, $\mathbf{X}_1 = \mathbf{T}\mathbf{V}_0^{1/2}$, $\mathbf{X}_2 = \mathbf{R}\mathbf{V}^{1/2}$, and

$$\mathbf{V}_0 = \mathbf{V} - \mathbf{X}_2^H\mathbf{X}_2 = \mathbf{V} - \mathbf{V}^{1/2}\mathbf{R}^H\mathbf{R}\mathbf{V}^{1/2} = \mathbf{V}^{1/2}(\mathbf{I}_m-\mathbf{R}^H\mathbf{R})\mathbf{V}^{1/2}.$$

This way, $\mathbf{X}_1 = \mathbf{T}\left(\mathbf{V}^{1/2}(\mathbf{I}_m - \mathbf{R}^H\mathbf{R})\mathbf{V}^{1/2}\right)^{1/2}$ and $(d\mathbf{V}) = (d\mathbf{V}_0)$. Therefore $(d\mathbf{V}_0) \wedge (d\mathbf{X}_1) \wedge (d\mathbf{X}_2)$ is

$$|\mathbf{V}|^{\beta(n_1+n_2)/2}|\mathbf{I}_m - \mathbf{R}^H\mathbf{R}|^{\beta n_1/2}(d\mathbf{V}) \wedge (d\mathbf{T}) \wedge (d\mathbf{R}).$$

Given that

$$|\mathbf{V}_0 + \mathbf{X}_2^H\mathbf{X}_2| = |\mathbf{V}| \text{ and } |\mathbf{V}_0| = |\mathbf{V}||\mathbf{I}_m - \mathbf{R}^H\mathbf{R}|.$$

Finally, observe that $\operatorname{tr}\mathbf{X}_1^H\mathbf{X}_1 = \operatorname{tr}\mathbf{V}^{1/2}(\mathbf{I}_m - \mathbf{R}^H\mathbf{R})\mathbf{V}^{1/2}\mathbf{T}^H\mathbf{T}$. Then, the required result is obtained.

□

Corollary 4.1. Under the Hypotheses of Theorem 4.1, define $\mathbf{F} = \mathbf{T}^H\mathbf{T}$ and $\mathbf{B} = \mathbf{R}^H\mathbf{R}$.

i) The density function of the termed *trimatrix variate generalised Gamma - Pearson type VII - Pearson type II distribution*, $dF_{\mathbf{T},\mathbf{R}}(\mathbf{T},\mathbf{R})$ is given by

$$\frac{\pi^{nm\beta/2}v^{nm\beta/2-1}|\mathbf{F}|^{\beta(n_1-m+1)/2-1}|\mathbf{B}|^{\beta(n_2-m+1)/2-1}}{\Gamma_1^\beta[n_0m\beta/2]\Gamma_m^\beta[n_1\beta/2]\Gamma_m^\beta[n_2m\beta/2]}(1-\operatorname{tr}\mathbf{B})^{(n_0+n_1)m\beta/2-1}$$
$$\times h\left\{\beta v\left[1+(1-\operatorname{tr}\mathbf{B})\operatorname{tr}\mathbf{F}\right]\right\}(dv)\wedge d\mathbf{F})\wedge d\mathbf{B}). \tag{4.5}$$

ii) Similarly, the density $dF_{\mathbf{V},\mathbf{F},\mathbf{U}}(\mathbf{V},\mathbf{F},\mathbf{U})$ is

$$\frac{\pi^{\beta nm/2}|\mathbf{V}|^{\beta(n-m+1)/2-1}}{\Gamma_m^\beta[\beta n_0/2]\Gamma_m^\beta[\beta n_1/2]\Gamma_m^\beta[\beta n_2/2]}h\left[\beta\operatorname{tr}\left(\mathbf{V}+(\mathbf{I}_m-\mathbf{B})\mathbf{V}^{1/2}\mathbf{F}\mathbf{V}^{1/2}\right)\right]$$
$$\times|\mathbf{I}_m-\mathbf{B}|^{\beta(n_0+n_1-m+1)/2-1}|\mathbf{F}|^{\beta(n_1-m1)/2-1}|\mathbf{B}|^{\beta(n_2-m+1)/2-1}$$
$$(d\mathbf{V})\wedge(d\mathbf{F})\wedge(d\mathbf{B}). \tag{4.6}$$

This density shall be termed *trimatricvariate generalised Wishart-beta type II-beta type I distribution*, where $n = n_0+n_1+n_2$,$V \in \mathfrak{P}_1^\beta$, $\mathbf{F} \in \mathfrak{P}_m^\beta$, $\mathbf{B} \in \mathfrak{P}_m^\beta$, $\mathbf{V} \in \mathfrak{P}_m^\beta$, $\operatorname{tr}\mathbf{B} \le 1$ and $\mathbf{I}_m - \mathbf{B} \in \mathfrak{P}_m^\beta$.

Proof. The proofs of (4.5) and (4.6) are follows from (4.1) and (4.2), respectively; making the change of variables $\mathbf{F} = \mathbf{T}^H\mathbf{T}$, $\mathbf{B} = \mathbf{R}^H\mathbf{R}$, and by Proposition 2.4, we have that $(d\mathbf{T}) \wedge d\mathbf{R})$ is

$$2^{-2m}|\mathbf{F}|^{\beta(n_1-m+1)/2-1}|\mathbf{B}|^{\beta(n_2-m+1)/2-1}(d\mathbf{F})\wedge d\mathbf{B})\bigwedge_{i=1}^{2}\left(\mathbf{H}_{1_i}^H d\mathbf{H}_{1_i}\right).$$

The desired results are achieved, integrating over $\mathbf{H}_{1_i} \in \mathcal{V}_{n_i,m}^\beta$ $i = 1,2$ using (2.1). Moreover

$$\int_{\mathbf{H}_{1_1}}\int_{\mathbf{H}_{1_2}}\bigwedge_{i=1}^{2}\left(\mathbf{H}_{1_i}^H d\mathbf{H}_{1_i}\right) = \frac{2^{2m}\pi^{\beta(n-n_0)m/2}}{\displaystyle\prod_{i=1}^{2}\Gamma_m^\beta[\beta n_i/2]}.$$

□

Additionally, we are interested in the distributions of the inverse of one or more of the arguments in the multimatrix variate or multimatricvariate distributions, which shall be termed inverse multimatrix variate or inverse multimatricvariate distributions.

Theorem 4.2. *Define* $\mathbf{A}_i = \mathbf{B}_i^{-1}$, $i = 1, \dots, r$.

i) *Assume that* $\mathbf{B}_1, \cdots, \mathbf{B}_r, \mathbf{B}_{r+1}, \cdots, \mathbf{B}_k$ *have a multimatrix variate beta type I distribution. The join density function*

$$dF_{\mathbf{A}_1,\cdots,\mathbf{A}_r,\mathbf{B}_{r+1},\cdots,\mathbf{B}_k}(\mathbf{A}_1, \cdots, \mathbf{A}_r, \mathbf{B}_{r+1}, \cdots, \mathbf{B}_k),$$

is

$$\frac{\Gamma_1^\beta[nm\beta/2]}{\Gamma_1^\beta[n_0 m\beta/2]\prod_{i=1}^k \Gamma_m^\beta[n_i\beta/2]} \prod_{i=1}^r \left(\frac{|\mathbf{A}_i|^{-\beta(n_i+m-1)/2-1}}{\left(1 - \operatorname{tr}\mathbf{A}_i^{-1}\right)^{n_i m\beta/2+1}} \right)$$
$$\times \left(1 + \sum_{i=1}^r \frac{\operatorname{tr}\mathbf{A}_i^{-1}}{\left(1 - \operatorname{tr}\mathbf{A}_i^{-1}\right)} + \sum_{i=r+1}^k \frac{\operatorname{tr}\mathbf{B}_i}{(1 - \operatorname{tr}\mathbf{B}_i)} \right)^{-nm\beta/2}$$
$$\times \prod_{i=r+1}^k \left(\frac{|\mathbf{B}_i|^{\beta(n_i-m+1)/2-1}}{(1 - \operatorname{tr}\mathbf{B}_i)^{n_i m\beta/2+1}} \right) \bigwedge_{i=1}^r (d\mathbf{A}_i) \bigwedge_{i=r+1}^k (d\mathbf{B}_i), \tag{4.7}$$

where $\mathbf{A}_i \in \mathfrak{P}_m^\beta$, $\mathbf{B}_i \in \mathfrak{P}_m^\beta$, $\operatorname{tr}\mathbf{A}_i < 1$, $\operatorname{tr}\mathbf{B}_i < 1$.

ii) *Consider that* $\mathbf{B}_1, \cdots, \mathbf{B}_r, \mathbf{B}_{r+1}, \cdots, \mathbf{B}_k$ *have a multimatricvariate beta type I distribution. Then, the join density function*

$$dF_{\mathbf{A}_1,\cdots,\mathbf{A}_r,\mathbf{B}_{r+1},\cdots,\mathbf{B}_k}(\mathbf{A}_1, \cdots, \mathbf{A}_r, \mathbf{B}_{r+1}, \cdots, \mathbf{B}_k),$$

is given by

$$\frac{\Gamma_m^\beta[\beta n/2]}{\prod_{i=0}^k \Gamma_m^\beta[\beta n_i/2]} \prod_{i=1}^r \left(\frac{|\mathbf{A}_i|^{-\beta(m-1)-2}}{|\mathbf{A}_i - \mathbf{I}_m|^{\beta(n_i-m+1)/2-1}} \right)$$
$$\times \left| \mathbf{I}_m + \sum_{i=1}^r (\mathbf{A}_i - \mathbf{I}_m)^{-1} + \sum_{i=r+1}^k (\mathbf{I}_m - \mathbf{B}_i)^{-1}\mathbf{B}_i \right|^{-\beta n/2}$$
$$\times \prod_{i=r+1}^k \left(\frac{|\mathbf{B}_i|}{|\mathbf{I}_m - \mathbf{B}_i|} \right)^{\beta(n_i-m+1)/2-1} \bigwedge_{i=1}^r (d\mathbf{A}_i) \bigwedge_{i=r+1}^k (d\mathbf{B}_i), \tag{4.8}$$

where $\mathbf{A}_i \in \mathfrak{P}_m^\beta$, $\mathbf{B}_i \in \mathfrak{P}_m^\beta$, $\mathbf{A}_i - \mathbf{I}_m \in \mathfrak{P}_m^\beta$, $\mathbf{I}_m - \mathbf{B}_i \in \mathfrak{P}_m^\beta$.

Proof. The density functions (4.7) and (4.8) are derived from (3.21) and (3.22), respectively, defining $\mathbf{A}_i = \mathbf{B_i}^{-1}$, $i = 1, \cdots, r$, and by using the Proposition 2.5. □

The parameter domain of the real normed division algebra in the multimatrix and multimatric variate distributions can be extended as in the real and complex cases. However, the statistical and/or geometrical interpretation, perhaps can be lost. In any case, these distributions are valid if we replace $n_i/2$ by a_i, $n_0m/2$ by a_0 and $nm/2$ by a. Here the a'^s are complex numbers with a positive real part. From a practical point of view, this parameter domain extension allows to use nonlinear optimisation rather integer nonlinear optimisation in the procedure of estimation, among other possibilities.

Each distribution can be reparametrised in order to obtain a general expression for its density function. As in the normal case, the expressions obtained in this article appear in their standard form.

Assume that $(\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_k)$ follows a *multimatricvariate or multimatrix variate generalised Whishart distribution*, with density function (3.6). Define $\mathbf{W}_i = \mathbf{\Sigma}_i^{1/2}\mathbf{V}_i\mathbf{\Sigma}_i^{1/2}$, $\mathbf{\Sigma}_i \in \mathfrak{P}_m^\beta$, $i = 0, 1, \dots, k$, then by Proposition 2.2, we have that

$$\bigwedge_{i=0}^{k} (d\mathbf{V}_i) = \prod_{i=0}^{k} |\mathbf{\Sigma}_i|^{-\beta(m-1)/2-1} \bigwedge_{i=0}^{k} (d\mathbf{W}_i) .$$

Then the density $dF_{\mathbf{W}_0,\mathbf{W}_1,\dots,\mathbf{W}_k}(\mathbf{W}_0, \mathbf{W}_1, \dots, \mathbf{W}_k)$ is

$$\pi^{\beta nm/2} \prod_{i=0}^{k} \left(\frac{|\mathbf{\Sigma}_i^{-1/2}\mathbf{W}_i\mathbf{\Sigma}_i^{-1/2}|^{\beta(n_i-m+1)/2-1}}{\Gamma_m^\beta[\beta n_i/2]} \right) h\left(\beta \operatorname{tr} \sum_{i=0}^{k} \mathbf{\Sigma}_i^{-1/2}\mathbf{W}_i\mathbf{\Sigma}_i^{-1/2}\right) \prod_{i=0}^{k} |\mathbf{\Sigma}_i|^{-\beta(m-1)/2-1} \bigwedge_{i=0}^{k} (d\mathbf{W}_i) .$$

Hence

$$\pi^{\beta nm/2} \prod_{i=0}^{k} \left(\frac{|\mathbf{W}_i|^{\beta(n_i-m+1)/2-1}}{\Gamma_m^\beta[\beta n_i/2]|\mathbf{\Sigma}_i|^{\beta n_i/2}} \right) h\left(\beta \operatorname{tr} \sum_{i=0}^{k} \mathbf{\Sigma}_i^{-1}\mathbf{W}_i\right) \bigwedge_{i=0}^{k} (d\mathbf{W}_i) ,$$

where $\mathbf{W}_i \in \mathfrak{P}_m^\beta$, $i = 0, 1, \dots, k$.

5 Example

In this Section we provide an example in quaternions, addressing the computability as the main aspect of the distributions previously derived. But first, we consider some important issues about the meaning of a plausible application in the context of real normed division algebras. Excepting the algebra of real numbers, we begin by emphasising that in the context of observational statistics is almost impossible to find experiments or inherent phenomena ruled by the fields of division algebras. Some isolated phenomena ruled by complex and quaternionic algebras appear in physics and random matric theory, see Edelman and Rao (2005), Dumitriu (2002) and Forrester (2009), among others. Furthermore, for the octonionic case, rigorous assessments are needed via Baez (2002) and from an applied point of view, the relevance of the octonions remains unclear. An excellent review of the history, construction and many other properties of octonions are given in Baez (2002). It highlights that:

> "Their relevance to geometry was quite obscure until 1925, when Élie Cartan described 'triality' — the symmetry between vector and spinors in 8-dimensional Euclidean space. Their potential relevance to physics was noticed in a 1934 paper by Jordan, von Neumann and Wigner on the foundations of quantum mechanics. Work along these lines continued quite slowly until the 1980s, when it was realised that the octionions explain some curious features of string theory. **However, there is still no proof that the octonions are useful for understanding the real world**. We can only hope that eventually this question will be settled one way or another".

For the sake of completeness, the octonions shall be considered in this work. Even so, some expectations are emerging, for example, Forrester (2009, Section 1.4.5, pp. 22-24) proved that the bi-dimensional eigenvalue density function of a 2×2 octonionic matrix Gaussian ensemble is obtained from the eigenvalue general joint density function of a Gaussian ensemble with $m = 2$ and $\beta = 8$. The intention of the example is not so much the applied aspect but to prove that the distributions proposed in the article are applicable, showing that it is possible to estimate their parameters. To a large extent, our interest in this area is motivated by the observation that in the statistical literature the problems of interest are first addressed for real algebra, then for complex algebra and in much fewer cases for quaternionic algebra. This pattern in the publications, together with the approach to distribution theory and associated problems in random matrix theory, has motivated the simultaneous study of some statistical problems from an integral point of view, developing the theory for real normed division algebras.

Now, finding a random suitable data base for this algebra is difficult, then we try first to explain a way of generating a number of applications by using data bases of the literature of shape theory.

We start with a known representation of a quaternion number in terms of 2×2 complex matrices. Let $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ be a quaternion, then q can be written in terms of a the following 2×2 matrix of complex entries:

$$\mathbf{Z} = \begin{bmatrix} a + b\mathbf{i} & c + d\mathbf{i} \\ -c + d\mathbf{i} & a - b\mathbf{i} \end{bmatrix}. \tag{5.1}$$

Thus $\mathbf{Z}$ can be seen as an array of 4 complex points, with a double symmetry: $a + b\mathbf{i}$ and $a - b\mathbf{i}$ are symmetric respect the $\Re$ axis; meanwhile $c + d\mathbf{i}$ and $-c + d\mathbf{i}$ are symmetric about the $\Im$ axis.

Now, shape theory deals, for example, with sets of planar figures summarised by corresponding landmarks between populations in order to obtain means, variability and discrimination via statistics on certain quotients spaces. Instead of the noisy Euclidean space, the statistics is performed with equivalent classes after filtering out some non meaning geometrical information, such as scaling, translation, rotation, reflection, etc.. Thus, a landmark data with the referred symmetries can be set into a vector variate quaternion sample for estimation of the extended real parameters $a_i = n_i/2, i = 1, \dots, k$ of the distributions here derived. We focus on the mouse vertebra landmark data given for example in Dryden and Mardia (1998). The sample consists of 23 small, 23 large and 30 control second thoracic vertebrae with 60 landmarks. After transforming the data for a suitable application of the complex matrix representation, we have in figure 1 an example of an small vertebra.

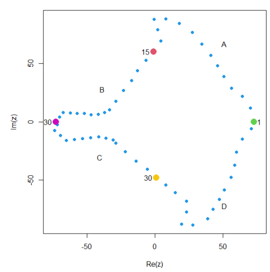
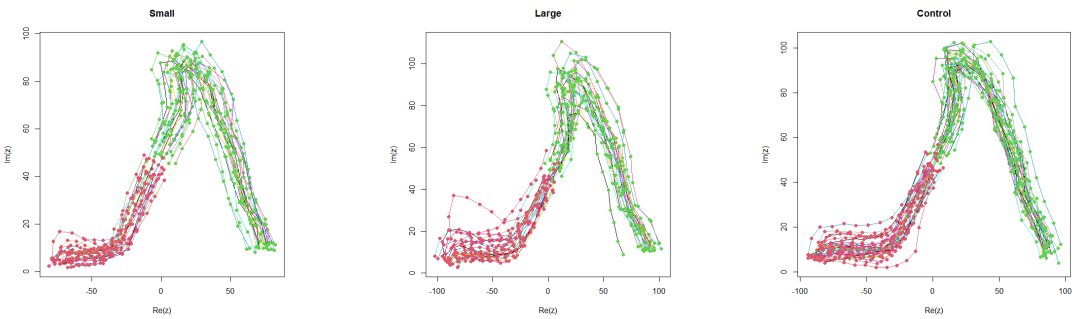


Figure 1: A 60 landmark small vertebra with high symmetry, divided in four parts A, B, C, D for suitable application of a quaternionic complex matrix representation.

The required sample most follows the symmetric suggestions of $\mathbf{Z}$. For getting this end, just cut the bones on landmarks 30 and 45 and two sectors are obtained: ABC, from landmarks 1 to 45; and, D, from landmarks 46 to 60. Now the free part D can be placed symmetrically to sector B (landmarks 15 to 30) as a reflection on the imaginary axis $\Im$. Finally, for each bone we have 14 pairs of landmarks $(a_u + b_u\mathbf{i}, c_{u+14} + d_{u+14}\mathbf{i})$, each one representing the quaternion $q_u = a_u + b_u\mathbf{i} + c_{u+14}\mathbf{j} + d_{u+14}\mathbf{k}$, where $u = 2, \dots, 15$. Namely, for each $u = 2, \dots, 15$, the first landmark $a_u + b_u\mathbf{i}$ belongs to the sector A (symmetric to sector D, respect $\Re$ axis) and it is paired with the second landmark $c_{u+15} + d_{u+15}\mathbf{i}$ (symmetric to the translated sector C). Summarising, the sample for the three classes of bones consists of upper landmarks 2 to 29, distributed by the pairs $(2, 16), (3, 17), \cdots, (14, 28), (15, 29)$. With each pair providing a quaternion, we have the following three dependent samples: 23 quaternion vectors of size 14 for the small group (Figure 2a), 23 quaternion vectors of size 14 for the large class (Figure 2b), and 30 quaternion vectors of size 14 for the control set (Figure 2c).

The mouse vertebra landmark data has been studied in several works based on the classical assumptions of normality and independent probabilistic sample, see for example Dryden and Mardia (1998) and the references therein. Gaussian restriction can not deal properly with the outlier shapes, meanwhile assuming a sample of independent small and large bones just facilitates the estimations via likelihood function, but it seems to be out of underlying sample extraction and population description given in the original source of the experiments back to the earlies 70s.



(a) (b) (c)

Figure 2: (a) Dependent sample of 23 quaternion vectors of size 14 for the small group. (b) Dependent sample of 23 quaternion vectors of size 14 for the large class. (c) Dependent sample of 30 quaternion vectors of size 14 for the control set. The sector A (green) with complex $a_u + b_u\mathbf{i}$ constitutes the first two components of the quaternion $q_u = a_u + b_u\mathbf{i} + c_{u+14}\mathbf{j} + d_{u+14}\mathbf{k}$, and the sector B (red) with complex $c_{u+15} + d_{u+15}\mathbf{i}$ indexes the last two components of $q_u, u = 2, \ldots, 15$.

This work provides two solutions of the previous problems. First, within the multimatricvariate and multimatrix variate distributions, the choice of distributions that are invariant under the family of $\beta-$elliptically contoured distributions. A notorious advantage facing the lack of knowledge about the random matrix law for the landmark data. And second, the distributions here derived give join density functions of dependent matrices, a fact eliminates the historical controversy between the probabilistic independence of the sample data. Another important fact, hidden for the classical studies based on Gaussianity (independently equivalent) resides on the real normed division algebra supporting the landmark data. In particular, the use of quaternions via the complex matrix representation integrates in the study the high symmetry of the bones. Finally, for this data, the distributions under quaternions are extremely simple, they are reduced to vectors instead of the real matrix setting given in the referred studies.

For the sake of illustration and simple computation, we consider the multimatrix variate beta type II distribution (3.17) as the likelihood function dependent sample invariant under the quaternionic elliptically contoured distribution.

For the small and large samples, $k = 23$, $\beta = 4$ and $m = 1$, meanwhile in the control group $k = 30$. In the three samples we use (3.17) for the maximum likelihood estimates of $a_0 = n_0/2$ and $a = n_i/2, i = 1, \ldots, k$. Here, $\mathbf{T}_i$ is a 14×1 quaternion vector, then $F_i = \mathbf{T}_i^H \mathbf{T}_i$ is a real value for all $i = 1, \ldots, k$. Thus the likelihood in (3.17) takes the form:

$$\frac{\Gamma_1^\beta\left[(a_0 + ka)m\beta\right]}{\Gamma_1^\beta\left[a_0 m\beta\right]\left(\Gamma_m^\beta\left[a\beta\right]\right)^k} \prod_{i=1}^{k} F_i^{\beta\left(a - \frac{m}{2} + \frac{1}{2}\right) - 1} \left(1 + \sum_{i=1}^{k} F_i\right)^{-(a_0 + ka)m\beta}.$$

The computations are performed in the Optimx package of R under several methods of optimisations and a wide range of seeds for a consistent estimation. The results are given in the following table

Sample	$\hat{a}_0$	$\hat{a}$
Small	0.040714	45.194923
Large	0.03294941	12.82131179
Control	0.03765324	32.99296063

Finally, we can use again the symmetry of the modified mouse vertebra data in order to define 14×2 quaternion matrices $\mathbf{T}_i, i = 1, \dots, k$. In this case the first column is the same formed by sectors A and B, and the second column corresponds to sector D (the reflection of sector A) and the translated and reflected sector C, which is symmetric respect sector B. Then we obtain the 2×2 quaternion matrix $\mathbf{F}_i = \mathbf{T}_i^H \mathbf{T}_i$, $i = 1, \dots, k$, and the likelihood function (3.17) can be computed in terms of the latent roots of the quaternionic Hermitian matrices $\mathbf{F}_i$.

The use of high symmetric planar landmark data for characterising quatenion applications opens an interesting perspective in shape theory, in particular the computation of probabilities are tractable expressions which can be implemented easily. Finally, under these symmetries, (5.1) shows a way of avoiding the quaternions by performing a study based on block 2×2 complex matrices, a real normed division algebra more easily understood and handled because its commutative property. These aspects are taking part of a future work.

6 Conclusions

This work has set the multimatrix and multimatric variate distributions in a unified approach for the real normed division algebras. The distributions are also indexed by the class of elliptical contoured models. The main advantages of the proposed theory are in agreement with the current paradigms of the distribution theory: 1) The distributions are computable in a simple PC. 2) After integrations, the results can be seen as joint distributions of several combinations of scalars, vectors and matrix variates, some of them invariant under the family of matrix variate elliptically contoured distributions. An ideal property for situations where the marginals and joint distributions are completely unknown. 3) The multimatrix variate and multimatricvariate distributions share the philosophy of copula theory, but without the restriction to reals, vectors and likelihood copula parameter estimation based on independent distributions. 4) The join distributions can be seen as likelihood functions of probabilistically dependent matrices, as a more real alternative a likelihood function of independent sample variate. 5) The multimatrix variate and multimatricvariate distributions emerged into as unified point of view for all the real normed division algebras, just modulated by a parameter $\beta = 1, 2, 4, 8$. 6) The properties presented here are valid for all real normed division algebras, then several applications can be switched according to the sample dependent origin. Finally, a application of symmetric landmark data popularised in real shape theory is translated into the quaternion setting. Current a research about multiple computation of probabilities on symmetric cones is considered.

Acknowledgements

The authors wish to thank the Editor and the anonymous Reviewers for their constructive comments on the preliminary version of this article.

References

Bekker, A., Roux, J. J. J., Ehlers, E., Arashi, M. (2001), Bimatrix variate beta type IV distribution: relation to Wilks's statistics and bimatrix variate Kummer-beta type IV distribution. Comm. Statist. (T&M) 40, 4165-4178.

Baez, J. C. (2002), The octonions. Bull. Amer. Math. Soc. 39, 145–205.

Chen, J. J., Novick, M. R. (1984), Bayesian analysis for binomial models with generalized beta prior distributions. J. Educational Statist. 9, 163–175.

Díaz-García, J. A. (2014), Integral properties of zonal spherical functions, hypergeometric functions and invariant polynomials. J Iranian Statist. Soc. 13 (1), 83-124.

Díaz-García, J. A., Gutiérrez-Jáimez, R. (2010a), Bimatrix variate generalised beta distributions. South African Statist. J. 44, 193-208.

Díaz-García, J. A., Gutiérrez-Jáimez, R. (2010b), Complex bimatrix variate generalised beta distributions. Linear Algebra Appl. 432 (2-3), 571-582.

Díaz-García, J. A., Gutiérrez-Jáimez, R. (2011), On Wishart distribution: Some extensions. Linear Algebra Appl. 435, 1296-1310.

Díaz-García, J. A., Gutiérrez-Jáimez, R. (2011), Noncentral bimatrix variate generalised beta distributions. Metrika. 73(3), 317-333.

Díaz-García, J. A., Gutiérrez-Jáimez, R. (2011), On Wishart distribution: Some extensions. Linear Algebra Appl. 435, 1296-1310.

Díaz-García, J. A., Gutiérrez-Jáimez, R. (2012), Matricvariate and matrix multivariate T distributions and associated distributions. Metrika, 75(7), 963-976.

Díaz-García, J. A., Gutiérrez-Jáimez, R. (2013), Spherical ensembles. Linear Algebra Appl. 438, 3174-3201.

Díaz-García, J. A., Caro-Lopera, F. J. (2022), Multimatricvariate distribution under elliptical models. J. Stat. Plann. Infer. 2016, 109-117.

Díaz-García, J. A., Caro-Lopera, F. J., Pérez Ramírez, F. O. (2022), Multivector variate distributions: An application in Finance. Sankhyā 84-A,Part 2, 534-555.

Díaz-García, J. A., Caro-Lopera, F. J. (2024), Multimatrix variate distribution. https://arxiv.org/abs/2405.02498.

Dickey, J. M. (1967), Matricvariate generalizations of the multivariate *t*- distribution and the inverted multivariate *t*-distribution. Ann. Math. Statist. 38, 511-518.

Dumitriu, I. (2002), Eigenvalue statistics for beta-ensembles. PhD thesis, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA.

Dray, T., Manogue, C. A. (1999), The exceptional Jordan eigenvalue problem, Inter. J. Theo. Phys. 38(11), 2901–2916.

Dryden, I. L., Mardia, K. V. (1998), Statistical Shape Analysis. Wiley, Chichester.

Ehlers, R. (2011, Bimatrix variate distributions of Wishart ratios with application. Doctoral dissertation, Faculty of Natural & Agricultural Sciences University of Pretoria, Pretoria. http://hdl.handle.net/2263/31284.

Edelman, A. Rao, R. R. (2005), Random matrix theory. Acta Numer. 14, 233–297.

Fang, K. T., Li, R. (1999), Bayesian Statistical Inference on Elliptical Matrix Distributions. J. Multivarialte Anal. 70, 66-85.

Fang, K. T., Zhang, Y. T. (1990), Generalized Multivariate Analysis. Science Press, Springer-Verlag, Beijing, 1990.

Fang, K. T., Zhang, Y. T., Ng, K. W. (1990), Symmetric Multivariate and realted distributions. Springer-Science+Business Media, B. V., New Delhi.

Forrester, P. J. (2009), Log-gases and random matrices. To appear. Available in: `http://www.ms.unimelb.edu.au/~matpjf/matpjf.html`.

Gross, K. I., Richards, D. ST, P. (1987), Special functions of matrix argument I: Algebraic induction zonal polynomials and hypergeometric functions. Trans. Amer. Math. Soc. 301(2), 475–501.

Goodall, C. R., Mardia, K. V. (1993), Multivariate Aspects of Shape Theory. Ann. Statist. 21, 848–866.

Gupta, A. K., Varga, T., Bondar, T. (2013), Elliptically Contoured Models in Statistics and Portfolio Theory. 2nd Edition, Springer.

Kabe, D. G. (1984), Classical statistical analysis based on a certain hypercomplex multivariate normal distribution. Metrika 31, 63–76.

Kotz, S., and Nadarajah, S. (2004), Multivariate *t* Distributions and Their Applications. Cambridge University Press, United Kingdom.

Libby, D. L., Novick, M. R. (1982), Multivariate Generalized beta distributions with applications to utility assessment. J. Educational Statist. 7, 271–294.

Li, F., Xue, Y. (2009), Zonal polynomials and hypergeometric functions of quaternion matrix argument. Comm. Statist. Theory Methods 38(8), 1184-1206.

Muirhead, R. J. (2005), Aspects of Multivariate Statistical Theory. John Wiley & Sons, New York.

Nadarajah, S. (2007), A bivariate gamma model for drought. Water Resour. Res. 43, W08501, doi:10.1029/2006WR005641.

Nadarajah, S. (2013), A bivariate distribution with gamma and beta marginals with application to drought data. J. App. Statist. 36(3), 277-301.

Olkin, I., Liu, R., (2003), A bivariate beta distribution. Statist. Prob. Letters, 62, 407–412.

Olkin, I., Rubin, H., (1964), Multivariate beta distributions and independence properties of Wishart distribution. Ann. Math. Statist. 35, 261–269. Correction 1966, 37(1), 297.

Sarabia, J. M., Prieto, F., Jordá, V. (2014), Bivariate beta-generated distributions with application to well-being data. J. Statist. Distributions Appl. 1:15. http://www.jsdajournal.com/content/1/1/15.