Hypergeometric Functions: From One Scalar Variable to Several Matrix Arguments, in Statistics and Beyond

Ayush Yajnik, Vibhor Gupta

New Delhi-110017, India

Abstract: ***Hypergeometric functions have been increasingly present in several disciplines including Statistics, but there is much confusion on their proper uses, as well as on their existence and domain of definition. In this article, we try to clarify several points and give a general overview of the topic, going from the univariate case to the matrix case, in one and then in several arguments. We also survey some results in fields close to Statistics, where hypergeometric functions are actively used, studied and developed.***

Keywords: Hypergeometric, Zonal Polynomial, Fractional Calculus

1. Hypergeometric Functions in Matrix Arguments

Three Proposed Approaches

In multivariate Analysis variables encountered can be matrices, which will be arguments of hypergeometric functions.

2. Functions in One Matrix Variate

In going from a scalar variable to a matrix, there are several difficulties to define the hypergeometric function. First, functions of matrices, square or rectangular, can only be defined under certain conditions (Higham), and they can be scalar-valued, or matrix-valued. Secondly, for scalar-valued matrix functions, they are usually based on symmetric functions of the matrix entries, or of the eigenvalues of the input square matrices. A simple introduction to this topic is given by Pham-Gia and Turkkan. We recall here some basic notions of calculus on matrices, that are not so obvious.

Domain of integration: Let $f(X)$ be a scalar function of the matrix X. Then $\int_{\Omega} f(X)\mathrm{d}X$ is the iterated integral of $f(X)$ for each entry of X separately, over the region Ω located within the space defined by the simplex bounding the ranges of the entries of X.

Since it is usually very difficult to carry out direct integration over a complex region Ω, integration on simple regions are frequently done by changes of variables, matrix decompositions, and finally identification with known expressions.

We have also the region $0 < X < I_m$ as the set of all square matrices such that X and $I_m - X$ are positive definite, which reduces to the continuous variable x being between 0 and 1 in a uni-dimensional space.

Jacobian and Exterior product: In carrying out the required changes of variables mentioned above we have to use jacobians, and using wedge products $\mathrm{d}x_1 \wedge \mathrm{d}x_2 = -\mathrm{d}x_2 \wedge \mathrm{d}x_1$ and exterior forms would be helpful. We have, for example, for $I = \int_A f(x_1,\cdots,x_m)\mathrm{d}x_1 \wedge\cdots\wedge \mathrm{d}x_m$ and transforms $x_i = x_i(y_1,\cdots,y_m), i = 1,\cdots,m$, the result $I = \int_A g(y_1,\cdots,y_m)\mathrm{d}y_1 \wedge\cdots\wedge \mathrm{d}y_m$, with $\bigwedge_{i=1}^{m} \mathrm{d}x_i = \det\left(\frac{\partial x_i}{\partial y_j}\right)\bigwedge_{i=1}^{m}\mathrm{d}y_i$ where the jacobian of the transformation is the absolute value of the determinant $\det\left(\frac{\partial x_i}{\partial y_j}\right)$.

The multigamma function:

Let $f(X) = etr(-X)|X|^{a-\frac{m+1}{2}}$, where $etr(X)$ is the exponential ofthe trace of X, with the domain of positive definite matrices being $\Omega = \{X : X > 0\}$, we have the multivariate gamma function

$$\Gamma_m(a) = \int_{X>0} etr(-X)|X|^{a-(m+1)/2}\,\mathrm{d}X.$$

Carrying out integration as explained above, we obtain a product of m ordinary gamma functions

$$\Gamma_m(a) = \pi^{m(m-1)/2}\prod_{i=1}^{m}\Gamma\left(a-\frac{i-1}{2}\right).$$

The Matrix Laplace Transform: Let $f(S)$ be a scalar function of the positive definite symmetric $m\times m$ matrix S. Its Laplace transform is defined by

$$g(Z)=\int_{S>0} etr(-ZS)f(S)\mathrm{d}S,\ Z=X+iY$$

symmetric.
We assume that the integral converges in the half-plane $\mathrm{Re}(Z)=X>X_0$, for some positive definite matrix X_0. Then $g(Z)$ is analytic in Z in the half-plane. If $\int|g(X+iY)|\mathrm{d}y<\infty$ and $\lim_{X\to\infty}\int|g(X+iY)|\mathrm{d}Y=0$, then the inverse Laplace transform is:

$$f(S)=\frac{1}{(2\pi i)^{\frac{m(m+1)}{2}}}\int_{\mathrm{Re}(Z)>X_0>0} etr(XZ)g(Z)\mathrm{d}Z$$
.

To define hypergeometric functions in one matrix argument, there are three approaches offered in the literature.

3. Laplace Transform Approach

This approach was pioneered by Bochner, developed by Herz , and uses the matrix forms of (10) and (11). We can then define ${}_pF_{q+1}(\Lambda)$ and ${}_pF_{q+1}(\Lambda)$. More precisely, we define ${}_pF_q\left[\begin{matrix}a_1,\cdots,a_p\\ b_1,\cdots,b_q\end{matrix};\Lambda\right]$ in a progressive way, with

$${}_{p+1}F_q\left[\begin{matrix}a_1,\cdots,a_p,\gamma\\ b_1,\cdots,b_q\end{matrix};-Z^{-1}\right]$$

$$=\det(Z)^{\gamma}\frac{1}{\Gamma_m(\gamma)}\int_{\Lambda>0} etr(-\Lambda Z)\cdot{}_pF_q\left[\begin{matrix}a_1,\cdots,a_p\\ b_1,\cdots,b_q\end{matrix};-\right.$$

. (25)

$${}_pF_{q+1}\left[\begin{matrix}a_1,\cdots,a_p\\ b_1,\cdots,b_q,\gamma\end{matrix};-\Lambda\right]$$

$$=\det(\Lambda)^{(m+1)/2-\gamma}\frac{\Gamma_m(\gamma)}{(2\pi i)^{m(m+1)/2}}\int_{\mathrm{Re}\,Z=X_0>0} etr($$

(26)

Here, m is the dimension of the matrices and in (25). Also, for the multivariate Laplace transform, the elements off-diagonal of Z are taken as $z_{ij}/2$. So, theoretically, hypergeometric functions can be defined in this way, and sometimes they can be computed by numerical methods.

<u>Zonal Polynomials Approach</u>
This approach was introduced by James, and developed by James and Constantine. It is based on group representation using matrices, aimed at replacing x^n of the scalar case, by a polynomial $C(X)$, when x is replaced by the random matrix X. $C(X)$ is called the zonal polynomial of X. We have, for example, instead of the multinomial form

$$[tr(X)]^k=(x_1+\cdots+x_m)^k=\sum_{\sum k_i=k}\frac{m!}{k_1!\cdots k_m!}x_1^{k_1}\cdots x_m^{k_m}$$

,

the expression $[tr(X)]^k=\sum_{\kappa}C_\kappa(X)$, where the zonal polynomial $C_\kappa(X)$ is a symmetric homogeneous polynomial of degree k in elements of X. Here, κ is the partition $(k_1,\cdots,k_m)$, with $k_1\geq k_2\geq\cdots\geq k_m\geq 0$, and $\sum_{i=1}^m k_i=k$. V_k isthe vector space of homogeneous polynomials of degree k in the $m(m+1)/2$ elements of the symmetric $m\times m$ matrix X, and $V_k=\oplus_\kappa V_\kappa$, i.e. V_κ is the direct sum of irreducible invariant subspaces V_κ in the representation of the real linear group $Gl(m,\mathbb{R})$ in the vector space V_κ.

When $m=1$ we have indeed $x^k=C_\kappa(x)$ and hence, zonal polynomials of a matrix are similar to powers of a scalar variable.
The decomposition into a direct sum of subrings is assured by ring theory and hence, zonal polynomials do exist. However, their values must be obtained by solving a differential equation of Laplace-Beltrami type

$$\Delta_Y C_\kappa(Y)=\alpha_\kappa C_\kappa(Y),\ \Delta_Y=\sum_{i=1}^m y_i^2\frac{\partial^2}{\partial y_i^2}+\sum_{i=1}^m\sum_{j=1,j\neq i}^m\frac{y_i^2}{y_i-y_j}\frac{\partial}{\partial y_i}$$

, which quickly becomes difficult to track. More precisely, we have:

$$\Delta_Y C_\kappa(Y) = \left[\rho_\kappa + k(m-1)\right] C_\kappa(Y),$$

with $\rho_\kappa = \sum_{i=1}^{m} k_i(k_i - i)$.

Alternately, we can obtain $C_\kappa(Y) = \sum_{\lambda \le \kappa} c_{\kappa,\lambda} M_\lambda(Y)$, where the monomial symmetric functions are $M_\lambda(Y) = \sum y_{i_1}^{k_1} \cdots y_{i_p}^{k_p}$ and the coefficients

$$c_{\kappa,\lambda} = \sum_{\lambda < \mu \le \kappa} \frac{\left[(l_i + t) - (l_j - t)\right]}{\rho_\kappa - \rho_\lambda} c_{\kappa,\mu},$$

$\lambda = (l_1, \cdots, l_m)$,

$\mu = (l_1, \cdots, l_i + t, \cdots, l_j - t, \cdots, l_m)$,

$t = (1, 2, \cdots, l_j)$.

For $k = 3$, for example, we have the values of $c_{\kappa,\lambda}$ as follows:

$$\begin{array}{cc|ccc} & & & \lambda & \\ & & (3) & (2,1) & (1,1,1) \\ \hline & (3) & 1 & 3/5 & 2/5 \\ \kappa & (2,1) & 0 & 12/5 & 18/5 \\ & (1,1,1) & 0 & 0 & 2 \end{array}$$

Other methods, not necessarily simpler, have been suggested. Values of $C_\kappa(Y)$ up to $k = 20$ are found by researchers. We have some basic results on integration associated with zonal polynomials, as follows:

$$\int_{O(m)} C_\kappa(XHYH')\,\mathrm{d}H = \frac{C_\kappa(X)C_\kappa(Y)}{C_\kappa(I_m)},$$

$$\int_{X>0} etr(-XZ)|X|^{a-\frac{m+1}{2}} C_\kappa(XY)\,\mathrm{d}X = (a)_\kappa \Gamma_m(a) C_\kappa(YZ^{-1}) \big/ |Z|^a, \quad (28)$$

and

$$\int_{0<X<I_m} |X|^{a-\frac{m+1}{2}} |I_m - X|^{b-\frac{m+1}{2}} C_\kappa(XY)\,\mathrm{d}X = \frac{(a)_\kappa}{(a+b)_\kappa} \frac{\Gamma_m(a)\Gamma_m(b)}{\Gamma_m(a+b)} C_\kappa(Y), \quad (29)$$

with

$$\Gamma_m(a) = \pi^{m(m-1)/4} \prod_{j=1}^{m} \Gamma\left(a - \frac{j-1}{2}\right)$$

and

$$(\alpha)_\kappa = \prod_{j=1}^{m} \left(\alpha - \frac{j-1}{2}\right)_{k_j}. \quad (30)$$

A hypergeometric functions of one matrix X then have the familiar form:

$${}_pF_q(a_1, \cdots, a_p; b_1, \cdots, b_q; Z) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_\kappa \cdots (a_p)_\kappa}{(b_1)_\kappa \cdots (b_q)_\kappa} \frac{C_\kappa(Z)}{k!} \quad (31)$$

and we have

$${}_2F_1(a; b; Z) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a)_\kappa (b)_\kappa}{(c)_\kappa} \frac{C_\kappa(Z)}{k!}.$$

Like the scalar variable case (see (9)), using zonal polynomials, we have the Euler-type representation:

$$
{}_{p+1}F_{q+1}\left[\begin{matrix} a_1,\cdots,a_p,c \\ b_1,\cdots,b_q,d \end{matrix};\Lambda\right]
= \frac{\Gamma_m(d)}{\Gamma_m(c)\Gamma_m(d-c)} \int_{0<T<I_m} |T|^{c-\frac{m+1}{2}} |I_m - T|^{d-c-\frac{m+1}{2}} {}_pF_q\left[\begin{matrix} a_1,\cdots,a_p \\ b_1,\cdots,b_q \end{matrix};T\Lambda\right] \mathrm{d}T . \quad (32)
$$

Similarly, again using zonal polynomials, the Laplace and inverse Laplace representations of ${}_pF_{q+1}$ in the scalar variable case can be extended to the matrix case, and we can prove (25) and (26).

This zonal polynomials approach is favored when we aim at deriving theoretical results, using and obtaining expressions similar to the scalar case. Since higher order zonal polynomials are difficult to obtain we have here a topic still under development. It is worth mentioning that numerical computations have been carried out successfully for low values of p and q only . Several breakthroughs are due to James and Constantine and Muirhead , as already mentioned. Contemporary research relies heavily on their results.

4. Matrix-Transforms Approach

Mathai introduced the M-Transform method, which can establish several relations between hypergeometric functions, by using the fact that Laplace transforms are unique. It is based on the Weyl fractional integral, and a function $f(X)$ is, by definition, a (r,s)-hypergeometric function, i.e.

$$f(Z) = {}_rF_s(a_1,\cdots,a_r;b_1,\cdots,b_s;Z)$$

if its M-transform, i.e. $\int_{Z=Z'>0} |Z|^{\rho-\frac{m+1}{2}} \cdot f(-Z)\mathrm{d}Z$,

is of the form $\dfrac{\prod_{j=1}^{s}\Gamma_m(b_j)\prod_{j=1}^{r}\Gamma_m(a_j-\rho)}{\prod_{j=1}^{r}\Gamma_m(a_j)\prod_{j=1}^{s}\Gamma_m(b_j-\rho)}\Gamma_m(\rho)$, with ρ arbitrary such that the above expression on gammas exists.

Similarly, the Lauricella function $F_A(\cdot)$ in n matrix arguments can be defined as the function that can be represented as a n-fold integral, i.e

$$
F_A(a;b_1,\cdots,b_n;c_1,\cdots,c_n;X_1,\cdots,X_n)
= K\int_0^{I_m}\cdots\int_0^{I_m} |U_1|^{b_1-\frac{m+1}{2}}\cdots|U_n|^{b_n-\frac{m+1}{2}} |I_m-U_1|^{c_1-b_1-\frac{m+1}{2}}\cdots|I_m-U_n|^{c_n-b_n-\frac{m+1}{2}} \left|I_m - X_1^{1/2}U_1X_1^{1/2} - \cdots - X_n^{1/2}U_nX_n^{1/2}\right|^{-a} \mathrm{d}U_1\cdots\mathrm{d}U_n,
$$

where

$$K = \frac{\prod_{j=1}^{n}\Gamma_m(c_j)}{\prod_{j=1}^{n}\left\{\Gamma_m(b_j)\Gamma_m(c_j-b_j)\right\}}.$$

Mathai [45] was able to define most hypergeometric functions of matrix arguments, including H and G, with this approach, which is favored when we seek pure theoretical results only, since numerical computations seem quite difficult to undertake.