Jordan Journal of Mathematics and Statistics.
Yarmouk University

DOI:https://doi.org/10.47013/18.3.11

On the Dirichlet-hypergeometric distribution on symmetric matrices

Mohamed Ben Farah *

Laboratory of Probability and Statistics, Sfax University, B.P. 1171, 3000, Sfax, Tunisia.

Received: April 24, 2024 Accepted: Jun. 29, 2025

Abstract: In this paper, we introduce an extension of the real Dirichlet-hypergeometric distribution to symmetric matrices, motivated by the need for structured matrix-valued distributions in multivariate analysis and random matrix theory. This matrix-variate generalization preserves the combinatorial and probabilistic foundations of the classical Dirichlet model while adapting them to the geometric and algebraic structure of symmetric matrices. We study distributional properties, including marginal and conditional distributions, the computation of moments, and derive the distribution of partial matrix sums. Furthermore, we introduce a class of related matrix distributions that naturally emerge in this framework. These results lay a foundation for further theoretical development and potential applications in areas requiring matrix-valued priors or structured random matrices.

Keywords: Dirichlet distribution; Gauss hypergeometric function; Liouville distribution; Transformation; symmetric matrices.

2010 Mathematics Subject Classification. 62E15; 60E05.

1 Introduction

The Dirichlet distribution has numerous applications across diverse fields. In particular, it has been employed as a conjugate prior for the multinomial distribution in Bayesian analysis. Additionally, it has been utilised in machine learning, where it has been employed in topic modelling and probabilistic graphical modelling. Furthermore, it has been applied in genetics, where it has been used to model allele frequencies within a given population. It has been extensively studied by Aitchison [1,2], Fang [3], Tang et al. [4], Tsagris and Stewart [5] and Balakrishnan and Nevzorov [6]. A number of different generalisations of the Dirichlet distribution have been proposed in the literature. These include Connor and Mosimann's distribution (Connor and Mosimann [7]), the scaled Dirichlet distribution (Libby and Novick [8], Chen and Novick [9], and Monti et al. [10]), the Liouville distribution (Marshall and Olkin [11], Bhattacharya [12] and Rayens and Srinivasan [13]), and the hyper-Dirichlet distribution (Hankin [14]). Nagar et al. [15] proposed another generalization of the Dirichlet distribution using the hypergeometric Gauss function. Its probability density function (pdf) is as follows:

$$\frac{\Gamma\left(\sum_{i=1}^n \alpha_i+\delta-\beta\right)\Gamma\left(\sum_{i=1}^n \alpha_i+\delta-\gamma\right)}{\prod_{i=1}^n \Gamma(\alpha_i)\Gamma(\delta)\Gamma\left(\sum_{i=1}^n \alpha_i+\delta-\beta-\gamma\right)}\prod_{i=1}^n x_i^{\alpha_i-1}\left(1-\sum_{i=1}^n x_i\right)^{\delta-1} {}_2F_1\left(\beta,\gamma;\delta;1-\sum_{i=1}^n x_i\right), \tag{1}$$

where $(x_1,\ldots,x_n)\in(0,+\infty)^n$ such that $\sum_{i=1}^n x_i<1$, $\alpha_i>0$, $i=1,\ldots,n$, $\delta>0$, $\beta,\gamma\in\mathbb{R}$ such that $\sum_{i=1}^n \alpha_i+\delta>\beta+\gamma$, $\Gamma(.)$ is the gamma function defined by

$$\Gamma(\alpha)=\int_0^\infty e^{-x}x^{\alpha-1}dx,$$

and ${}_2F_1$ is the Gauss hypergeometric function (see Luke [16]).

The extension of the generalized real Dirichlet distribution to symmetric matrices is of significant importance, as it provides a powerful and flexible tool for the modelling of complex, symmetric compositional data that arises in various

* Corresponding author e-mail: mohamed.benfarah@isbs.rnu.tn

fields. It facilitates the capture of intricate dependencies, enables Bayesian inference for matrix-valued parameters, and opens up new avenues for the analysis of structured multivariate data, particularly in network analysis and other domains where symmetric relationships are fundamental. In this paper, we introduce and study an extension of the density (1) to symmetric matrices that involves the Gauss hypergeometric function of the matrix argument.

Let r be a positive integer. We denote the space of real symmetric $r\times r$ matrices by V_r and the cone of positive definite elements of V_r by Ω_r. The identity matrix is denoted by I_r, the determinant of an element x in V_r by $\det(x)$, and its trace by $\mathrm{tr}(x)$. In our study, we use the concept of a "quotient" which is defined by the division algorithm on matrices. More precisely, we exploit the fact that a symmetric positive definite matrix x can be uniquely expressed as the square of a matrix denoted $x^{\frac{1}{2}}$. The quotient of y by x is then defined as $x^{-\frac{1}{2}}yx^{-\frac{1}{2}}$, where $x^{-\frac{1}{2}}$ represents the inverse of the matrix $x^{\frac{1}{2}}$.

Furthermore, we define the multivariate gamma function, which is commonly used in multivariate statistical analysis, by

$$\begin{aligned}\Gamma_r(\alpha) &= \int_{\Omega_r} e^{-\mathrm{tr}(x)}\det(x)^{\alpha-\frac{r+1}{2}}dx\\ &= \pi^{\frac{r(r-1)}{4}}\prod_{i=1}^{r}\Gamma\left(\alpha-\frac{i-1}{2}\right), \qquad Re(\alpha)>\frac{r-1}{2}.\end{aligned}$$

The generalized hypergeometric coefficient is defined by

$$(\alpha)_\rho = \prod_{i=1}^{r}\left(\alpha-\frac{i-1}{2}\right)_{\rho_i},$$

$$(\alpha)_j = \alpha(\alpha+1)\dots(\alpha+j-1),\ j=1,2,\dots \ \ and \ \ (\alpha)_0=1,$$

where $\rho=(\rho_1,\dots,\rho_r)$, $\rho_1\geq\dots\geq\rho_r\geq 0$. The generalized hypergeometric function of matrix argument is defined by

$${}_pF_q(\alpha_1,\dots,\alpha_p;\beta_1,\dots,\beta_q;x)=\sum_{k=0}^{\infty}\sum_{\kappa}\frac{(\alpha_1)_\kappa\dots(\alpha_p)_\kappa}{(\beta_1)_\kappa\dots(\beta_q)_\kappa}\frac{C_\kappa(x)}{k!} \tag{2}$$

where α_i, $i=1,\dots,p$, β_j, $j=1,\dots,q$ are complex numbers so long as none of β_j is zero or an integer or half integer $\leq\frac{r-1}{2}$, $C_\kappa(x)$ is the zonal polynomial of $r\times r$ complex symmetric matrix x corresponding to the partition κ and $\sum_\kappa$ denotes summation over all partition κ. Conditions of the convergence of the serie in (2) have been documented in the literature, see, for example, Constantine [17] and James [18]. From (2), it can be seen that

$${}_0F_0(x)=\sum_{k=0}^{\infty}\sum_{\kappa}\frac{C_\kappa(x)}{k!}=\sum_{k=0}^{\infty}\frac{(\mathrm{tr}(x))^k}{k!}=e^{\mathrm{tr}(x)}$$

$${}_1F_0(\alpha_1;x)=\sum_{k=0}^{\infty}\sum_{\kappa}(\alpha_1)_\kappa\frac{C_\kappa(x)}{k!}=\det(I_r-x)^{-\alpha_1},\ x\in\Omega_r\cap(I_r-\Omega_r)$$

and

$${}_2F_1(\alpha_1,\alpha_2;\beta_1;x)=\sum_{k=0}^{\infty}\sum_{\kappa}\frac{(\alpha_1)_\kappa(\alpha_2)_\kappa}{(\beta_1)_\kappa}\frac{C_\kappa(x)}{k!},\ x\in\Omega_r\cap(I_r-\Omega_r).$$

The Gauss hypergeometric function ${}_2F_1$ is defined by the following integral

$${}_2F_1(\alpha,\beta;\gamma;x)=\frac{\Gamma_r(\gamma)}{\Gamma_r(\alpha)\Gamma_r(\gamma-\alpha)}\int_{\Omega_r\cap(I_r-\Omega_r)}\det(u)^{\alpha-\frac{r+1}{2}}\det(I_r-u)^{\gamma-\alpha-\frac{r+1}{2}}\det(I_r-xu)^{-\beta}du \tag{3}$$

where $\beta\in\mathbb{C}$, $Re(\alpha)>(r-1)/2$ and $Re(\gamma-\alpha)>(r-1)/2$. It can be observed from the above that

$${}_2F_1(\alpha,\beta;\gamma;I_r)=\frac{\Gamma_r(\gamma)\Gamma_r(\gamma-\alpha-\beta)}{\Gamma_r(\gamma-\beta)\Gamma_r(\gamma-\alpha)},\ Re(\gamma-\alpha-\beta)>\frac{r-1}{2}. \tag{4}$$

Also, we have

$$\int_{\Omega_r\cap(I_r-\Omega_r)}\det(x)^{\alpha-\frac{r+1}{2}}\det(I_r-x)^{\beta-\frac{r+1}{2}}\,{}_2F_1(\nu,\gamma;\delta;u(I_r-x))dx=\frac{\Gamma_r(\alpha)\Gamma_r(\beta)}{\Gamma_r(\alpha+\beta)}\,{}_3F_2(\beta,\nu,\gamma;\alpha+\beta,\delta;u), \tag{5}$$

for all $u\in V_r$, where $\nu,\gamma\in\mathbb{C}$, $Re(\alpha),Re(\beta),Re(\delta)>(r-1)/2$.

Finally, we define the Dirichlet distribution on symmetric matrices. To explore the properties of this distribution, readers are advised to refer to the works of Ben Farah and Hassairi [19,20], Ben Farah and Ghorbel [21], Gupta and Nagar [22], and Olkin and Rubin [23]. The vector of random matrices $(X_1,\dots,X_n)$ is said to have a Dirichlet distribution on V_r with parameters $(\alpha_1,\dots,\alpha_n;\alpha_{n+1})$, denoted by $D_r(\alpha_1,\dots,\alpha_n;\alpha_{n+1})$, if its pdf is given by

$$\frac{\Gamma_r\left(\sum_{i=1}^{n+1}\alpha_i\right)}{\prod_{i=1}^{n+1}\Gamma_r(\alpha_i)}\prod_{i=1}^{n}\det(x_i)^{\alpha_i-\frac{r+1}{2}}\det\left(I_r-\sum_{i=1}^{n}x_i\right)^{\alpha_{n+1}-\frac{r+1}{2}},\qquad (x_1,\dots,x_n)\in T_n, \tag{6}$$

where $\alpha_i > \frac{r-1}{2}$, $i=1,\dots,n+1$ and T_n is given by

$$T_n=\left\{(x_1,\dots,x_n)\in\Omega_r^n;\sum_{i=1}^{n}x_i\in\Omega_r\cap(I_r-\Omega_r)\right\}.$$

For $n=1$ in (6), the distribution of the random matrix X_1 is called the beta distribution on V_r with parameters (α_1,α_2) and is denoted by $B_r(\alpha_1,\alpha_2)$. Its pdf is given by

$$\frac{\Gamma_r(\alpha_1+\alpha_2)}{\Gamma_r(\alpha_1)\Gamma_r(\alpha_2)}\det(x_1)^{\alpha_1-\frac{r+1}{2}}\det(I_r-x_1)^{\alpha_2-\frac{r+1}{2}},\qquad x_1\in\Omega_r\cap(I_r-\Omega_r). \tag{7}$$

Let $f(.)$ be a continuous function and $\alpha_i>\frac{r-1}{2}$, $i=1,\dots,n$. The Liouville-Dirichlet's integral is defined (see Gupta and Richards [24]) by

$$L_n(\alpha_1,\dots,\alpha_n;f)=\int_{T_n}\prod_{i=1}^{n}\det(x_i)^{\alpha_i-\frac{r+1}{2}}f\left(\sum_{i=1}^{n}x_i\right)\prod_{i=1}^{n}dx_i.$$

Substituting

$$y_i=x^{-\frac{1}{2}}x_ix^{-\frac{1}{2}},\qquad i=1,\dots,n-1 \text{ and } x=\sum_{i=1}^{n}x_i,$$

with the Jacobian $J(x_1,\dots,x_{n-1},x_n\longrightarrow y_1,\dots,y_{n-1},x)=\det(x)^{(n-1)\frac{r+1}{2}}$ in the above integral and using (6), we obtain

$$L_n(\alpha_1,\dots,\alpha_n;f)=\frac{\prod_{i=1}^{n}\Gamma_r(\alpha_i)}{\Gamma_r\left(\sum_{i=1}^{n}\alpha_i\right)}\int_{\Omega_r\cap(I_r-\Omega_r)}\det(x)^{\sum_{i=1}^{n}\alpha_i-\frac{r+1}{2}}f(x)dx. \tag{8}$$

2 The Dirichlet-hypergeometric distribution on symmetric matrices

In this section, we begin by extending the definition of the generalized real Dirichlet distribution given by Nagar et al. [15], to symmetric matrices. We then prove that this distribution is considered to be the distribution of the beta matrix multiplied by a vector of Dirichlet matrices.

Definition 1. *The random matrices $X_1,\dots,X_n$ are said to have a Dirichlet-hypergeometric distribution on V_r, denoted by $DH_r(\alpha_1,\dots,\alpha_n;\beta,\gamma;\delta)$, if their pdf is given by*

$$C(\alpha_1,\dots,\alpha_n;\beta,\gamma;\delta)\prod_{i=1}^{n}\det(x_i)^{\alpha_i-\frac{r+1}{2}}\det\left(I_r-\sum_{i=1}^{n}x_i\right)^{\delta-\frac{r+1}{2}}{}_2F_1\left(\beta,\gamma;\delta;I_r-\sum_{i=1}^{n}x_i\right), \tag{9}$$

where $(x_1,\dots,x_n)\in T_n$, $\alpha_i>\frac{r-1}{2}$, $i=1,\dots,n$, $\delta>\frac{r-1}{2}$, $\beta,\gamma\in\mathbb{R}$ such that $\sum_{i=1}^{n}\alpha_i+\delta>\beta+\gamma+\frac{r-1}{2}$ and $C(\alpha_1,\dots,\alpha_n;\beta,\gamma;\delta)$ is the normalizing constant.

Using (8) , the normalizing constant in (9) is given by

$$\{C(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)\}^{-1} = \int_{T_n}\prod_{i=1}^{n}\det(x_i)^{\alpha_i-\frac{r+1}{2}}\det\left(I_r-\sum_{i=1}^{n}x_i\right)^{\delta-\frac{r+1}{2}}\ {}_2F_1\left(\beta,\gamma;\delta;I_r-\sum_{i=1}^{n}x_i\right)\prod_{i=1}^{n}dx_i$$
$$= \frac{\prod_{i=1}^{n}\Gamma_r(\alpha_i)}{\Gamma_r\left(\sum_{i=1}^{n}\alpha_i\right)}\int_{\Omega_r\cap(I_r-\Omega_r)}\det(x)^{\sum_{i=1}^{n}\alpha_i-\frac{r+1}{2}}\det(I_r-x)^{\delta-\frac{r+1}{2}}\ {}_2F_1(\beta,\gamma;\delta;I_r-x)dx,$$

Using (5) and (4), we get

$$\{C(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)\}^{-1} = \frac{\prod_{i=1}^{n}\Gamma_r(\alpha_i)\Gamma_r(\delta)\Gamma_r\left(\sum_{i=1}^{n}\alpha_i+\delta-\beta-\gamma\right)}{\Gamma_r\left(\sum_{i=1}^{n}\alpha_i+\delta-\beta\right)\Gamma_r\left(\sum_{i=1}^{n}\alpha_i+\delta-\gamma\right)}. \tag{10}$$

The particular case when $n=1$ in (9), the distribution of the random matrix X_1 is called the beta-hypergeometric distribution on V_r with parameters $(\alpha_1;\beta,\gamma;\delta)$ and is denoted by $BH_r(\alpha_1;\beta,\gamma;\delta)$. Its pdf is given by

$$C(\alpha_1;\beta,\gamma;\delta)\det(x_1)^{\alpha_1-\frac{r+1}{2}}\det(I_r-x_1)^{\delta-\frac{r+1}{2}}\ {}_2F_1(\beta,\gamma;\delta;I_r-x_1),$$

where $x_1\in\Omega_r\cap(I_r-\Omega_r)$ and $C(\alpha_1;\beta,\gamma;\delta)=\dfrac{\Gamma_r(\alpha_1+\delta-\beta)\Gamma_r(\alpha_1+\delta-\gamma)}{\Gamma_r(\alpha_1)\Gamma_r(\delta)\Gamma_r(\alpha_1+\delta-\beta-\gamma)}$.
This distribution have been introduced by Gupta and Nagar [22].

Note that if $\beta=\delta$, the beta-hypergeometric distribution $BH_r(\alpha_1;\beta,\gamma;\delta)$ reduces to the beta distribution $B_r(\alpha_1-\gamma,\beta)$. Further, if we take $\beta=\delta$ and $\gamma=0$ in (9), the Dirichlet-hypergeometric distribution $DH_r(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$ reduces to the Dirichlet distribution $D_r(\alpha_1,\ldots,\alpha_n;\beta)$.

The following theorem shows that the Dirichlet-hypergeometric distribution on symmetric matrices is K-invariant, where K is the orthogonal group of V_r.

Theorem 1. *Let $X=(X_1,\ldots,X_n)\sim DH_r(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$. Then the distribution of X is K-invariant.*

Proof. Let k be an orthogonal $r\times r$ matrix independent of X. By transforming $y_i=kx_ik'$, $i=1,\ldots,n$, in the pdf (9), with the Jacobian $J(x_1,\ldots,x_n\longrightarrow y_1,\ldots,y_n)=1$, we get that the conditional distribution of $kXk'\mid k$ is $DH_r(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$. As this distribution is independent of k, it follows that $kXk'\sim DH_r(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$.

In the next theorem, the Dirichlet-hypergeometric distribution on V_r is derived from the Dirichlet distribution.

Theorem 2. *Let $(Y_1,\ldots,Y_n)\sim D_r(\alpha_1,\ldots,\alpha_n;\delta)$ and $Z\sim B_r(\beta,\gamma)$ be independent. Define*

$$X_i = Z^{\frac12}Y_iZ^{\frac12},\ \ i=1,\ldots,n.$$

Then $(X_1,\ldots,X_n)\sim DH_r(\alpha_1,\ldots,\alpha_n;\gamma,\sum_{i=1}^{n}\alpha_i+\delta-\beta;\gamma+\delta)$.

Proof. According to the conditions of the theorem, the pdf of $(Y_1,\ldots,Y_n,Z)$ is given by

$$c\prod_{i=1}^{n}\det(y_i)^{\alpha_i-\frac{r+1}{2}}\det\left(I_r-\sum_{i=1}^{n}y_i\right)^{\delta-\frac{r+1}{2}}\det(z)^{\beta-\frac{r+1}{2}}\det(I_r-z)^{\gamma-\frac{r+1}{2}}, \tag{11}$$

where $c=\dfrac{\Gamma_r\left(\sum_{i=1}^{n}\alpha_i+\delta\right)\Gamma_r(\beta+\gamma)}{\prod_{i=1}^{n}\Gamma_r(\alpha_i)\Gamma_r(\delta)\Gamma_r(\beta)\Gamma_r(\gamma)}$, $(y_1,\ldots,y_n)\in T_n$ and $z\in\Omega_r\cap(I_r-\Omega_r)$.
Making the transformation

$$x_i = z^{\frac12}y_iz^{\frac12},\ \ i=1,\ldots,n$$

with the Jacobian $J(y_1,\ldots,y_n,z\to x_1,\ldots,x_n,z)=\det(z)^{-n\frac{r+1}{2}}$ in (11), we get the pdf of $(X_1,\ldots,X_n,Z)$ as

$$c\prod_{i=1}^{n}\det(x_i)^{\alpha_i-\frac{r+1}{2}}\det\left(z-\sum_{i=1}^{n}x_i\right)^{\delta-\frac{r+1}{2}}\det(z)^{-(\sum_{i=1}^{n}\alpha_i+\delta-\beta)}\det(I_r-z)^{\gamma-\frac{r+1}{2}}, \tag{12}$$

where $(x_1,\ldots,x_n)\in T_n$, $z\in\Omega_r\cap(I_r-\Omega_r)$ and $z-\sum_{i=1}^{n}x_i\in\Omega_r$.

Now, Transform z to $w=\left(I_r-\sum_{i=1}^{n}x_i\right)^{-\frac{1}{2}}(I_r-z)\left(I_r-\sum_{i=1}^{n}x_i\right)^{-\frac{1}{2}}$, with the Jacobian $J(z\to w)=\det\left(I_r-\sum_{i=1}^{n}x_i\right)^{\frac{r+1}{2}}$ in (12), the pdf of $(X_1,\ldots,X_n,W)$ is given by

$$c\prod_{i=1}^{n}\det(x_i)^{\alpha_i-\frac{r+1}{2}}\det\left(I_r-\sum_{i=1}^{n}x_i\right)^{\gamma+\delta-\frac{r+1}{2}}\det(w)^{\gamma-\frac{r+1}{2}}\det(I_r-w)^{\delta-\frac{r+1}{2}}\det\left(I_r-\left(I_r-\sum_{i=1}^{n}x_i\right)w\right)^{-(\sum_{i=1}^{n}\alpha_i+\delta-\beta)},$$

where $(x_1,\ldots,x_n)\in T_n$ and $w\in\Omega_r\cap(I_r-\Omega_r)$. Finally, Integrating out the variable w using (3) yields the desired result.

As corollary of this theorem, we have

Corollary 1. *Let $Y\sim B_r(\alpha,\delta)$ and $Z\sim B_r(\beta,\gamma)$ be independent. Then $X=Z^{\frac{1}{2}}YZ^{\frac{1}{2}}\sim BH_r(\alpha;\gamma,\alpha+\delta-\beta;\gamma+\delta)$.*

3 Properties

This section is devoted to the study of certain properties of the Dirichlet-hypergeometric distribution on symmetric matrices. The marginal distributions of this distribution are presented in the following theorem.

Theorem 3. *Let $(X_1,\ldots,X_n)\sim DH_r(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$. Then, for any integer $k=1,\ldots,n$,*

$$(X_1,\ldots,X_k)\sim DH_r(\alpha_1,\ldots,\alpha_k;\beta,\gamma;\sum_{i=k+1}^{n}\alpha_i+\delta).$$

Proof. We first find the pdf of $(X_1,\ldots,X_{n-1})$ by integrating x_n from the pdf of $(X_1,\ldots,X_n)$. Substituting

$$z=\left(I_r-\sum_{i=1}^{n-1}x_i\right)^{-\frac{1}{2}}x_n\left(I_r-\sum_{i=1}^{n-1}x_i\right)^{-\frac{1}{2}},$$

with the Jacobian $J(x_n\to z)=\det\left(I_r-\sum_{i=1}^{n-1}x_i\right)^{\frac{r+1}{2}}$ in (9), the pdf of $(X_1,\ldots,X_{n-1},Z)$ is obtained as

$$C(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)\prod_{i=1}^{n-1}\det(x_i)^{\alpha_i-\frac{r+1}{2}}\det\left(I_r-\sum_{i=1}^{n-1}x_i\right)^{\alpha_n+\delta-\frac{r+1}{2}}$$

$$\times\det(z)^{\alpha_n-\frac{r+1}{2}}\det(I_r-z)^{\delta-\frac{r+1}{2}}\,{}_2F_1\left(\beta,\gamma;\delta;\left(I_r-\sum_{i=1}^{n-1}x_i\right)(I_r-z)\right).$$

Integrate the previous expression with respect to z using (3), we obtain that the pdf of $(X_1,\ldots,X_{n-1})$ is given by

$$C(\alpha_1,\ldots,\alpha_{n-1},\beta,\gamma;\alpha_n+\delta)\prod_{i=1}^{n-1}\det(x_i)^{\alpha_i-\frac{r+1}{2}}\det\left(I_r-\sum_{i=1}^{n-1}x_i\right)^{\alpha_n+\delta-\frac{r+1}{2}}\,{}_2F_1\left(\beta,\gamma;\alpha_n+\delta;\left(I_r-\sum_{i=1}^{n-1}x_i\right)\right),$$

where $(x_1,\ldots,x_{n-1})\in T_{n-1}$. Performing this process $n-k$ times yields the pdf of $(X_1,\ldots,X_k)$.

Corollary 2. *Let $(X_1,\ldots,X_n)\sim DH_r(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$. Then, for any integer $k=1,\ldots,n$,*

$$X_k\sim BH_r(\alpha_k;\beta,\gamma;\alpha-\alpha_k+\delta),$$

where $\alpha=\sum_{i=1}^n\alpha_i$.

If $(X_1,\ldots,X_n)\sim D_r(\alpha_1,\ldots,\alpha_n;\alpha_{n+1})$, and define, for $i=k+1,\ldots,n$, $Y_i=\left(I_r-\sum_{i=1}^k X_i\right)^{-\frac{1}{2}}X_i\left(I_r-\sum_{i=1}^k X_i\right)^{-\frac{1}{2}}$. Then, it is well know that $(Y_{k+1},\ldots,Y_n)\sim D_r(\alpha_{k+1},\ldots,\alpha_n;\alpha_{n+1})$.In the subsequent theorem, we present an analogous result for the Dirichlet-hypergeometric distribution.

Theorem 4. *Let $(X_1,\ldots,X_n)\sim DH_r(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$. Define, for any integer $k=1,\ldots,n-1$,*

$$Y_i=\left(I_r-\sum_{i=1}^k X_i\right)^{-\frac{1}{2}}X_i\left(I_r-\sum_{i=1}^k X_i\right)^{-\frac{1}{2}},\quad i=k+1,\ldots,n.$$

Then, the pdf of $(Y_{k+1},\ldots,Y_n)$ is given by

$$K\prod_{i=k+1}^n\det(y_i)^{\alpha_i-\frac{r+1}{2}}\det\left(I_r-\sum_{i=k+1}^n y_i\right)^{\delta-\frac{r+1}{2}}{}_3F_2\left(\beta,\gamma;\sum_{i=k+1}^n\alpha_i+\delta;\delta,\sum_{i=1}^n\alpha_i+\delta;I_r-\sum_{i=k+1}^n y_i\right),$$

where $(y_{k+1},\ldots,y_n)\in T_{n-k}$ and $K=\dfrac{\Gamma_r\left(\sum_{i=1}^n\alpha_i+\delta-\beta\right)\Gamma_r\left(\sum_{i=1}^n\alpha_i+\delta-\gamma\right)\Gamma_r\left(\sum_{i=k+1}^n\alpha_i+\delta\right)}{\prod_{i=k+1}^n\Gamma_r(\alpha_i)\Gamma_r(\delta)\Gamma_r\left(\sum_{i=1}^n\alpha_i+\delta-\beta-\gamma\right)\Gamma_r\left(\sum_{i=1}^n\alpha_i+\delta\right)}$.

Proof. Making the transformation

$$y_i=\left(I_r-\sum_{i=1}^k x_i\right)^{-\frac{1}{2}}x_i\left(I_r-\sum_{i=1}^k x_i\right)^{-\frac{1}{2}},\ i=k+1,\ldots,n,$$

with the Jacobian $J(x_{k+1},\ldots,x_n\to y_{k+1},\ldots,y_n)=\det(I_r-\sum_{i=1}^k x_i)^{-(n-k)\frac{r+1}{2}}$ in (9) and integrating with respect to $x_1,\ldots,x_k$, we get the pdf of $(Y_{k+1},\ldots,Y_n)$ as

$$C(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)\prod_{i=k+1}^n\det(y_i)^{\alpha_i-\frac{r+1}{2}}\det\left(I_r-\sum_{i=k+1}^n y_i\right)^{\delta-\frac{r+1}{2}}I, \tag{13}$$

where $(y_{k+1},\ldots,y_n)\in T_{n-k}$ and

$$I=\int_{T_k}\prod_{i=1}^k\det(x_i)^{\alpha_i-\frac{r+1}{2}}\det\left(I_r-\sum_{i=1}^k x_i\right)^{\sum_{i=k+1}^n\alpha_i+\delta-\frac{r+1}{2}}{}_2F_1\left(\beta,\gamma;\delta;\left(I_r-\sum_{i=k+1}^n y_i\right)\left(I_r-\sum_{i=1}^k x_i\right)\right)\prod_{i=1}^k dx_i.$$

By evaluating I using (8) and (3), we obtain

$$I=\frac{\prod_{i=1}^k\Gamma_r(\alpha_i)\Gamma_r\left(\sum_{i=k+1}^n\alpha_i+\delta\right)}{\Gamma_r\left(\sum_{i=1}^n\alpha_i+\delta\right)}{}_3F_2\left(\beta,\gamma;\sum_{i=k+1}^n\alpha_i+\delta;\delta,\sum_{i=1}^n\alpha_i+\delta;I_r-\sum_{i=k+1}^n y_i\right). \tag{14}$$

Substituting (14) in (13) and simplifying using (10), the required result is obtained.

This next theorem determines the joint distribution of partial sums derived from random matrices that are distributed jointly according to a Dirichlet-hypergeometric distribution.

Theorem 5. *Let $(X_1,\ldots,X_n)\sim DH_r(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$ and $n_1,\ldots,n_l$ be non-negative integers such that $\sum_{i=1}^l n_i=n$. Denote, for $j=1,\ldots,l$,*

$$\alpha_{(j)}=\sum_{i=k_{j-1}+1}^{k_j}\alpha_i,\ \ k_0=0,\ \ k_j=\sum_{i=1}^j n_i.$$

Define, for $j = 1,\ldots,l,$

$$S_j = \sum_{i=k_{j-1}+1}^{k_j} X_i \ \textit{and} \ Y_i = S_j^{-\frac{1}{2}} X_i S_j^{-\frac{1}{2}}, \ i = k_{j-1}+1,\ldots,k_j-1.$$

Then

(i) $(S_1,\ldots,S_l)$ *and* $(Y_{k_{j-1}+1},\ldots,Y_{k_j-1})$, $j = 1,\ldots,l$, *are independent.*

(ii) $(Y_{k_{j-1}+1},\ldots,Y_{k_j-1}) \sim D_r(\alpha_{k_{j-1}+1},\ldots,\alpha_{k_j-1};\alpha_{k_j})$, $j = 1,\ldots,l$.

(iii) $(S_1,\ldots,S_l) \sim DH_r(\alpha_{(1)},\ldots,\alpha_{(l)};\beta,\gamma;\delta)$.

*Proof.*Substituting

$$s_j = \sum_{i=k_{j-1}+1}^{k_j} x_i \ \textit{and} \ y_i = s_j^{-\frac{1}{2}} x_i s_j^{-\frac{1}{2}}, \ i = k_{j-1}+1,\ldots,k_j-1, \ j = 1,\ldots,l,$$

with the Jacobian

$$\begin{aligned} J(x_1,\ldots,x_n \to y_1,\ldots,y_{n_1-1},s_1,\ldots,y_{k_{l-1}+1},\ldots,y_{n-1},s_l) &= \prod_{j=1}^{l} J(x_{k_{j-1}+1},\ldots,x_{k_j} \to y_{k_{j-1}+1},\ldots,y_{k_j-1},s_j) \\ &= \prod_{j=1}^{l} \det(s_j)^{(n_j-1)\frac{r+1}{2}} \end{aligned}$$

in the pdf of $(X_1,\ldots,X_n)$ given by (9), we get the joint pdf of $Y_{k_{j-1}+1},\ldots,Y_{k_j-1}$, S_j, where $j = 1,\ldots,l$, as

$$\begin{aligned} C(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta) \prod_{j=1}^{l} \det(s_j)^{\alpha_{(j)}-\frac{r+1}{2}} \det\left(I_r - \sum_{j=1}^{l} s_j\right)^{\delta-\frac{r+1}{2}} {}_2F_1\left(\beta,\gamma;\delta;I_r - \sum_{j=1}^{l} s_j\right) \\ \times \prod_{j=1}^{l} \left[\prod_{i=k_{j-1}+1}^{k_j-1} \det(y_j)^{\alpha_j-\frac{r+1}{2}} \det\left(I_r - \sum_{i=k_{j-1}+1}^{k_j-1} y_i\right)^{\alpha_{k_j}-\frac{r+1}{2}} \right] \end{aligned} \tag{15}$$

where $(s_1,\ldots,s_{l)}) \in T_l$ and $(Y_{k_{j-1}+1},\ldots,Y_{k_j-1}) \in T_{k_j-k_{j-1}-1}$, $j = 1,\ldots,l$. From the factorization in (15), it can be observed that $(S_1,\ldots,S_l)$ and $(Y_{k_{j-1}+1},\ldots,Y_{k_j-1})$, $j = 1,\ldots,l$, are distributed independently. Moreover, $(Y_{k_{j-1}+1},\ldots,Y_{k_j-1}) \sim D_r(\alpha_{k_{j-1}+1},\ldots,\alpha_{k_j-1};\alpha_{k_j})$, $j = 1,\ldots,l$, and $(S_1,\ldots,S_l) \sim DH_r(\alpha_{(1)},\ldots,\alpha_{(l)};\beta,\gamma;\delta)$.

Corollary 3.*Let* $(X_1,\ldots,X_n) \sim DH_r(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$. *Define*

$$S = \sum_{j=1}^{n} X_j \ \textit{and} \ Y_i = S^{-\frac{1}{2}} X_i S^{-\frac{1}{2}}, \ i = 1,\ldots,n-1.$$

Then $(Y_1,\ldots,Y_{n-1})$ *and S are independent,* $(Y_1,\ldots,Y_{n-1}) \sim D_r(\alpha_1,\ldots,\alpha_{n-1};\alpha_n)$ *and* $S \sim BH_r\left(\sum_{i=1}^{n} \alpha_i;\beta,\gamma;\delta\right)$.

Corollary 4.*Let* $(X_1,\ldots,X_n) \sim DH_r(\alpha_1,\ldots,\alpha_n;\alpha,\gamma;\delta)$. *Then, for any integer* $k = 1,\ldots,n-1$,

$$\left(\sum_{i=1}^{n} X_i\right)^{-\frac{1}{2}} \left(\sum_{i=1}^{k} X_i\right) \left(\sum_{i=1}^{n} X_i\right)^{-\frac{1}{2}} \sim B_r\left(\sum_{i=1}^{k} \alpha_i, \sum_{i=k+1}^{n} \alpha_i\right).$$

Next, we derive a factorization of the Dirichlet-hypergeometric distribution on V_r.

Theorem 6.*Let* $(X_1,\ldots,X_n) \sim DH_r(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$. *Define*

$$Y_n = \sum_{i=1}^{n} X_i \ \textit{and} \ Y_k = \left(\sum_{i=1}^{k+1} X_i\right)^{-\frac{1}{2}} \left(\sum_{i=1}^{k} X_i\right) \left(\sum_{i=1}^{k+1} X_i\right)^{-\frac{1}{2}}, \ k = 1,\ldots,n-1.$$

Then the random matrices $Y_1,\ldots,Y_n$ *are independent,* $Y_k \sim B_r\left(\sum_{i=1}^{k} \alpha_i, \alpha_{k+1}\right)$, $k = 1,\ldots,n-1$, *and* $Y_n \sim BH_r\left(\sum_{i=1}^{n} \alpha_i;\beta,\gamma;\delta\right)$.

*Proof.*The transformation mentioned above makes it possible to observe that

$$\det(x_k) = \det(I_r - y_{k-1}) \prod_{i=k}^{n} \det(y_i), \quad k = 2, \dots, n$$

and

$$\det(x_1) = \prod_{i=1}^{n} \det(y_i).$$

Substituting this with the Jacobian of transformation

$$J(x_1, \dots, x_n \longrightarrow y_1, \dots, y_n) = \prod_{k=2}^{n} \det(y_k)^{\frac{r+1}{2}}$$

in the density (9), one obtains the pdf of $(Y_1, \dots, Y_n)$ as

$$C(\alpha_1, \dots, \alpha_n; \beta, \gamma; \delta) \prod_{k=1}^{n-1} \left\{ \det(y_k)^{\sum_{i=1}^{k} \alpha_i - \frac{r+1}{2}} \det(I_r - y_k)^{\alpha_{k+1} - \frac{r+1}{2}} \right\}$$
$$\times \left\{ \det(y_n)^{\sum_{i=1}^{n} \alpha_i - \frac{r+1}{2}} \det(I_r - y_n)^{\delta - \frac{r+1}{2}} \; {}_2F_1\left(\beta, \gamma; \delta; (I_r - y_n)^{-1}\right) \right\},$$

where $y_k \in \Omega_r \cap (I_r - \Omega_r)$, $k = 1, \dots, n$. Now, observing that

$$C(\alpha_1, \dots, \alpha_n; \beta, \gamma; \delta) = C\left(\sum_{i=1}^{n} \alpha_i; \beta, \gamma; \delta\right) \frac{\Gamma_r\left(\sum_{i=1}^{n} \alpha_i\right)}{\prod_{i=1}^{n} \Gamma_r(\alpha_i)}$$

$$= C\left(\sum_{i=1}^{n} \alpha_i; \beta, \gamma; \delta\right) \prod_{k=1}^{n-1} \frac{\Gamma_r\left(\sum_{i=1}^{k+1} \alpha_i\right)}{\Gamma_r(\alpha_{k+1}) \Gamma_r\left(\sum_{i=1}^{k} \alpha_i\right)},$$

we get the desired result.

We will now calculate the moments of various functions associated with the random matrices that jointly follow a Dirichlet-hypergeometric distribution on V_r.

Theorem 7.*Let $(X_1, \dots, X_n) \sim DH_r(\alpha_1, \dots, \alpha_n; \beta, \gamma; \delta)$. Then*

(i) $E\left[\prod_{i=1}^{n} \det(X_i)^{p_i}\right] = \prod_{i=1}^{n} \frac{\Gamma_r(\alpha_i + p_i)}{\Gamma_r(\alpha_i)} \frac{\Gamma_r(\alpha + \delta - \beta)}{\Gamma_r(\alpha + \delta - \beta + p)} \frac{\Gamma_r(\alpha + \delta - \gamma)}{\Gamma_r(\alpha + \delta - \gamma + p)} \frac{\Gamma_r(\alpha + \delta - \beta - \gamma + p)}{\Gamma_r(\alpha + \delta - \beta - \gamma)},$

where $p_i \geq 0$, $i = 1, \dots, n$, $\alpha = \sum_{i=1}^{n} \alpha_i$ and $p = \sum_{i=1}^{n} p_i$.

(ii) $E\left[\det\left(I_r - \sum_{i=1}^{n} X_i\right)^h\right] = \frac{\Gamma_r(\delta + h)}{\Gamma_r(\alpha + \delta + h)} \frac{\Gamma_r(\alpha + \delta - \beta) \Gamma_r(\alpha + \delta - \gamma)}{\Gamma_r(\delta) \Gamma_r(\alpha + \delta - \beta - \gamma)} \; {}_3F_2(\delta + h, \alpha, \gamma; \alpha + \delta + h, \delta; I_r),$

where $h \geq 0$ and $\alpha = \sum_{i=1}^{n} \alpha_i$.

Proof.(i) Using (9) and (8), we obtain

$$E\left[\prod_{i=1}^{n} \det(X_i)^{p_i}\right] = C(\alpha_1, \dots, \alpha_n; \beta, \gamma; \delta) \frac{\prod_{i=1}^{n} \Gamma_r(\alpha_i + p_i)}{\Gamma_r(\alpha + p)} J, \tag{16}$$

where $\alpha = \sum_{i=1}^{n} \alpha_i$, $p = \sum_{i=1}^{n} p_i$ and

$$J = \int_{\Omega_r \cap (I_r - \Omega_r)} \det(x)^{\alpha + p - \frac{r+1}{2}} \det(I_r - x)^{\delta - \frac{r+1}{2}} \; {}_2F_1(\beta, \gamma; \delta; I_r - x) dx.$$

Now, using (5) and (4), the integral J is equal to

$$J = \frac{\Gamma_r(\alpha+p)\Gamma_r(\delta)\Gamma_r(\alpha+\delta+p-\alpha-\gamma)}{\Gamma_r(\alpha+\delta+p-\alpha)\Gamma_r(\alpha+\delta+p-\gamma)}. \tag{17}$$

Substituting (17) in (16) and simplifying by using (10) yields the desired result.
(ii) By following the same steps, we obtain

$$E\left[\det\left(I_r-\sum_{i=1}^{n}X_i\right)^h\right] = C(\alpha_1,\dots,\alpha_n;\beta,\gamma;\delta)\frac{\prod_{i=1}^{n}\Gamma_r(\alpha_i)}{\Gamma_r(\alpha)}\int_{\Omega_r\cap(I_r-\Omega_r)}\det(x)^{\alpha-\frac{r+1}{2}}\det(I_r-x)^{\delta+h-\frac{r+1}{2}}\,{}_2F_1(\beta,\gamma;\delta;I_r-x)dx,$$

The desired result is obtained by evaluating the integral above using (5).

4 Related distributions

Within the framework of related distributions to the Dirichlet-hypergeometric distribution, two notable extensions are the inverted Dirichlet-hypergeometric distribution and the ordered Dirichlet-hypergeometric distribution. These distributions emerge from specific transformations or considerations of the Dirichlet-hypergeometric distribution and are designed to address diverse data and modelling requirements.

4.1 The inverted Dirichlet-hypergeometric distribution on symmetric matrices

In this subsection, we introduce the inverted Dirichlet-hypergeometric distribution on symmetric matrices.

Theorem 8.*Let* $(X_1,\dots,X_n)\sim DH_r(\alpha_1,\dots,\alpha_n;\beta,\gamma;\delta)$. *Define*

$$U_i = \left(I_r-\sum_{j=1}^{n}X_j\right)^{-\frac{1}{2}} X_i \left(I_r-\sum_{j=1}^{n}X_j\right)^{-\frac{1}{2}}, \quad i=1,\dots,n.$$

Then the pdf of $(U_1,\dots,U_n)$ *is given by*

$$C(\alpha_1,\dots,\alpha_n;\beta,\gamma;\delta)\prod_{i=1}^{n}\det(u_i)^{\alpha_i-\frac{r+1}{2}}\det\left(I_r+\sum_{i=1}^{n}u_i\right)^{-(\sum_{i=1}^{n}\alpha_i+\delta)}{}_2F_1\left(\beta,\gamma;\delta;\left(I_r+\sum_{i=1}^{n}u_i\right)^{-1}\right), \tag{18}$$

where $u_i\in\Omega_r$, $i=1,\dots,n$.

*Proof.*The pdf of $(X_1,\dots,X_n)$ is given by (9). Making the transformation

$$u_i = \left(I_r-\sum_{j=1}^{n}x_j\right)^{-\frac{1}{2}} x_i \left(I_r-\sum_{j=1}^{n}x_j\right)^{-\frac{1}{2}}, \quad i=1,\dots,n, \tag{19}$$

so that

$$x_i = \left(I_r+\sum_{j=1}^{n}u_j\right)^{-\frac{1}{2}} u_i \left(I_r+\sum_{j=1}^{n}u_j\right)^{-\frac{1}{2}}, \quad i=1,\dots,n.$$

The Jacobian of this transformation is

$$\begin{aligned} J = J(x_1,\dots,x_n \longrightarrow u_1,\dots,u_n) &= J(x_1,\dots,x_{n-1},x_n \longrightarrow x_1,\dots,x_{n-1},z) \\ &\times J(x_1,\dots,x_{n-1},z \longrightarrow u_1,\dots,u_{n-1},z) \\ &\times J(u_1,\dots,u_{n-1},z \longrightarrow u_1,\dots,u_{n-1},u_n), \end{aligned}$$

where

$$z = I_r - \sum_{j=1}^{n} x_j, \quad u_i = z^{-\frac{1}{2}} x_i z^{-\frac{1}{2}}, \quad i = 1, \ldots, n-1$$

and

$$u_n = z^{-1} - \left(I_r + \sum_{j=1}^{n-1} u_j \right).$$

Using the results given in Deemer and Olkin [25], we obtain

$$J(x_1, \ldots, x_{n-1}, x_n \longrightarrow x_1, \ldots, x_{n-1}, z) = 1$$

$$J(x_1, \ldots, x_{n-1}, z \longrightarrow u_1, \ldots, u_{n-1}, z) = \det \left(I_r + \sum_{j=1}^{n} u_j \right)^{-\frac{r+1}{2}(n-1)}$$

$$J(u_1, \ldots, u_{n-1}, z \longrightarrow u_1, \ldots, u_{n-1}, u_n) = \det \left(I_r + \sum_{j=1}^{n} u_j \right)^{-(r+1)}.$$

Then

$$J = \det \left(I_r + \sum_{j=1}^{n} u_j \right)^{-\frac{r+1}{2}(n+1)}.$$

Substituting this and (19) in (9), we obtain that the pdf of $(U_1, \ldots, U_n)$ is (18).

Definition 2. *The distribution of $(U_1, \ldots, U_n)$ given by (18) is called the inverted Dirichlet-hypergeometric distribution on V_r with parameters $(\alpha_1, \ldots, \alpha_n; \beta, \gamma; \delta)$, and is denoted by $IDH_r(\alpha_1, \ldots, \alpha_n; \beta, \gamma; \delta)$.*

When $n = 1$, that is when we have the beta-hypergeometric random matrix X with parameters $(\alpha; \beta, \gamma; \delta)$, the random matrix $U = (I_r - X)^{-\frac{1}{2}} X (I_r - X)^{-\frac{1}{2}}$ has the distribution

$$\frac{\Gamma_r(\delta + \alpha - \beta)\Gamma_r(\delta + \alpha - \gamma)}{\Gamma_r(\delta)\Gamma_r(\alpha)\Gamma_r(\delta + \alpha - \beta - \gamma)} \det(u)^{\alpha - \frac{r+1}{2}} \det(I_r + u)^{-(\alpha+\delta)} \, {}_2F_1\left(\beta, \gamma; \delta; (I_r + u)^{-1}\right) \mathbf{1}_{\Omega_r}(u), \tag{20}$$

which we call the inverted beta-hypergeometric distribution on V_r with parameters $(\alpha; \beta, \gamma; \delta)$ and is denoted by $IBH_r(\alpha; \beta, \gamma; \delta)$.
For $\beta = \delta$, the inverted beta-hypergeometric distribution $IBH_r(\alpha; \beta, \gamma; \delta)$ reduces to the inverted beta distribution $IB_r(\alpha - \gamma, \beta)$.

The specific case where $r = 1$, $\beta = \delta$ and $\gamma = 0$ corresponds to the real inverted Dirichlet distribution which has numerous properties documented within the statistical literature and a lot of applications across diverse fields, including statistics, finance, ecology and genetics. See, for example, Ben Farah [26], Otto et al. [27] and Ling et al. [28].

Theorem 9. *Let $(U_1, \ldots, U_n) \sim IDH_r(\alpha_1, \ldots, \alpha_n; \beta, \gamma; \delta)$. Define*

$$X_i = \left(I_r + \sum_{j=1}^{n} U_j \right)^{-\frac{1}{2}} U_i \left(I_r + \sum_{j=1}^{n} U_j \right)^{-\frac{1}{2}}, \quad i = 1, \ldots, n.$$

Then $(X_1, \ldots, X_n) \sim DH_r(\alpha_1, \ldots, \alpha_n; \beta, \gamma; \delta)$.

Proof. The proof is similar to that of Theorem 8.

Corollary 5. *Let $U \sim IBH_r(\alpha; \beta, \gamma; \delta)$ and define*

$$X = (I_r + U)^{-\frac{1}{2}} U (I_r + U)^{-\frac{1}{2}}.$$

Then $X \sim BH_r(\alpha; \beta, \gamma; \delta)$.

4.2 The ordered Dirichlet-hypergeometric distribution on symmetric matrices

The ordered Dirichlet-hypergeometric distribution is a closely related distribution to the Dirichlet-hypergeometric distribution.

Definition 3.*Let* $(X_1,\ldots,X_n) \sim DH_r(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$. *Define*

$$S_i = \sum_{j=1}^{i} X_j, \;\; i = 1,\ldots,n.$$

Then the distribution of $(S_1,\ldots,S_n)$ *is called the ordered Dirichlet-hypergeometric distribution on* V_r *with parameters* $(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$*, denoted by* $ODH_r(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)$*. Its pdf is given by*

$$C(\alpha_1,\ldots,\alpha_n;\beta,\gamma;\delta)\prod_{i=1}^{n-1}\det(s_{i+1}-s_i)^{\alpha_{i+1}-\frac{r+1}{2}}\det(s_1)^{\alpha_1-\frac{r+1}{2}}\det(I_r-s_n)^{\delta-\frac{r+1}{2}}\;{}_2F_1(\beta,\gamma;\delta;I_r-s_n), \tag{21}$$

where $s_1 \in \Omega_r$*,* $s_{i+1}-s_i \in \Omega_r$*,* $i=1,\ldots,n-1$ *and* $s_n \in \Omega_r \cap (I_r-\Omega_r)$*.*

5 Conclusion

The Dirichlet-hypergeometric distribution on symmetric matrices provides a robust framework for modeling complex multivariate data structures, particularly in the context of covariance matrices and other symmetric positive definite matrices. This distribution extends the classical Dirichlet distribution, allowing for greater flexibility in capturing relationships among variables and accommodating various constraints inherent to symmetric matrices. The theoretical properties established in this paper, including marginal distributions and conditional distributions, distribution of partial sums, moments, and relationships to other distributions, highlight its utility in statistical applications such as Bayesian inference, machine learning, and multivariate analysis. Future work may focus on computational methods for estimation and inference, as well as applications in diverse fields such as finance, bioinformatics, and network analysis, where understanding the underlying structure of data is crucial. Overall, the Dirichlet-hypergeometric distribution on symmetric matrices represents a significant advancement in the field of multivariate statistics, offering new avenues for research and practical applications.

Declarations

Competing interests: The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Funding: There's no funding for this work.
Availability of data and materials: No data was used for the research described in the article.
Acknowledgments: The author thanks the referee for his helpful comments and suggestions.

References

[1] J. Aitchison, *The Statistical Analysis of Compositional Data*, London: The Blackburn Press (2003).
[2] J. Aitchison, *The Analysis of Compositional Data*, Chapman and Hall, London, (1986).
[3] K.T. Fang, S. Kotz, and K.W. Ng, *Symmetric Multivariate and Related Distributions*, Chapman and Hall, London (1990).
[4] M-L .Tang, Q. Wu, S. Yang and G-L. Tian, Dirichlet composition distribution for compositional data with zero components: An application to fluorescence in situ hybridization (FISH) detection of chromosome, *Biometrical Journal*, **64**(4)(2022), 714 − 732.
[5] M. Tsagris and C. Stewart, A dirichlet regression model for compositional data with zeros, *Lobachevskii Journal of Mathematic*, **39**(2018), 398 − 412.
[6] N. Balakrishnan, and V.B. Nevzorov, *A Primer on Statistical Distributions*, John Wiley and Sons, New York (2003).
[7] R.J. Connor and J.E. Mosimann, Concepts of Independence for Proportions with a Generalization of the Dirichlet Distribution, *Journal of the American Statistical Association*, **64** (1969), 194 − 206.
[8] D.L. Libby and M. R. Novick, Multivariate generalized beta distributions with applications to utility assessment, *Journal of educational statistics*, **7**(4)(1982), 271 − 294.

[9] J.J. Chen and M.R. Novick, Bayesian analysis for binomial models with generalized beta prior distributions, *Journal of Educational Statistics*, **9**(2)(1984), 163 – 175.
[10] G.S. Monti, G. Mateu-Figueras and V. Pawlowsky-Glahn, Notes on the scaled Dirichlet distribution, *Compositional data analysis*, (2011), 128 – 138.
[11] A.W. Marshall, I. Olkin, and B.C. Arnold, *Inequalities: theory of majorization and its applications*, Springer (1979).
[12] B. Bhattacharya, Maximum entropy characterizations of the multivariate Liouville distributions, *Journal of multivariate analysis*, **97**(6)(2006), 1272 – 1283.
[13] W.S. Rayens and C. Srinivasan, Dependence Properties of Generalized Liouville Distributions on the Simplex, *Journal of the American Statistical Association*, **89** (428)(1994), 1465 – 1470.
[14] R. KS Hankin, A generalization of the Dirichlet distribution*Journal of Statistical Software*, **33** (2010), 1 – 18.
[15] D.K. Nagar, P.A. Bran-Cardona and A.K. Gupta, Multivariate Generalization of the Hypergeometric Function Type I Distribution, *Acta Appl. Math.*, **105** (2009), 111 – 122.
[16] Y.L. Luke, *Special Functions and Their Approximations: v. 2*, Academic press, (1969).
[17] A.G. Constantine, Some non-central distribution problems in multivariate analysis, *The Annals of Mathematical Statistics*, **34**(4) (1963), 1270 – 1285.
[18] A.T. James, Distributions of matrix variates and latent roots derived from normal samples, *The Annals of Mathematical Statistics*, **35**(2) (1964), 475 – 501.
[19] M. Ben Farah and A. Hassairi, Characterization of the Dirichlet distribution on symmetric matrices, *Statistics & probability letters*, **77**(4) (2007), 357 – 364.
[20] M. Ben Farah and A. Hassairi, On the Dirichlet distribution on symmetric matrices, *J. Stat. Plann. Inf.*, **139** (2009), 2559 – 2570.
[21] M. Ben Farah and M. Ghorbel, Dirichlet partition on symmetric matrices, *Indian J. Pure Appl. Math.*, **46**(1) (2015), 73 – 83.
[22] A.K. Gupta and D.K. Nagar, *Matrix Variate Distributions*, Chapman and Hall/CRC, Boca Raton, FL, (2000).
[23] I. Olkin and H. Rubin, Multivariate Beta Distributions and Independence Properties of the Wishart Distribution, *Ann. Math. Statist*, **35** (1964), 261 – 269.
[24] R.D. Gupta and D.ST.P. Richards, Multivariate liouville distributions, *Journal of Multivariate Analysis*,**23**(2) (1987), 233 – 256.
[25] W.L. Deemer and I. OIkin, The Jacobians of certain matrix transformations useful in multivariate analysis, *Biometrika*, **38** (1951), 345 – 367.
[26] M. Ben Farah, Characterization of the Three-Variate Inverted Dirichlet Distributions, *Armenian Journal of Mathematics*, **12**(15) (2023), 1 – 9.
[27] A.F. Otto, J.T. Ferreira, A. Bekker, A. Punzo and S.D. Tomarchio, A refreshing take on the inverted Dirichlet via a mode parameterization with some statistical illustrations, *Journal of the Korean Statistical Society*, **54** (2025), 314 – 341.
[28] M.H. Ling, N. Balakrishnan and S.J. Bae, On the application of inverted Dirichlet distribution for reliability inference of completely censored components with dependence structure, *Computers and Industrial Engineering*, **196** (2024), 110452.