Matrix Distributions under Classical Group Actions

Haoming Wang

Sun Yat-sen University

October 19, 2025

Background I

- With the development of science and technology and the accumulation of production materials, distributions involving matrices have entered our normal lives. Traditional multivariate statistical analysis, such as the vector autoregressive (ARMA) model for multivariate time series, considers that vector sequences exhibit temporal correlation and, more importantly, correlation between their components. Classical t and F statistics are defined for single variables or vectors and cannot meet the needs of real-world applications.
- This difficulty lies in the complexity of the tensor form of the covariance structure of multivariate time series. Specifically, they are often multidimensional, heterogeneous in variance, and non-stationary, posing a challenge to mathematical modeling.

Background II

- One approach is to draw on the theory of random processes, treating time series as second-order random fields, with the vector dimension as the primary indicator and the time dimension as the secondary, to study their properties. The classical Karhunen-Loève theorem and Fourier analysis techniques provide a strong mathematical foundation for this approach.
- Another approach is to consider the distribution of matrices under classical group actions, generally referring to orthogonal groups, unitary groups, symplectic groups, etc. This approach is more direct and does not require the introduction of redundant hyperparameters. However, direct modeling of panel data requires more complex mathematical theory and more sophisticated techniques.

Background III

- Recent textbooks on multivariate statistical analysis, such as Gupta and Nagar (2018); Mathai et al. (2022), including more classic ones Anderson (1958); Srivastava and Khatri (1979); Muirhead (1982), have mostly discussed statistical distribution theory based on the Wishart distribution. However, relatively little literature exists on the general case of non-independent and identically distributed data.
- This paper will adopt the second approach, utilizing polynomials that are invariant under various matrix transformations to derive the exact distribution of matrices with dependent entries and their derived statistics. The construction of this distribution theory relies on simultaneous diagonalization in rectangular coordinates, or singular value decomposition, a core concept throughout this paper.

Background IV

- On the other hand, direct integration over classical groups is often difficult. We need to select a set of orthonormal bases in an appropriate polynomial space to represent the solutions of the Laplace-Beltrami operator. A typical solution is zonal polynomials, namely spherical harmonics in mathematical physics. This basis plays a role similar to the angular part of the radial-angular decomposition in polar coordinates, making it easy to perform integral operations.
- References include classical group representation theory in Weyl (1946); Hua (1958); Macdonald (1998), classical multivariate statistical analysis in James (1960, 1961a,b, 1964); Constantine (1963, 1966), and zonal polynomials and hypergeometric functions in Takemura (1984); Gross and Richards (1987); Shimizu (2022); Richards (2024). Reading these materials will be helpful in understanding the main tools and ideas of this article.

Karhunen-Loève theorem I

In the following, we study the zero-mean two-parameter random process $Z(t,x)$, usually abbreviated as time-space random field or time-frequency random field satisfying

$$\mathrm{E}\iint_{T\times S}|Z(t,x)|^2 dxdt < \infty$$

on Euclidean compact domains T and S, which is equivalent to its covariance

$$C(t_1,t_2,x_1,x_2) = \mathrm{E}Z(t_1,x_1)\overline{Z(t_2,x_2)}$$

is integrable on $(T\times S)\times(T\times S)$ by the Cauchy-Schwartz inequality. We abbreviate $Z(t,x)$ as

Karhunen-Loève theorem II

T_1 (with respect to common spatial basis). if there exists in $L^2(S)$ an orthonormal basis $\{\psi_i(x)\}$ such that $Z(t,x)$ can be developed for each $t \in T$, in the sense of convergence in quadratic mean, i.e. convergence in the norm of the Hilbert space $L^2(\Omega)$, into series

$$Z(t,x) = \sum_i \xi_i(t)\psi_i(x)$$

where $\xi_i(t) = \int Z(t,x)\overline{\psi_i(x)}dt$ and $\mathrm{E}\xi_i(t)\overline{\xi_{i'}(t)} = 0$ whenever $i \neq i'$. In a word, the coefficients of random fields $\xi_i(t)$ are pairwise uncorrelated *at each* t.

$T_{1\frac{1}{2}}$ (with uncorrelated coefficients on temporal domain). if it is T_1 separated and additionally, the coefficients $\xi_i(t)$ are *totally uncorrelated with each other*: $\mathrm{E}\xi_i(t_1)\overline{\xi_{i'}(t_2)} = 0$ whenever $i \neq i', t_1, t_2 \in T$.

Karhunen-Loève theorem III

T_2 (with respect to temporal basis and spatial basis respectively). if it is T_1 separated and there exists further in $L^2(T)$ an orthonormal basis $\{\phi_j(t)\}$ such that $Z(t,x)$ can be developed, in the sense of convergence in quadratic mean, into series

$$Z(t,x) = \sum_i \sum_j \eta_{ij} \phi_j(t) \psi_i(x),$$

where $\eta_{ij} = \int \int Z(t,x) \overline{\phi_j(t)\psi_i(x)} dsdt$ and the random variables η_{ij} are *pairwise uncorrelated with each other*: $\mathrm{E}\eta_{ij}\overline{\eta_{i'j'}} = 0$ whenever $i \neq i'$ or $j \neq j'$.

T_3 (with separated covariance on time and space domains). if it is T_2 separated and additionally, there exist σ_i^2, τ_j^2 such that the covariance of η_{ij} are *degenerate to rank one*: $\mathrm{E}\eta_{ij}\overline{\eta_{i'j'}} = (\sigma_i^2 \delta_{ii})(\tau_j^2 \delta_{j'j'})$.

Karhunen-Loève theorem IV

(*) (degenerate to time and space random fields) if it is T_3 and there exist random variables λ_i, μ_j such that $\eta_{ij} = \lambda_i \mu_j$, or equivalently,

$$Z(s,t) = \left(\sum_j \mu_j \phi_j(t) \right) \left(\sum_i \lambda_i \psi_i(x) \right)$$

where $\mathrm{E}\lambda_i \overline{\lambda_{i'}} = \sigma_i^2 \delta_{ii'}$ and $\mathrm{E}\mu_j \overline{\mu_{j'}} = \tau_j^2 \delta_{jj'}$.

Denote $c_i(t) = \mathrm{E}|\xi_i(t)|^2$, $c_i(t_1, t_2) = \mathrm{E}\xi_i(t_1)\xi_i(t_2)$ and $c_{ij}^2 = \mathrm{E}|\eta_{ij}|^2$. We tabulate some equivalent conditions for these assumptions to hold explicitly

Karhunen-Loève theorem V

Table 1: Comparison of $T_1, T_{1\frac{1}{2}}, T_2$ and T_3.

T_1	$\Leftrightarrow$	$\mathrm{E}Z(t,x_1)\overline{Z(t,x_2)} = \sum_i c_i(t)\psi_i(x_1)\overline{\psi_i(x_2)}$
	$\Leftrightarrow$	$\psi_i(x)$ is the proper functions of $C(t,t,x_1,x_2)$, independent of t
$T_{1\frac{1}{2}}$	$\Leftrightarrow$	$\mathrm{E}Z(t_1,x_1)\overline{Z(t_2,x_2)} = \sum_i c_i(s_1,s_2)\psi_i(t_1)\overline{\psi_i(t_2)}$
	$\Leftrightarrow$	$\psi_i(t)$ is the proper functions of $C(t_1,t_2,x_2,x_2)$, independent of t_1, t_2
T_2	$\Leftrightarrow$	$\mathrm{E}Z(t_1,x_1)\overline{Z(t_2,x_2)} = \sum_i \sum_j c_{ij}\phi_j(t_1)\overline{\phi_j(t_2)}\psi_i(x_1)\overline{\psi_i(x_2)}$
	$\Leftrightarrow$	$c_{ij}, \phi_j(t)\psi_i(x)$ is the proper pairs of $C(t_1,t_2,x_1,x_2)$
T_3	$\Leftrightarrow$	$\mathrm{E}Z(t_1,x_1)\overline{Z(t_2,x_2)} = C_2(t_1,t_2)C_1(s_1,s_2)$
	$\Leftrightarrow$	$\tau_j^2, \phi_j(t)$ and $\sigma_i^2, \psi_i(x)$ are the proper pairs of $C_2(t_1,t_2)$ and $C_1(x_1,x_2)$ respectively
$(*)$	$\Leftrightarrow$	There exist $X(x)$ and $Y(t)$ such that $Z(t,x) = Y(t)X(x)$ and $\mathrm{E}Y(t_1)Y(t_2) = C_2(t_1,t_2)$ and $\mathrm{E}X(x_1)X(x_2) = C_1(x_1,x_2)$

Karhunen-Loève theorem VI

Theorem

From the above table, there are some obvious deductions.

1. *For any* $t \in T$, *if* $\mathrm{E} \int |Z(t,x)|^2 dx < \infty$, *then* $T_{1\frac{1}{2}} \Rightarrow T_1$.
2. *If* $\mathrm{E} \iint |Z(t,x)|^2 dt dx < \infty$, *then* $T_3 \Rightarrow T_2 \Rightarrow T_{1\frac{1}{2}}$.
3. *If* $T_{1\frac{1}{2}}$ *holds true and there exist* $\phi_j(t)$ *forming a proper function of* $c_i(t_1, t_2)$ *that is independent of* i, *then* $T_{1\frac{1}{2}}$ *implies* T_2.
4. *If* $T_{1\frac{1}{2}}$ *holds and there exist constants* ω_i *such that* $c_i(t_1, t_2) = \omega_i \rho(t_1, t_2)$, *for some symmetric function* ρ, *then If* T_3 *holds.*

Karhunen-Loève theorem VII

Proof.

1 is a direct consequences of the Karhunen-Loève theorem. To prove 2, note that the series converges in the quadratic mean, so we can change the order of summation and integral, which means

$$\mathrm{E}\xi_i(t_1)\overline{\xi_{i'}(t_2)} = \mathrm{E}\int Z(t_1,x)\overline{\psi_i(x)}\int Z(t_2,x)\overline{\psi_{i'}(x)} = c_i(t_1,t_2)\delta_{ii'}.$$

This equals to 0 for any $i \neq i'$ and $t_1, t_2 \in T$. If there exist $\phi_j(t)$ forming a proper function of $c_i(t_1,t_2)$ that is independent of i, then there exist c_{ij} such that

$$c_i(t_1,t_2) = \sum c_{ij}\phi_j(t_1)\phi_j(t_2).$$

This proves 3. The proof for 4 is similar. □

Karhunen-Loève theorem VIII

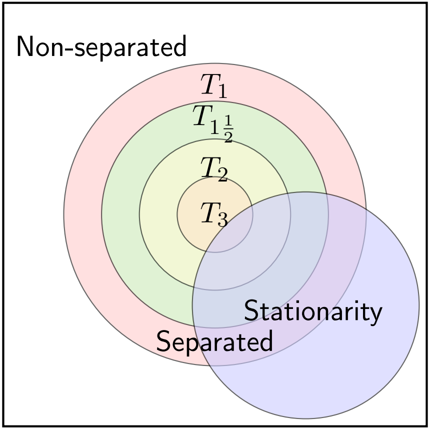

Figure 1: Venn diagram of $T_1, T_{1\frac{1}{2}}, T_2$, and T_3 separation assumptions, within the largest circle the separated and outside the non-separated covariance structures. The stationarity circle overlaps each of them.

Rectangular coordinates I

In this article, we use $A = (a_1, a_2, \ldots, a_p)$ to represent a matrix A arranged in columns, and $A = [\mathrm{a}_1; \mathrm{a}_2; \ldots; \mathrm{a}_n]$ to represent a matrix A arranged in rows. Both vectors and matrices are written in lowercase. The symbol $A \otimes B$ represents the Kronecker product of the matrix $A = (a_{ij})$ and the matrix B, i.e., the block matrix $(a_{ij}B)$. The symbol $\mathrm{vec}(A), A = (a_1, a_2, \ldots, a_n)$ represents the vectorized $[a_1; a_2; \ldots; a_n]$ of the matrix A. Let $\mathbb{R}, \mathbb{C}, \mathbb{O}$ denote the field of real numbers, the field of complex numbers, and the quaternion division ring, respectively. Consider the $np \times np$ block matrix $\mathit{\Theta}$ on $\mathbb{F} = \mathbb{R}, \mathbb{C}, \mathbb{O}$ as follows

$$\mathit{\Theta} = \begin{bmatrix} \mathit{\Theta}_{11} & \mathit{\Theta}_{12} & \ldots & \mathit{\Theta}_{1n} \\ \mathit{\Theta}_{21} & \mathit{\Theta}_{22} & \ldots & \mathit{\Theta}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathit{\Theta}_{n1} & \mathit{\Theta}_{n2} & \ldots & \mathit{\Theta}_{nn} \end{bmatrix},$$

where $\mathit{\Theta}_{ii}$ $(i = 1, 2, \ldots, n)$ is a positive definite matrix of dimension p.

Rectangular coordinates II

Rectangular coordinates refer to a $p \times p$ matrix $B = (b_1, b_2, \ldots, b_p)$, satisfying $\bar{b}_i' b_j = \delta_{ij}$, such that Θ has the following structure

T_1 The following expansion holds

$$\Theta = \sum_{j=1}^{n} \sum_{j'=1}^{n} A_{jj'} \otimes B_{jj'},$$

Where $B_{jj'} = b_i \bar{b}_{j'}'$ and $A_{jj'}(k,l) = \bar{b}_j' \Theta_{kl} b_{j'}$ satisfy

$$A_{jj'}(k,k) = 0 \, (j \neq j').$$

In other words, the diagonal elements of $A_{jj'}$ $(j \neq j')$ are zero.

$T_{1\frac{1}{2}}$ The expansion T_1 holds, and

$$A_{jj'} = A_{jj} \delta_{jj'},$$

That is, $A_{jj'}$ $(j \neq j')$ is zero.

Rectangular coordinates III

T_2 The expansion T_1 holds, There exists another set of $n \times n$ rectangular coordinates $A = (a_1, a_2, \ldots, a_n)$, satisfying $\bar{a}_i' a_j = \delta_{ij}$, such that

$$\Theta = \sum_{i=1}^{n} \sum_{j=1}^{p} \gamma_{ij} A_i \otimes B_j,$$

where $A_i = a_i \bar{a}_i'$, $B_j = b_j \bar{b_j}'$.

T_3 The expansion T_2 holds, and there exist positive constants α_i, β_j such that $\gamma_{ij} = \alpha_i \beta_j$, or equivalently, the following holds

$$\Theta = \Phi^{-1} \otimes \Psi^{-1},$$

where $\Phi^{-1} = \sum_{i=1}^{n} \alpha_i a_i \bar{a}_i'$, $\Psi^{-1} = \sum_{j=1}^{p} \beta_j b_j \bar{b}_j'$.

Rectangular coordinates IV

It is not difficult to see from the table 1.

Theorem

Rectangular coordinates are unique.

Another corollary is the discretization case of Schoenberg (1942).

Theorem

Σ *is positive definite if and only if* A_{jj} $(j = 1, 2, \ldots, p)$ *is positive definite.*

Spherical coordinates I

Let n, p be positive integers and $n \geq p$. Without loss of generality, assume that Z is a real matrix with full column rank. Any $n \times p$ matrix Z can be uniquely decomposed into $Z = HR^{\frac{1}{2}}$, where $H\,(n \times p)$ satisfies $H'H = I_p$ and $R\,(p \times p)$ is a symmetric positive definite matrix.

Definition (Spherical Coordinates)

The unique $p \times p$ matrix H is called the spherical coordinates of Z. In particular, $H = Z(Z'Z)^{-\frac{1}{2}}$, $R = Z'Z$.

We call the set of $n \times p$ matrices that satisfy the condition $H'H = I_p$ a Stiefel manifold, denoted by $V_{n,p}$. For a set of spherical coordinates H $(n \times p)$, we can always expand it into an orthogonal matrix $K = (H, H_\perp) \in O(n)$, where the column vectors of $H_\perp \in V_{n,n-p}$ are

Spherical coordinates II

orthogonal to the column vectors of $H \in V_{n,p}$. James (1954) (Equation (8.19)) was the first to prove that

$$\wedge_{i,j}^{n} dz_{ij} = 2^{-p} |Z'Z|^{\frac{1}{2}(n-p-1)} \wedge_{i \leq j}^{n} d(z_i' z_j) \wedge_{i<j}^{n} h_i' dh_j \wedge_{i=k+1}^{n} \wedge_{j=1}^{k} h_i' dh_j,$$

where $H = (h_1, \ldots, h_k), H_\perp = (h_{k+1}, \ldots, h_n)$. In matrix form,

Theorem (Spherical Decomposition)

$(dZ) = 2^{-p} |R|^{\frac{1}{2}(n-p-1)} \cdot (dR) \cdot (dK)$, *where* $(dK) = (H'dH) \cdot (H_\perp' dH)$.

The geometric meaning of (dK) is, $(H'dH)$ describes the rotation of the column vectors of the object H within the p-dimensional subspace it spans, thus corresponding to the volume element of $O(p)$, and $(H_\perp' dH)$ describes the change of direction of the p-dimensional subspace in the n-dimensional space, that is, the movement of the subspace in the Grassmann manifold $G_{n,p}$. The two together describe the Stiefel manifold $V_{n,p}$. The orthogonal group $O(p)$ is the principal bundle of the fiber on the Grassmann $G_{n,p}$.

Classical distributions I

Consider a population Z of $n \times p$ matrices whose elements are random variables, with a covariance matrix $\Sigma = \mathrm{Cov}(\mathrm{vec}(X'))$.

Definition (Matrix Normal Distribution)

A matrix population X of $n \times p$ is called a normal matrix if $\mathrm{vec}(X')$ is an np-dimensional normal vector.

Hereafter, we call the $n \times p$ matrix X a central matrix normal distribution, and the non-central distribution with mean M is defined if $X - M \sim T_i$, $i = 1, 1\frac{1}{2}, 2, 3$.

Definition (Product Moment Distribution)

$X'X$ is called a product moment distribution, where $X \sim T_i$, $i = 1, 1\frac{1}{2}, 2, 3$ is a matrix normal distribution. In particular, when $X \sim N_{n,p}(M, I_n, \Psi)$, $X'X$ is a Wishart distribution.

Classical distributions II

Definition (Matrix t Distribution)

The $m \times p$ matrix population Z is called a matrix t distribution if it can be written as $Z = XS^{-\frac{1}{2}}$, where $X \in N_{m,p}(M, I_m, I_p)$ is a matrix normal distribution, and $S = Y'Y$, $Y \in T_i$ a central matrix normal distribution.

Definition (Matrix F Distribution)

The $n \times p$ matrix population Z is called a matrix F distribution if it can be written as $Z = S_1 S_2^{-1}$, where $S_1 = X'X', X \in T_i$ is a matrix normal distribution, and $S_2 = Y'Y$, $Y \in T_{i'}$ a central matrix normal distribution. If X is non-central and Y shares a common decision matrix $\Theta \in T_i$ with X, this matrix F distribution is called non-central matrix F distribution.

Zonal polynomials I

Let U be a $n \times n$ symmetric matrix. Let $(a)^k = a(a+1)\dots(a+k-1)$ represent the Pochhammer symbol, and $(a)^\kappa = \prod_{i=1}^m (a - \frac{1}{2}(i-1))^{k_i}$, where $\kappa \vdash k$ is a partition. Define

$$
\begin{aligned}
C_\kappa(I_n) &= \frac{2^k (\frac{1}{2}n)^\kappa}{(2k-1)!!} \chi_{2\kappa}(1), \\
C_\kappa(U) = C_\kappa(I_n)|U|^{k_n} &\int_{O(n)} \prod_{i=1}^{n-1} |HXH'(1,\dots,i)|^{k_i - k_{i+1}} [dH].
\end{aligned} \tag{1}
$$

Here, $(1,\dots,i)$ represents the principal submatrix, $\chi_{2k_1,\dots,2k_n}(1)$ is the character for identity matrix by Schur-Weyl duality.

Theorem (Linearization formula)

$C_\lambda(U) \cdot C_\mu(U) = \sum_{|\nu|=|\lambda|+|\mu|} g_{\lambda,\mu}^{\nu} C_\nu(U)$.

Hypergeometric functions I

The multivariate gamma function, denoted $\Gamma_p(a)$, is defined as

$$\Gamma_n(a) = \int_{A>0} \operatorname{etr}(-A)|A|^{a-\frac{n+1}{2}}(dA), \tag{2}$$

where the real part of a is denoted $\Re(a) > \frac{1}{2}(n-1)$, and is integrated over all $n \times n$ real symmetric positive definite matrices.
A closely related function is the multivariate beta function, denoted $B_n(a,b)$, is defined by

$$B_n(a,b) = \int_{0<X<I} |X|^{a-\frac{n+1}{2}}|I-X|^{b-\frac{n+1}{2}}(dX), \tag{3}$$

where $\Re(a), \Re(b) > \frac{1}{2}(n-1)$, which integrates over all $n \times n$ real symmetric matrices X such that X, $I - X$ are positive definite.

Hypergeometric functions II

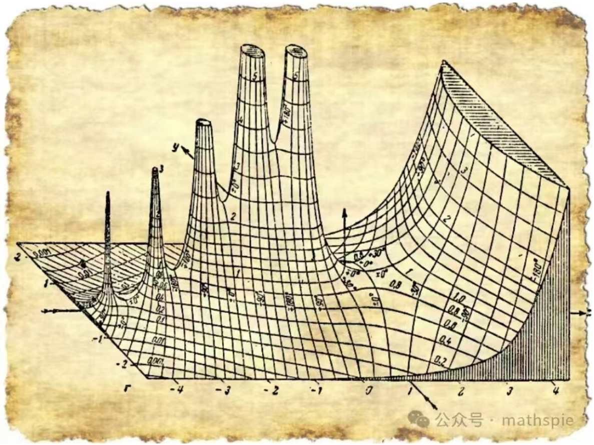

Figure 2: A hand-drawn graph of the absolute value of the complex gamma function, from "Tables of Higher Functions" by Jahnke and Emde, 1909.

Hypergeometric functions III

The Beta function is related to the multivariate gamma function by the following relation

$$B_n(a, b) = \frac{\Gamma_n(a)\Gamma_n(b)}{\Gamma_n(a + b)}. \tag{4}$$

First, we need to show that the hypergeometric function of a single matrix argument can be given by the recurrence relation in Herz (1955). Constantine (1963) solved the system of equations for the hypergeometric function by introducing polynomials. This makes practical computations using hypergeometric functions feasible.

Definition

Let X be an $n \times n$ real symmetric positive definite matrix. Then

$${}_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; X) = \sum_{k=0}^{\infty} \sum_{\kappa \vdash k} \frac{1}{k!} \frac{(a_1)^\kappa \ldots (a_p)^\kappa}{(b_1)^\kappa \ldots (b_q)^\kappa} C_\kappa(X). \tag{5}$$

Hypergeometric functions IV

Some special cases include

$$ {}_0F_0(X) = \mathrm{etr}(X) = \sum_{k=0}^{\infty} \sum_{\kappa \vdash k} \frac{C_\kappa(X)}{k!}, \tag{6} $$

$$ {}_1F_0(a; X) = |I - X|^{-a} = \sum_{k=0}^{\infty} \sum_{\kappa \vdash k} (a)^{\kappa} \frac{C_\kappa(X)}{k!}. \tag{7} $$

In addition, the hypergeometric function of two matrix arguments can also be defined by zonal polynomials.

Hypergeometric functions V

Definition

Let X, Y be $n \times n, m \times m$ real symmetric positive definite matrix ($n \geq m$) respectively. Then

$$
{}_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; X, Y) = \sum_{k=0}^{\infty} \sum_{\kappa \vdash k} \frac{1}{k!} \frac{(a_1)^{\kappa} \ldots (a_p)^{\kappa}}{(b_1)^{\kappa} \ldots (b_q)^{\kappa}} \frac{C_{\kappa}(X) C_{\kappa}(Y)}{C_{\kappa}(I_n)}. \tag{8}
$$

It is worth mentioning that when the two matrix parameters have the same size, the following equation holds

$$
\begin{aligned}
\int_{O(n)} {}_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; XHYH')[dH] \\
= {}_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; X, Y),
\end{aligned} \tag{9}
$$

Hypergeometric functions VI

where X, Y are $n \times n$ symmetric matrices, integrated over all $n \times n$ orthogonal matrices $O(n)$, The normalized Haar measure is defined as $[dH] = \frac{1}{\mathrm{vol}\left(O(n)\right)}(dH)$. If the two matrices are of different sizes, a similar formula is given by the following lemma.

Lemma (Shimizu (2022))

Let A be $n \times n$, B be a $p \times p$ real symmetric matrix. If $n \geq p$,

$$\int_{V_{n,p}} {}_0F_0(AH_1BH_1')[dH_1] = {}_0F_0(A, B),$$

where H_1 runs over all $n \times p$ matrices in the Stiefel manifold $V_{n,p} = \{H_1 : H_1'H_1 = I_p\}$, and $[dH_1] = \frac{1}{\mathrm{vol}(V_{n,p})}(dH_1)$.

The following lemma is used to derive the non-central distribution.

Hypergeometric functions VII

Lemma (James (1961b))

Let X be a $n \times p$ real matrix.

$$\int_{V_{n,p}} \operatorname{etr}(XH_1')[dH_1] = {}_0F_1\left(\frac{1}{2}n; \frac{1}{4}X'X\right),$$

where H_1 runs over the Stiefel manifold $V_{n,p}$.

Lemma (Khatri (1966))

Let A be a $n \times n$ real symmetric matrix, B be a $p \times p$ real symmetric positive definite matrix, and $n > p - 1$. For any fixed $p \times p$ matrix S, we have

$$\int_{X'X=S} \operatorname{etr}(AXBX')(dX) = \frac{\pi^{\frac{np}{2}}}{\Gamma_p(\frac{n}{2})} |S|^{\frac{n-p-1}{2}} {}_0F_0(A, BS).$$

The following two lemma is used for the derivation of the distribution of largest and smallest roots in MANOVA.

Hypergeometric functions IX

Lemma (Constantine (1963, 1966); Davis (1979, 1981))

The incomplete Gamma and Beta integrals involving hypergeometric functions are

$$\int_0^R \operatorname{etr}(-AS)|S|^{c-\frac{p+1}{2}} {}_pF_q\begin{pmatrix} a_1,\dots,a_p \\ b_1,\dots,b_q \end{pmatrix}; BS\Bigg) \, dS$$

$$= \frac{|R|^c}{\mathrm{B}_p(c,\frac{p+1}{2})} {}_{p+1}F_{q+1}\left(\begin{matrix} c, a_1,\dots,a_p \\ c+\frac{p+1}{2}, b_1,\dots,b_q \end{matrix}; -AR, BR\right)$$

$$\int_0^R |S|^{c-\frac{p+1}{2}}|I-S|^{d-\frac{p+1}{2}} {}_pF_q\left(\begin{matrix} a_1,\dots,a_p \\ b_1,\dots,b_q \end{matrix}; AS\right) dS$$

$$= \frac{|R|^c}{\mathrm{B}_p\left(c,\frac{p+1}{2}\right)} {}_{p+2}F_{q+1}\left(\begin{matrix} c, -d+\frac{p+1}{2}, a_1,\dots,a_p \\ c+\frac{p+1}{2}, b_1,\dots,b_q \end{matrix}; A, AR\right)$$

Product Moment Distribution

Product moment distribution I

Consider an $n \times p$ matrix population X with general normal entries, which has a probability density function $p(X)$ with respect to the Lebesgue measure on $\mathbb{R}^{n \times p}$.

Theorem

The product moment distribution $X'X$, where $X \sim T_i, i = 1, 1\frac{1}{2}, 2, 3$, is positive definite with probability one if and only if $n > p - 1$.

This is an observation due to Dykstra (1970) that the p-dimensional normal distribution lies in the $(p-1)$-dimensional subspace with probability zero.

Product moment distribution II

Theorem

Let the rectangular coordinates of $X \in T_1$ *be* $B = (b_1, b_2, \ldots, b_p)$*. Assume that the* $p \times p$ *real symmetric matrix* $U = (u_{ij})$ *satisfies* $u_{ij} = \mathrm{tr}(A_{ij})$ *and all elements on its diagonal are positive, that is,* $u_{ii} > 0 \, (i = 1, 2, \ldots, p)$*. Then when* $n > p - 1$*, the probability density function of* $S = X'X$ *depends only on* $T = (t_{ij})$*,* $t_{ij} = \mathrm{tr}(B_{ij}S)$*,*

$$\begin{aligned} &\frac{|\Theta_1|^{\frac{1}{2}}}{2^{\frac{np}{2}}\Gamma_p(\frac{n}{2})} \mathrm{etr}\left(-\frac{1}{2}UT\right)|T|^{\frac{n-p-1}{2}}, \text{ when } \; M = 0; \\ &\times \mathrm{etr}\left(-\frac{1}{2}\Omega\right) {}_0F_1\left(\frac{n}{2}; \frac{1}{4}\Delta T\right), \text{ when } \; M \neq 0; \end{aligned} \tag{10}$$

Where $\Omega = \sum_{i,j=1}^{p} B'_{ij}M'A'_{ij}M$ *and* $\Delta = \sum_{i,j,k,l=1}^{p} B'_{ij}M'A'_{ij}A_{kl}MB_{kl}$*.*

Product moment distribution III

Theorem

The moment generating function of $S = (s_{ij})$ is

$$\begin{aligned} \mathrm{E}\exp\left(\sum_{i\le j}\gamma_{ij}s_{ij}\right) &= |\Theta_1|^{\frac{1}{2}}|U|^{-\frac{n}{2}}{}_1F_0\left(\frac{n}{2};W\right), \text{ When } \ M=0; \\ &\times \operatorname{etr}\left(-\frac{1}{2}\Omega\right)\operatorname{etr}\left(\frac{1}{2}\Delta U^{-1}(I-W)^{-1}\right), \text{ when } \ M\neq 0; \end{aligned} \tag{11}$$

Where $W = U^{-\frac{1}{2}}BRB'U^{-\frac{1}{2}}$, $2R = \Gamma + I$, and $\Gamma = (\gamma_{ij})$ is symmetric.

Product moment distribution IV

Theorem

The joint distribution of the eigenvalues $l_1, l_2, \ldots, l_p$ *of* S *is*

$$\frac{\pi^{\frac{p^2}{2}} |\Theta_1|^{\frac{1}{2}}}{2^{\frac{np}{2}} \Gamma_p(\frac{n}{2}) \Gamma_p(\frac{p}{2})} \prod_{i<j}^{p} (l_i - l_j) \prod_{i=1}^{p} (l_i)^{\frac{n-p-1}{2}} {}_0F_0\left(-\frac{1}{2}U, L\right), \text{ when } M = 0;$$

$$\times \operatorname{etr}\left(-\frac{1}{2}\Omega\right) {}_0F_1\left(\frac{n}{2}; \frac{1}{4}\Delta, L\right), \text{ when } M \neq 0;$$

where $L = \operatorname{diag}(l_i)$, $l_1 > l_2 > \cdots > l_p > 0$; *and zero elsewhere.*

Although the above three theorems only consider real distributions, the results for complex cases are similar. Furthermore, the results for the cases $T_{1\frac{1}{2}}$, T_2, and T_3 should also be parallel. Rectangular coordinates are closely related to elliptical countered distritbutions.

Product moment distribution V

Table 2: Characterisations of left-spherical (LS), multivariate spherical (MS), vector spherical (VS) distributions and their extensions to elliptical distributions.

Type	$-2\log$(M.G.F.)	DOF	Class	M.G.F.
T_1	$\sum_{j=1}^{p}\sum_{j'=1}^{p} t_j' A_{jj'} t_{j'}$	$\frac{1}{2}n(n-1)p^2 + np$	LE	$\phi(t_j' A_{jj'} t_{j'})$
$T_{1\frac{1}{2}}$	$\sum_{j=1}^{p} t_j' A_{jj} t_j$	$\frac{1}{2}n(n+1)p$	ME	$\phi(t_j' A_{jj} t_j)$
T_2	$\sum_{i=1}^{n}\sum_{j=1}^{p} \gamma_{ij} t_j' A_i t_j$	np	ME	$\phi(t_j' A_{jj} t_j)$
T_3	$\sum_{i=1}^{n}\sum_{j=1}^{p} \alpha_i \beta_j t_j' A_i t_j$	$n+p$	VE	$\phi(\sum t_j' A_{jj} t_j)$

where $T = (t_1, \ldots, t_n)$ and $X = (x_1, \ldots, x_n)$ are both $n \times p$ matrices, $A_{ij}' = A_{ji}$, and the M.G.F. is $\mathrm{E}\exp(\sum t_{ri} x_{ri})$. These terminologies LE, ME, VE are slightly modified from Fang and Zhang (1990).

Product moment distribution VI

Question

Can a 3×3 left elliptical distribution be constructed that is not $T_{1\frac{1}{2}}$?

Question

Can the density theorem be proved by mathematical induction?

Question

Can it be proved that (11) *for $M = 0$ defines a moment generating function if and only if $n \in \{0, 1, 2, \dots, p-1\} \cup (p-1, \infty)$ and* (11) *for $M \neq 0$ defines a moment generating function if and only if, in addition, $n \geq \max\{\mathrm{rank}(\mathit{\Omega}), \mathrm{rank}(\mathit{\Delta})\}$ when $n < p-1$?*

Matrix t Distribution

Matrix t distribution I

Lemma

Consider the positive definite matrices $A\ (n\times n), B\ (p\times p)$*, and the fixed matrix* $C\ (n\times p)$*. When* $n\geq p$

$$\int_{M_{n,p}} \operatorname{etr}(-AXBX'+CX')|X'X|^{a-\frac{p+1}{2}}(dX)=\frac{\pi^{\frac{np}{2}}\Gamma_p\left(a+\frac{1}{2}(n-p-1)\right)}{|A|^{\frac{p}{2}}|B|^{a+\frac{n-p-1}{2}}\Gamma_p(\frac{n}{2})}$$
$$\times {}_1F_1\left(a+\frac{1}{2}(n-p-1),\frac{1}{2}n;A,\frac{1}{4}C'A^{-1}CB^{-1}\right)$$

where $\Re(a)>\frac{1}{2}(p-1)$*, and* X *runs over all* $n\times p$ *matrices.*

Matrix t distribution II

Theorem

The matrix t distribution Z has the following density

$$\frac{\pi^{\frac{p(p+1)}{4}}\Gamma_p\left(n+\frac{1}{2}m-\frac{1}{2}\right)}{(2)^{\frac{1}{2}(n+m)p}\Gamma_p(\frac{n}{2})\Gamma_p(\frac{p}{2})}|U+Z'Z|^{-n-\frac{m}{2}-\frac{1}{2}}\operatorname{etr}\left(-\frac{1}{2}M'M\right) \\ \times {}_1F_1\left(n+\frac{1}{2}m-\frac{1}{2},\frac{1}{2}p;\frac{1}{4}M'ZZ'MB(U+Z'Z)^{-1}B^{'}\right). \tag{12}$$

Matrix F Distribution

Matrix F distribution I

Theorem

The central matrix F distribution $S = S_1 S_2^{-1}$ has probability density function

$$p(S) = \frac{|\Theta_1|^{\frac{1}{2}}|\Theta_2|^{\frac{1}{2}}\Gamma_p(\frac{n_1+n_2+p+1}{2})}{2^{(n_1+n_2)p}\mathrm{B}_p(\frac{n_1}{2},\frac{n_2}{2})\Gamma_p(\frac{p+1}{2})|U_2|^{\frac{n_1+n_2}{2}}}|S|^{\frac{n_1-p-1}{2}}$$
$$\times {}_1F_1\left(\frac{n_1+n_2}{2};\frac{p+1}{2};-\frac{1}{2}(U_1U_2^{-1}+S)\right),$$

Matrix F distribution II

Theorem

The joint distribution of latent roots $f_1, f_2, \ldots, f_p$ of non-central F distribution $S = S_1(S_1 + S_2)^{-1}$ is

$$\frac{\pi^{p^2/2}|\Theta|}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) \Gamma_p(\frac{p}{2}) |U|^{\frac{n_1+n_2}{2}}} \prod_{i<j}^{p} (f_i - f_j) |F|^{\frac{n_1}{2}} |I - F|^{\frac{n_1-p-1}{2}}$$

$$\times \operatorname{etr}\left(-\frac{1}{2}\Omega\right) {}_1F_1\left(\frac{n_1+n_2}{2}; \frac{n_1}{2}; \Delta U^{-1}, F\right),$$

where $F = \operatorname{diag}(f_1, f_2, \ldots, f_p)$, $f_1 > f_2 > \cdots > f_p$; elsewhere zero.

Matrix F distribution III

Theorem

The probability distribution function for the largest root f_1 is

$$P(f_1 < x) = \frac{|\Theta| x^{\frac{n_1 p}{2}}}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) B_p(\frac{n_1}{2}, \frac{p+1}{2}) |U|^{\frac{n_1+n_2}{2}}} \times \operatorname{etr}\left(-\frac{1}{2}\Omega\right) {}_2F_1\left(a, b; c; \Delta U^{-1}, \Delta U^{-1} R\right).$$

where $a = -\frac{1}{2}(n_2 - p - 1), b = \frac{1}{2}(n_1 + n_2)$, $c = \frac{1}{2}(n_2 + p + 1)$. *Similarly,*

$$1 - P(f_p \le y) = P(f_p > y) = \frac{|\Theta|(1-y)^{\frac{n_2 p}{2}}}{2^{(n_1+n_2)p} B_p(\frac{n_1}{2}, \frac{n_2}{2}) B_p(\frac{n_2}{2}, \frac{p+1}{2}) |U|^{\frac{n_1+n_2}{2}}} \times \operatorname{etr}\left(-\frac{1}{2}\Omega\right) {}_2F_1\left(a, b; c; \Delta U^{-1}, \Delta U^{-1} R\right).$$

where $a = -\frac{1}{2}(n_1 - p - 1), b = \frac{1}{2}(n_1 + n_2)$, $c = \frac{1}{2}(n_1 + p + 1)$.

References

[1] Wang, Haoming. On the distribution of the sample covariance from a matrix normal population. Preprint.
[2] Wang, Haoming. On the distribution of the ratio of a matrix normal distribution and the sample covariance. Preprint.
[3] Wang, Haoming. On incomplete Gamma and Beta integrals. Preprint.
[4] Wang, Haoming. Matrix Distributions under Classical Group Actions. SYSU Thesis.

T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*, volume 2. Wiley New York, 1958.

A. G. Constantine. Some non-central distribution problems in multivariate analysis. *Ann. Math. Stat.*, 34(4):1270–1285, 1963.

A. G. Constantine. The distribution of Hotelling's generalized ${T_0}^2$. *Ann. Math. Statist.*, 37:215–225, 1966.

A. W. Davis. Invariant polynomials with two matrix arguments extending the zonal polynomials: Applications to multivariate distribution theory. *Ann. Inst. Stat. Math.*, 31:465–485, 1979.

A. W. Davis. On the construction of a class of invariant polynomials in several matrices, extending the zonal polynomials. *Ann. Inst. Stat. Math.*, 33:297–313, 1981.

R. L. Dykstra. Establishing the positive definiteness of the sample covariance matrix. *Ann. Math. Stat.*, 41(6):2153–2154, 1970.

K.-T. Fang and Y.-T. Zhang. Generalized multivariate analysis. *Science Press*, 1990.

K. I. Gross and S. P. Richards, D. Special functions of matrix argument. I: Algebraic Induction, Zonal Polynomials, and Hypergeometric Functions. *Trans. Amer. Math. Soc.*, 301(2):781–811, 1987.

A. K. Gupta and D. K. Nagar. *Matrix variate distributions*. Chapman and Hall/CRC, 2018.

C. S. Herz. Bessel functions of matrix argument. *Ann. Math.*, 61(3): 474–523, 1955.

L.-K. Hua. *Harmonic analysis of functions of several complex variables in the classical domains*. 6. American Mathematical Society, 1958.

A. T. James. Normal multivariate analysis and the orthogonal group. *Ann. Math. Stat.*, 25(1):40–75, 1954.

A. T. James. The distribution of the latent roots of the covariance matrix. *Ann. Math. Stat.*, 31:151–158, 1960.

A. T. James. The distribution of noncentral means with known covariance. *Ann. Math. Stat.*, 32:874–882, 1961a.

A. T. James. Zonal polynomials of the real positive definite symmetric matrices. *Ann. Math.*, 74(2):456–469, 1961b.

A. T. James. Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Stat.*, 35:475–501, 1964.

C. G. Khatri. On certain distribution problems based on positive definite quadratic functions in normal vectors. *Ann. Math. Stat.*, 37(2):468 – 479, 1966.

I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford university press, 1998.

A. M. Mathai, S. B. Provost, and H. J. Haubold. *Multivariate statistical analysis in the real and complex domains*. Springer Nature, 2022.

R. J. Muirhead. *Aspects of Multivariate Statistical Theory*. John Wiley & Sons, 1982.

D. S. P. Richards. Functions of matrix argument. In *NIST Digital Library of Mathematical Functions*, chapter 35. 2024. Available online at `http://dlmf.nist.gov/35.4.E3`.

I. J. Schoenberg. Positive definite functions on spheres. *Duke Math. J.*, 9 (1):96 – 108, 1942.

K. Shimizu. *Distribution theory of eigenvalues for a singular beta-Wishart matrix*. PhD thesis, Tokyo Rika University, 2022.

M. S. Srivastava and C. G. Khatri. *An introduction to multivariate statistics*. Elsevier Science Ltd, 1979.

A. Takemura. *Zonal polynomials*, volume 4. Inst. Math. Stat., 1984.

H. Weyl. *The classical groups: their invariants and representations*, volume 45. Princeton university press, 1946.