CHAPTER 7

JACOBIANS OF MATRIX TRANSFORMATIONS

[*This Chapter is based on the lectures of Professor A.M. Mathai of McGill University, Canada (Director of the SERC Schools).*]

7.0. Introduction

Real scalar functions of matrix argument, when the matrices are real, will be dealt with. It is difficult to develop a theory of functions of matrix argument for general matrices. Let $X = (x_{ij}), i = 1 \cdots, m$ and $j = 1, \cdots, n$ be an $m \times n$ matrix where the x_{ij}'s are real elements. It is assumed that the readers have the basic knowledge of matrices and determinants. The following standard notations will be used here. A prime denotes the transpose, X' = transpose of X, $|(.)|$ denotes the determinant of the square matrix, $m \times m$ matrix $(\cdot)$. The same notation will be used for the absolute value also. $\mathrm{tr}(X)$ denotes the trace of a square matrix X, $\mathrm{tr}(X)$ = sum of the eigenvalues of X = sum of the leading diagonal elements in X. A real symmetric positive definite X (definiteness is defined only for symmetric matrices when real) will be denoted by $X = X' > 0$. Then $0 < X = X' < I \Rightarrow X = X' > 0$ and $I - X > 0$. Further, $\mathrm{d}X$ will denote the wedge product or skew symmetric product of the differentials $\mathrm{d}x_{ij}$'s.

That is, when $X = (x_{ij})$, an $m \times n$ matrix

$$\mathrm{d}X = \mathrm{d}x_{11} \wedge \cdots \wedge \mathrm{d}x_{1n} \wedge \mathrm{d}x_{21} \wedge \cdots \mathrm{d}x_{2n} \wedge \cdots \wedge \mathrm{d}x_{mn}. \tag{7.0.1}$$

If $X = X'$, that is symmetric, and $p \times p$, then

$$\mathrm{d}X = \mathrm{d}x_{11} \wedge \mathrm{d}x_{21} \wedge \mathrm{d}x_{22} \wedge \mathrm{d}x_{31} \wedge \cdots \wedge \mathrm{d}x_{pp} \tag{7.0.2}$$

a wedge product of $1 + 2 + \cdots + p = p(p+1)/2$ differentials.

A wedge product or skew symmetric product is defined in Chapter 1.

7.1. Jacobians of Linear Matrix Transformations

Some standard Jacobians, that we will need later, will be illustrated here. For more on Jacobians see Mathai (1997). First we consider a very basic linear transformation involving a vector of real variables going to a vector of real variables.

Theorem 7.1.1. *Let X and Y be $p \times 1$ vectors of real scalar variables, functionally independent (no element in X is a function of the other elements in X and similarly no element in Y is a function of the other elements in Y), and let $Y = AX$, $|A| \neq 0$, where $A = (a_{ij})$ is a nonsingular $p \times p$ matrix of constants (A is free of the elements in X and Y; each element in Y is a linear function of the elements in X, and vice versa). Then*

$$Y = AX,\ |A| \neq 0 \Rightarrow \mathrm{d}Y = |A|\mathrm{d}X. \tag{7.1.1}$$

Proof 7.1.1.

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} = AX = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \Rightarrow$$

$$y_i = a_{i1}x_1 + ... + a_{ip}x_p, i = 1, ..., p.$$

$$\frac{\partial y_i}{\partial x_j} = a_{ij} \Rightarrow \left(\frac{\partial y_i}{\partial x_j}\right) = (a_{ij}) = A \Rightarrow J = |A|.$$

Hence,

$$\mathrm{d}Y = |A|\mathrm{d}X.$$

That is, Y and X, $p \times 1$, A is $p \times p$, $|A| \neq 0$, A is a constant matrix, then

$$Y = AX,\ |A| \neq 0 \Rightarrow \mathrm{d}Y = |A|\mathrm{d}X.$$

Example 7.1.1. Consider the transformation $Y = AX$ where $Y' = (y_1, y_2, y_3)$ and $X' = (x_1, x_2, x_3)$ and let the transformation be

$$y_1 = x_1 + x_2 + x_3$$
$$y_2 = 3x_2 + x_3$$
$$y_3 = 5x_3.$$

Then write $\mathrm{d}Y$ in terms of $\mathrm{d}X$.

Solution 7.1.1. From the above equations, by taking the differentials we have

$$\begin{aligned}\mathrm{d}y_1 &= \mathrm{d}x_1 + \mathrm{d}x_2 + \mathrm{d}x_3\\ \mathrm{d}y_2 &= 3\mathrm{d}x_2 + \mathrm{d}x_3\\ \mathrm{d}y_3 &= 5\mathrm{d}x_3.\end{aligned}$$

Then taking the product of the differentials we have

$$\begin{aligned}\mathrm{d}y_1 \wedge \mathrm{d}y_2 \wedge \mathrm{d}y_3 = [&\mathrm{d}x_1 + \mathrm{d}x_2 + \mathrm{d}x_3]\\ &\times \wedge[3\mathrm{d}x_2 + \mathrm{d}x_3] \wedge [5\mathrm{d}x_3].\end{aligned}$$

Taking the product directly and then using the fact that $\mathrm{d}x_2 \wedge \mathrm{d}x_2 = 0$ and $\mathrm{d}x_3 \wedge \mathrm{d}x_3 = 0$ we have

$$\begin{aligned}\mathrm{d}Y &= \mathrm{d}y_1 \wedge \mathrm{d}y_2 \wedge \mathrm{d}y_3\\ &= 15\mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3 = 15\mathrm{d}X\\ &= |A|\mathrm{d}X\end{aligned}$$

where

$$A = \begin{bmatrix}1 & 1 & 1\\ 0 & 3 & 1\\ 0 & 0 & 5\end{bmatrix}.$$

This verifies Theorem 7.1.1 also. This theorem is the standard result that is seen in elementary textbooks. Now we will investigate more elaborate linear transformations.

Theorem 7.1.2. *Let X and Y be $m \times n$ matrices of functionally independent real variables and let $A, m \times m$ be a nonsingular constant matrix. Then*

$$Y = AX \Rightarrow \mathrm{d}Y = |A|^n \mathrm{d}X. \tag{7.1.2}$$

Proof 7.1.2. Let $Y = AX = (AX^{(1)}, AX^{(2)}, ..., AX^{(n)})$ where $X^{(1)}, ..., X^{(n)}$ are the columns of X. Then the Jacobian matrix for X going to Y is of the form

$$\begin{bmatrix} A & O & \dots & O \\ O & A & \dots & O \\ \vdots & \vdots & \dots & \vdots \\ O & O & \dots & A \end{bmatrix} \Rightarrow \begin{vmatrix} A & O & \dots & O \\ \vdots & \vdots & \dots & \vdots \\ O & O & \dots & A \end{vmatrix} = |A|^n = J \qquad (7.1.3)$$

where O denotes a null matrix and J is the Jacobian for the transformation of X going to Y or $\mathrm{d}Y = |A|^n \mathrm{d}X$.

In the above linear transformation the matrix X was pre-multiplied by a nonsingular constant matrix A. Now let us consider the transformation of the form $Y = XB$ where X is post-multiplied by a nonsingular constant matrix B.

Theorem 7.1.3. *Let X be a $m \times n$ matrix of functionally independent real variables and let B be an $n \times n$ nonsingular matrix of constants. Then*

$$Y = XB,\ |B| \neq 0 \Rightarrow \mathrm{d}Y = |B|^m \mathrm{d}X, \qquad (7.1.4)$$

Proof 7.1.3.

$$Y = XB = \begin{bmatrix} X^{(1)}B \\ \vdots \\ X^{(m)}B \end{bmatrix}$$

where $X^{(1)}, ..., X^{(m)}$ are the rows of X. The Jacobian matrix is of the form,

$$\begin{bmatrix} B & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & B \end{bmatrix} \Rightarrow \begin{vmatrix} B & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & B \end{vmatrix} = |B|^m \Rightarrow \mathrm{d}Y = |B|^m \mathrm{d}X.$$

Then combining the above two theorems we have the Jacobian for the most general linear transformation.

Theorem 7.1.4. *Let X and Y be $m \times n$ matrices of functionally independent real variables. Let A be $m \times m$ and B be $n \times n$ nonsingular matrices of constants. Then*

$$Y = AXB, |A| \neq 0, |B| \neq 0\ , Y, m \times n, X, m \times n, \Rightarrow \mathrm{d}Y = |A|^n |B|^m \mathrm{d}X. \qquad (7.1.5)$$

Proof 7.1.4. For proving this result first consider the transformation $Z = AX$ and then the transformation $Y = ZB$, and make use of Theorems 7.1.2 and 7.1.3.

In Theorems 7.1.2 to 7.1.4 the matrix X was rectangular. Now we will examine a situation where the matrix X is square and symmetric. If X is $p \times p$ and symmetric then there are only $1 + 2 + \cdots + p = p(p+1)/2$ functionally independent elements in X because, here $x_{ij} = x_{ji}$ for all i and j. Let $Y = Y' = AXA', X = X', |A| \neq 0$. Then we can obtain the following result:

Theorem 7.1.5. *Let $X = X'$ be a $p \times p$ real symmetric matrix of $p(p+1)/2$ functionally independent real elements and let A be a $p \times p$ nonsingular constant matrix. Then*

$$Y = AXA', X = X', |A| \neq 0, \Rightarrow \mathrm{d}Y = |A|^{p+1}\mathrm{d}X. \tag{7.1.6}$$

Proof 7.1.5. This result can be proved by using the fact that a nonsingular matrix such as A can be written as a product of elementary matrices in the form

$$A = E_1 E_2 \cdots E_k$$

where $E_1, \cdots, E_k$ are elementary matrices. Then

$$Y = AXA' \Rightarrow E_1 E_2 \cdots E_k X E_k' \cdots E_1'.$$

where E_j' is the transpose of E_j. Let $Y_k = E_k X E_k'$, $Y_{k-1} = E_{k-1} Y_k E_{k-1}'$, and so on, and finally $Y = Y_1 = E_1 Y_2 E_1'$. Evaluate the Jacobians in these transformations to obtain the result, observing the following facts. If, for example, the elementary matrix E_k is formed by multiplying the i-th row of an identity matrix by the nonzero scalar c then taking the wedgeproduct of differentials we have $\mathrm{d}Y_k = c^{p+1}\mathrm{d}X$. Similarly, for example, if the elementary matrix E_{k-1} is formed by adding the i-th row of an identity matrix to its j-th row then the determinant remains the same as 1 and hence $\mathrm{d}Y_{k-1} = \mathrm{d}Y_k$. Since these are the only two types of basic elementary matrices, systematic evaluation of successive Jacobians gives the final result as $|A|^{p+1}$.

Note 7.1.1. From the above theorems the following properties are evident: If X is a $p \times q$ matrix of functionally independent real variables and if c is a scalar quantity and B is a $p \times q$ constant matrix then

$$Y = c\,X \Rightarrow \mathrm{d}Y = c^{pq}\mathrm{d}X \tag{7.1.7}$$

$$Y = c\,X + B \Rightarrow \mathrm{d}Y = c^{pq}\mathrm{d}X. \tag{7.1.8}$$

Note 7.1.2. If X is a $p \times p$ symmetric matrix of functionally independent real variables, a is a scalar quantity and B is a $p \times p$ symmetric constant matrix then

$$Y = a\,X + B, X = X',\ B = B' \Rightarrow \mathrm{d}Y = a^{p(p+1)/2}\mathrm{d}X. \tag{7.1.9}$$

Note 7.1.3. For any $p \times p$ lower triangular (or upper triangular) matrix of $p(p+1)/2$ functionally independent real variables, $Y = X + X'$ is a symmetric matrix, where X' denoting the transpose of $X = (x_{ij})$, then observing that the diagonal elements in $Y = (y_{ij})$ are multiplied by 2, that is, $y_{ii} = 2x_{ii},\ i = 1, ..., p$, we have

$$Y = X + X' \Rightarrow \mathrm{d}Y = 2^{p}\mathrm{d}X. \tag{7.1.10}$$

Example 7.1.2. Let X be a $p \times q$ matrix of pq functionally independent random variables having a matrix-variate Gaussian distribution with the density given by

$$f(X) = c\ \exp\{-\mathrm{tr}[A(Z - M)B(X - M)']\}$$

where, A is a $p \times p$ positive defi nite constant matrix, B is a $q \times q$ positive defi nite constant matrix, M is $p \times q$ constant matrix, tr(·) denotes the trace of the matrix (·) and c is the normalizing constant, then evaluate c.

Solution 7.1.2. Since A and B are positive definite matrices we can write $A = A_1A_1'$ and $B = B_1B_1'$ where A_1 and B_1 are nonsingular matrices, that is, $|A_1| \neq 0,\ |B_1| \neq 0$. Also we know that for any two matrices P and Q, $\mathrm{tr}(PQ) = \mathrm{tr}(QP)$ as long as PQ and QP are defined, PQ need not be equal to QP. Then

$$\begin{aligned}\mathrm{tr}[A(X - M)B(X - M)'] &= \mathrm{tr}[A_1A_1'(X - M)B_1B_1'(X - M)'] \\ &= \mathrm{tr}[A_1'(X - M)B_1B_1'(X - m)'A_1] \\ &= \mathrm{tr}(YY')\end{aligned}$$

where

$$Y = A_1'(X - M)B_1,$$

But from Theorem 7.1.4

$$\mathrm{d}Y = |A_1'|^q |B_1|^p \mathrm{d}(X - M) = |A_1|^q |B_1|^p \mathrm{d}(X - M)$$

since $|A_1|' = |A_1|$

$$= |A|^{q/2} |B|^{p/2} \mathrm{d}X$$

since $|A| = |A_1|^2,\ |B| = |B_1|^2, \mathrm{d}(X - M) = \mathrm{d}(X)$, M being a constant matrix. If $f(X)$ is a density then the total integral is unity, that is,

$$\begin{aligned} 1 &= \int_X f(X)\mathrm{d}X \\ &= c \int_X \exp\{-\mathrm{tr}[A(X - M)B(X - M)']\}\mathrm{d}X \\ &= c \int_Y \exp\{-\mathrm{tr}[YY']\}\mathrm{d}Y \end{aligned}$$

where, for example, $\int_X$ denotes the integral over all elements in X. Note that for any real matrix P, trace of PP' is the sum of squares of all the elements in P. Hence

$$\begin{aligned} \int_Y \exp\{-\mathrm{tr}[YY']\}\mathrm{d}Y &= \int_Y \exp\{-\sum_{i,j} y_{ij}^2\}\mathrm{d}Y \\ &= \prod_{i,j} \int_{-\infty}^{\infty} \mathrm{e}^{-y_{ij}^2} \mathrm{d}y_{ij}. \end{aligned}$$

But

$$\int_{-\infty}^{\infty} \mathrm{e}^{-u^2} \mathrm{d}u = \sqrt{\pi}$$

and therefore

$$\begin{aligned} 1 &= c\,|A|^{q/2}|B|^{p/2}\sqrt{\pi}^{pq} \Rightarrow \\ c &= (|A|^{q/2}|B|^{p/2}\pi^{pq/2})^{-1}. \end{aligned}$$

Note 7.1.4. What happens in the transformation $Y = X + X'$ where both X and Y are $p \times p$ matrices of functionally independent real elements. When $X = X'$, then $Y = 2X$ and this case is already covered before. If $X \neq X'$ then Y has become symmetric with $p(p+1)/2$ variables whereas in X there are p^2 variables and hence this is not a one-to-one transformation.

Example 7.1.3. Consider the transformaton

$$y_{11} = x_{11} + x_{21},\ y_{12} = x_{11} + x_{21} + 2x_{12} + 2x_{22},$$
$$y_{13} = x_{11} + x_{21} + 2x_{13} + x_{23},\ y_{21} = x_{11} + 3x_{21},$$
$$y_{22} = x_{11} + 3x_{21} + 2x_{12} + 6x_{22},\ y_{23} = x_{11} + 3x_{21} + 2x_{13} + 6x_{23}.$$

Write this transformatin in the form $Y = AXB$ and then evaluate the Jacobian in this transformation.

Solution 7.1.3. Writing the transformation in the form $Y = AXB$ we have

$$Y = \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \end{bmatrix},\ X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix},$$
$$A = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix},\ B = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Hence the Jacobian is

$$J = |A|^3|B|^2 = (2^3)(4^2) = 128.$$

This can also be verified by taking the differentials in the starting explicit forms and then taking the wedge products. This verification is left to the reader.

Exercises 7.1.

7.1.1. If X and A are $p \times p$ lower triangular matrices where $A = (a_{ij})$ is a constant matrix with $a_{jj} > 0,\ j = 1, ..., p, X = (x_{ij})$ and x_{ij}'s, $i \geq j$ are functionally independent real variables then show that

$$Y = XA \Rightarrow \mathrm{d}Y = \left\{\prod_{j=1}^{p} a_{jj}^{p-j+1}\right\}\mathrm{d}X,$$
$$Y = AX \Rightarrow \mathrm{d}Y = \left\{\prod_{j=1}^{p} a_{jj}^{j}\right\}\mathrm{d}X,$$

and

$$Y = aX \Rightarrow \mathrm{d}Y = a^{p(p+1)/2}\mathrm{d}X \qquad (7.1.11)$$

where a is a scalar quantity.

7.1.2. Let X and B be uppper triangular $p \times p$ matrices where $B = (b_{ij})$ is a constant matrix with $b_{jj} > 0, j = 1, ..., p$, $X = (x_{ij})$ where the x_{ij}'s, $i \le j$ be functionally independent real variables and b be a scalar quantity, then show that

$$Y = XB \Rightarrow \mathrm{d}Y = \left\{\prod_{j=1}^{p} b_{jj}^{j}\right\}\mathrm{d}X,$$

$$Y = BX \Rightarrow \mathrm{d}Y = \left\{\prod_{j=1}^{p} b_{jj}^{p+1-j}\right\}\mathrm{d}X,$$

and

$$Y = bX \Rightarrow \mathrm{d}Y = b^{p(p+1)/2}\mathrm{d}X. \tag{7.1.12}$$

7.1.3. Let X, A, B be $p \times p$ lower triangular matrices where $A = (a_{ij})$ and $B = (b_{ij})$ be constant matrices with $a_{jj} > 0, b_{jj} > 0,\ j = 1, ..., p$ and $X = (x_{ij})$ with x_{ij}'s, $i \ge j$ be functionally independent real variables. Then show that

$$Y = AXB \Rightarrow \mathrm{d}Y = \left\{\prod_{j=1}^{p} a_{jj}^{j} b_{jj}^{p+1-j}\right\}\mathrm{d}X,$$

and

$$Z = A'X'B' \Rightarrow \mathrm{d}Z = \left\{\prod_{j=1}^{p} b_{jj}^{j} a_{jj}^{p+1-j}\right\}\mathrm{d}X. \tag{7.1.13}$$

7.1.4. Let $X = -X'$ be a $p \times p$ skew symmetric matrix of functionally independent $p(p-1)/2$ real variables and let $A, |A| \neq 0$, be a $p \times p$ constant matrix. Then prove that

$$Y = AXA', X' = -X, |A| \neq 0 \Rightarrow \mathrm{d}Y = |A|^{p-1}\mathrm{d}X. \tag{7.1.14}$$

7.1.5. Let X be a lower triangular $p \times p$ matrix of functionally independent real variables and $A = (a_{ij})$ be a lower trinagular matrix of constants with $a_{jj} > 0,\ j =$

$1, ..., p$. Then show that

$$Y = XA + A'X' \Rightarrow \mathrm{d}Y = 2^p \left\{ \prod_{j=1}^{p} a_{jj}^{p+1-j} \right\} \mathrm{d}X, \qquad (7.1.15)$$

and

$$Y = AX + X'A' \Rightarrow \mathrm{d}Y = 2^p \left\{ \prod_{j=1}^{p} a_{jj}^{j} \right\} \mathrm{d}X. \qquad (7.1.16)$$

7.1.6. Let X and A be as defined in Exercise 7.1.5. Then show that

$$Y = A'X + X'A \Rightarrow \mathrm{d}Y = 2^p \left\{ \prod_{j=1}^{p} a_{jj}^{j} \right\} \mathrm{d}X, \qquad (7.1.17)$$

and

$$Y = AX' + XA' \Rightarrow \mathrm{d}Y = 2^p \left\{ \prod_{j=1}^{p} a_{jj}^{p+1-j} \right\} \mathrm{d}X. \qquad (7.1.18)$$

7.1.7. Consider the transformation $Y = AX$ where

$$Y = \begin{bmatrix} y_{11} & y_{12} \\ 0 & y_{22} \end{bmatrix}, \; X = \begin{bmatrix} x_{11} & x_{12} \\ 0 & x_{22} \end{bmatrix}, \; A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}.$$

Writing AX explicitly and then taking the differentials and wedge products show that $\mathrm{d}Y = 12\mathrm{d}X$ and verifty that $J = \prod_{j=1}^{p} a_{jj}^{p+1-j} = (2^2)(3^1) = 12$.

7.1.8. Let Y and X be as in Exercise 7.1.7 and consider the transformation $Y = XB$ where, $B = A$ in Exercise 7.1.7. Then writing XB explicitly, taking differentials and then the wedge products show that $\mathrm{d}Y = 18\mathrm{d}X$ and verify the result that $J = \prod_{j=1}^{p} b_{jj}^{j} = (2^1)(3^2) = 18$.

7.1.9. Let Y, X, A be as in Exercise 7.1.7. Consider the transformation $Y = AX + X'A'$. Evaluate the Jacobian from first principles of taking differentials and wedge products and then verify the result that the Jacobian is $2^p = 2^2 = 4$ times the Jacobian in Exercise 7.1.7.

7.1.10. Let Y, X, B be as in Exercise 7.1.8. Consider the transformation $Y = XB + B'X'$. Evaluate the Jacobian from first principles and then verify that the Jacobian is $2^p = 2^2 = 4$ times the Jacobian in Exercise 7.1.8.

7.2. Jacobians in Some Nonlinear Transformations

Some basic nonlinear transformations will be considered in this section and some more results will be given in the exercises at the end of this section. The most popular nonlinear transformation is when a positive definite (naturally real symmetric also) matrix is decomposed into a triangular matrix and its transpose. This will be discussed first.

Example 7.2.1. Let X be $p \times p$, symmetric positive definite and let $T = (t_{ij})$ be a lower triangular matrix. Consider the transformation $Z = TT'$. Obtain the conditions for this transformation to be one-to-one and then evaluate the Jacobian.

Solution 7.2.1.

$$X = (x_{ij}) = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ \vdots & \vdots & \dots & \vdots \\ x_{p1} & x_{p2} & \dots & x_{pp} \end{bmatrix}$$

with $x_{ij} = x_{ji}$ for all i and j, $X = X' > 0$. When X is positive definite, that is, $X > 0$ then $x_{jj} > 0, j = 1, ..., p$ also.

$$TT' = \begin{bmatrix} t_{11} & 0 & \dots & 0 \\ t_{21} & t_{22} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ t_{p1} & t_{p2} & \dots & t_{pp} \end{bmatrix} \begin{bmatrix} t_{11} & t_{21} & \dots & t_{p1} \\ 0 & t_{22} & \dots & t_{p2} \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \cdots & t_{pp} \end{bmatrix} = X \Rightarrow$$

$x_{11} = t_{11}^2 \Rightarrow t_{11} = \pm\sqrt{x_{11}}$. This can be made unique if we impose the condition $t_{11} > 0$. Note that $x_{12} = t_{11}t_{21}$ and this means that t_{21} is unique if $t_{11} > 0$. Continuing like this, we see that for the transformation to be unique it is sufficient that $t_{jj} > 0, j = 1, ..., p$. Now, observe that,

$$x_{11} = t_{11}^2, x_{22} = t_{21}^2 + t_{22}^2, ..., x_{pp} = t_{p1}^2 + ... + t_{pp}^2$$

and $x_{12} = t_{11}t_{21}, ..., x_{1p} = t_{11}t_{p1}$, and so on.

$$
\begin{aligned}
\frac{\partial x_{11}}{\partial t_{11}} &= 2t_{11}, \frac{\partial x_{11}}{\partial t_{21}} = 0, ..., \frac{\partial x_{11}}{\partial t_{p1}} = 0, \\
\frac{\partial x_{12}}{\partial t_{21}} &= t_{11}, ..., \frac{\partial x_{1p}}{\partial t_{p1}} = t_{11}, \\
\frac{\partial x_{22}}{\partial t_{22}} &= 2t_{22}, \frac{\partial x_{22}}{\partial t_{31}} = 0, \cdots, \frac{\partial x_{22}}{\partial t_{p1}} = 0,
\end{aligned}
$$

and so on. Taking the x_{ij}'s in the order $x_{11}, x_{12}, \cdots, x_{1p}, x_{22}, \cdots, x_{2p}, \cdots, x_{pp}$ and the t_{ij}'s in the order $t_{11}, t_{21}, t_{22}, \cdots, t_{pp}$ we have the Jacobian matrix a triangular matrix with the diagonal elements as follows: t_{11} is repeated p times, t_{22} is repeated $p-1$ times and so on, and finally t_{pp} appearing once. The number 2 is appearing a total of p times. Hence the determinant is the product of the diagonal elements, giving,

$$2^p t_{11}^p t_{22}^{p-1} \cdots t_{pp}.$$

Therefore, for $X = X' > 0, T = (t_{ij}),\ t_{ij} = 0, i < j,\ t_{jj} > 0, j = 1, \cdots p$ we have

Theorem 7.2.1. *Let $X = X' > 0$ be a $p \times p$ real symmetric positive definite matrix and let $X = TT'$ where T is lower triangular with positive diagonal elements, $t_{jj} > 0,\ j = 1, ..., p$. Then*

$$X = TT' \Rightarrow \mathrm{d}X = 2^p \left\{ \prod_{j=1}^{p} t_{jj}^{p+1-j} \right\} \mathrm{d}T. \tag{7.2.1}$$

Example 7.2.2. If X is $p \times p$ real symmetric positive defi nite then evaluate the following integral, we will call it *matrix-variate real gamma* , denoted by $\Gamma_p(\alpha)$:

$$\Gamma_p(\alpha) = \int_X |X|^{\alpha - \frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(X)} \mathrm{d}X \tag{7.2.2}$$

and show that

$$\Gamma_p(\alpha) = \pi^{\frac{p(p-1)}{4}} \Gamma(\alpha) \Gamma\left(\alpha - \frac{1}{2}\right) \cdots \Gamma\left(\alpha - \frac{p-1}{2}\right) \tag{7.2.3}$$

for $\Re(\alpha) > \frac{p-1}{2}$.

Solution 7.2.2. Make the transformation $X = TT'$ where T is lower triangular with positive diagonal elements. Then

$$|TT'| = \prod_{j=1}^{p} t_{jj}^2, \ \mathrm{d}X = 2^p \left\{ \prod_{j=1}^{p} t_{jj}^{p+1-j} \right\} \mathrm{d}T$$

and

$$\mathrm{tr}(X) = t_{11}^2 + (t_{21}^2 + t_{22}^2) + ... + (t_{p1}^2 + ... + t_{pp}^2).$$

Then substituting these, the integral over X reduces to the following:

$$\int_X |X|^{\alpha - \frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(X)} \mathrm{d}X = \int_T \left\{ \prod_{j=1}^{p} \int_0^\infty 2 t_{jj}^{\alpha - \frac{j}{2}} \mathrm{e}^{-t_{jj}^2} \mathrm{d}t_{jj} \right\} \prod_{i>j} \int_{-\infty}^{\infty} \mathrm{e}^{-t_{ij}^2} \mathrm{d}t_{ij}.$$

Observe that

$$2 \int_0^\infty t_{jj}^{\alpha - \frac{j}{2}} \mathrm{e}^{-t_{jj}^2} \mathrm{d}t_{jj} = \Gamma\left(\alpha - \frac{j-1}{2}\right), \ \Re(\alpha) > \frac{j-1}{2},$$

$$\int_{-\infty}^{\infty} \mathrm{e}^{-t_{ij}^2} \mathrm{d}t_{ij} = \sqrt{\pi}$$

and there are $p(p-1)/2$ factors in $\prod_{i>j}$ and hence

$$\Gamma_p(\alpha) = \pi^{\frac{p(p-1)}{2}} \Gamma(\alpha) \Gamma\left(\alpha - \frac{1}{2}\right) \cdots \Gamma\left(\alpha - \frac{p-1}{2}\right)$$

and the condition $\Re(\alpha - \frac{j-1}{2}), \ j = 1, \cdots, p \Rightarrow \Re(\alpha) > \frac{p-1}{2}$. This establishes the result.

Notation 7.2.1.

$\Gamma_p(\alpha)$: **Real matrix-variate gamma**

Definition 7.2.1. **Real matrix-variate gamma** $\Gamma_p(\alpha)$: It is defined by equations (7.2.2) and (7.2.3) where (7.2.2) gives the integral representation and (7.2.3) gives the explicit form.

Remark 7.2.1. If we try to evaluate the integral $\int_X |X|^{\alpha-\frac{p+1}{2}} e^{-\text{tr}(X)} \mathrm{d}X$ from first principles, as a multiple integral, notice that even for $p = 3$ the integral is practically impossible to evaluate. For $p = 2$ one can evaluate after going through several stages.

The transformation in (7.2.1) is a nonlinear transformation whereas in (7.1.4) it is a general linear transformation involving mn functionally independent real x_{ij}'s. When X is a square and nonsingular matrix its regular inverse X^{-1} exists and the transformation $Y = X^{-1}$ is a one-to-one nonlinear transformation. What will be the Jacobian in this case?

Theorem 7.2.2. *For $Y = X^{-1}$ where X is a $p \times p$ nonsingular matrix we have*

$$\begin{aligned} Y = X^{-1}, |X| \neq 0 \Rightarrow \mathrm{d}Y &= |X|^{-2p} \text{ for a general } X \\ &= |X|^{-(p+1)} \text{ for } X = X'. \end{aligned} \tag{7.2.4}$$

Proof 7.2.1. This can be proved by observing the following: When X is nonsingular, $XX^{-1} = I$ where I denotes the identity matrix. Taking differentials on both sides we have

$$\begin{aligned} (\mathrm{d}X)X^{-1} + X(\mathrm{d}X^{-1}) = O \Rightarrow \\ (\mathrm{d}X^{-1}) = -X^{-1}(\mathrm{d}X)X^{-1} \end{aligned} \tag{7.2.5}$$

where $(\mathrm{d}X)$ means the matrix of differentials. Now we can apply Theorem 7.1.4 treating X^{-1} as a constant matrix because it is free of the differentials since we are taking only the wedge product of differentials on the left side.

Note 7.2.1. If the square matrix X is nonsingular and skew symmetric then proceeding as above it follows that

$$Y = X^{-1},\ |X| \neq 0,\ X' = -X \Rightarrow \mathrm{d}Y = |X|^{-(p-1)}\mathrm{d}X. \tag{7.2.6}$$

Note 7.2.2. If X is nonsingular and lower or upper triangular then, proceeding as before we have

$$Y = X^{-1} \Rightarrow \mathrm{d}Y = |X|^{-(p+1)} \tag{7.2.7}$$

where $|X| \neq 0,\ X$ is lower or upper triangular.

Theorem 7.2.3. *Let $X = (x_{ij})$ be $p \times p$ symmetric positive definite matrix of functionally independent real variables with $x_{jj} = 1,\ j = 1, ..., p$. Let $T = (t_{ij})$ be a* *lower triangular matrix of functionally independent real variables with* $t_{jj} > 0,\ j = 1,...,p$. *Then*

$$X = TT',\ with\ \sum_{j=1}^{i} t_{ij}^2 = 1,\ i = 1,...,p \Rightarrow$$
$$\mathrm{d}X = \left\{\prod_{j=2}^{p} t_{jj}^{p-j}\right\}\mathrm{d}T, \quad (7.2.8)$$

and

$$X = T'T,\ with\ \sum_{i=j}^{p} t_{ij}^2 = 1,\ j = 1,...,p \Rightarrow$$
$$\mathrm{d}X = \left\{\prod_{j=1}^{p-1} t_{jj}^{j-1}\right\}\mathrm{d}T. \quad (7.2.9)$$

Proof 7.2.2. Since X is symmetric with $x_{jj} = 1,\ j = 1,...,p$ there are only $p(p-1)/2$ variables in X. When $X = TT'$ take the x_{ij}'s in the order $x_{21},...,x_{p1},x_{32},...,x_{p2},...,x_{pp-1}$ and the t_{ij}'s also in the same order and form the matrix of partial derivatives. We obtain a triangular format and the product of the diagonal elements gives the required Jacobian.

Example 7.2.3. Let $R = (r_{ij})$ be a $p \times p$ real symmetric positive definite matrix such that $r_{jj} = 1,\ j = 1,...,p,\ -1 < r_{ij} = r_{ji} < 1,\ i \neq j$. (This is known as the correlation matrix in statistical theory). Then show that

$$f(R) = \frac{[\Gamma(\alpha)]^p}{\Gamma_p(\alpha)}|R|^{\alpha-\frac{p+1}{2}}$$

is a density function for $\Re(\alpha) > \frac{p-1}{2}$.

Solution 7.2.3. Since R is positive definite $f(R) \geq 0$ for all R. Let us check the total integral. Let T be a lower triangular matrix as defined in the above theorem and let $R = TT'$. Then

$$\int_R |R|^{\alpha-\frac{p+1}{2}}\mathrm{d}R = \int_T \left\{\prod_{j=2}^{p}(t_{jj}^2)^{\alpha-\frac{j+1}{2}}\right\}\mathrm{d}T.$$

Observe that

$$t_{jj}^2 = 1 - t_{j1}^2 - ... - t_{jj-1}^2$$

where $-1 < t_{ij} < 1,\ i > j$. Then let

$$B = \int_R |R|^{\alpha - \frac{p+1}{2}} \mathrm{d}R = \prod_{i>j} \Delta_j$$

where

$$\Delta_j = \int_{w_j} (1 - t_{j1}^2 - ... - t_{jj-1}^2)^{\alpha - \frac{j+1}{2}} \mathrm{d}t_{j1} \cdots \mathrm{d}t_{jj-1}$$

where $w_j = (t_{jk}),\ -1 < t_{jk} < 1,\ k = 1, ..., j-1, \sum_{k=1}^{j-1} t_{jk}^2 < 1.$

Evaluating the integral with the help of Dirichlet integral of Chapter 1 and then taking the product we have the final result showing that $f(R)$ is a density.

Exercises 7.2.

7.2.1 Let $X = X' > 0$ be $p \times p$. Let $T = (t_{ij})$ be an upper triangular matrix with positive diagonal elements. Then show that

$$X = TT' \Rightarrow \mathrm{d}X = 2^p \left\{ \prod_{j=1}^{p} t_{jj}^{j} \right\} \mathrm{d}T. \tag{7.2.10}$$

7.2.2. Let $x_1, \cdots, x_p$ be real scalar variables. Let $y_1 = x_1 + \cdots + x_p,\ y_2 = x_1x_2 + x_1x_3 + \cdots x_{p-1}x_p$ (sum of products taken two at a time), $\cdots, y_k = x_1 \cdots x_k$. Then for $x_j > 0, j = 1, \cdots, k$ show that

$$\mathrm{d}y_1 \wedge \cdots \wedge \mathrm{d}y_k = \left\{ \prod_{i=1}^{p-1} \prod_{j=i+1}^{p} |x_i - x_j| \right\} \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_p. \tag{7.2.11}$$

7.2.3. Let $x_1, \cdots, x_p$ be real scalar variables. Let

$$\begin{aligned} x_1 &= r\sin\theta_1 \\ x_j &= r\cos\theta_1\cos\theta_2\cdots\cos\theta_{j-1}\sin\theta_j,\ \ j=2,3,\cdots,p-1 \\ x_p &= r\cos\theta_1\cos\theta_2\cdots\cos\theta_{p-1} \end{aligned}$$

for $r>0, -\frac{\pi}{2}<\theta_j\le\frac{\pi}{2}, j=1,\cdots,p-2, -\pi<\theta_{p-1}\le\pi$. Then show that

$$\mathrm{d}x_1\wedge\cdots\wedge\mathrm{d}x_p = r^{p-1}\left\{\prod_{j=1}^{p-1}|\cos\theta_j|^{p-j-1}\right\}\mathrm{d}r\wedge\mathrm{d}\theta_1\wedge\cdots\wedge\mathrm{d}\theta_{p-1}. \tag{7.2.12}$$

7.2.4. Let $X=\frac{T}{|T|}$ where X and T are $p\times p$ lower triangular or upper triangular matrices of functionally independent real variables with positive diagonal elements. Then show that

$$\mathrm{d}X = (p-1)|T|^{-p(p+1)/2}\mathrm{d}T. \tag{7.2.13}$$

7.2.5. For real symmetric positive definite matrices X and Y show that

$$\lim_{t\to\infty}\left|I+\frac{XY}{t}\right|^{-t} = \mathrm{e}^{-\mathrm{tr}(XY)} = \lim_{t\to\infty}\left|I-\frac{XY}{t}\right|^{t}. \tag{7.2.14}$$

7.2.6. Let $X=(x_{ij}), W=(w_{ij})$ be lower triangular $p\times p$ matrices of distinct real variables with $x_{jj}>0, w_{jj}>0, j=1,\cdots,p, \sum_{k=1}^{j}w_{jk}^2=1, j=1,\cdots,p$. Let $D=\mathrm{diag}(\lambda_1,\cdots,\lambda_p), \lambda_j>0, j=1,\cdots,p$, real and distinct where $\mathrm{diag}(\lambda_1,\cdots,\lambda_p)$ denotes a diagonal matrix with diagonal elements $\lambda_1,\cdots,\lambda_p$. Show that

$$X=DW \Rightarrow \mathrm{d}X = \left\{\prod_{j=1}^{p}\lambda_j^{j-1}w_{jj}^{-1}\right\}\mathrm{d}D\wedge\mathrm{d}W. \tag{7.2.15}$$

7.2.7. Let X, A, B be $p\times p$ nonsingular matrices where A and B are constant matrices and X is a matrix of functionally independent real variables. Then, ignoring the sign, show that

$$\begin{aligned} Y = AX^{-1}B \Rightarrow \mathrm{d}Y &= |AB|^p|X|^{-2p}\mathrm{d}X \text{ for a general } X, && (7.2.16)\\ &= |AX^{-1}|^{p+1}\mathrm{d}X \text{ for } X=X',\ B=A', && (7.2.17)\\ &= |AX^{-1}|^{p-1} \text{ for } X'=-X, B=A'. && (7.2.18) \end{aligned}$$

7.2.8. Let X and A be $p\times p$ matrices where A is a nonsingular constant matrix and X is a matrix of functionally independent real variables such that $A+X$ is nonsingular. Then, ignoring the sign, show that

$$Y = (A+X)^{-1}(A-X) \text{ or } (A-X)(A+X)^{-1} \Rightarrow$$

$$\mathrm{d}Y = 2^{p^2}|A|^p|A+X|^{-2p}\mathrm{d}X \text{ for a general } X, \tag{7.2.19}$$

$$= 2^{\frac{p(p+1)}{2}}|I+X|^{-(p+1)}\mathrm{d}X \text{ for } A=I,\ X=X'. \tag{7.2.20}$$

7.2.9. Let X and A be real $p\times p$ lower triangular matrices where A is a constant matrix and X is a matrix of functionally independent real variables such that A and $A+X$ are nonsingular. Then, ignoring the sign, show that

$$Y = (A+X)^{-1}(A-X) \Rightarrow$$

$$\mathrm{d}Y = 2^{\frac{p(p+1)}{2}}|A+X|_+^{-(p+1)}\left\{\prod_{j=1}^{p}|a_{jj}|^{p+1-j}\right\}\mathrm{d}X, \tag{7.2.21}$$

and

$$Y = (A-X)(A+X)^{-1} \Rightarrow$$

$$\mathrm{d}Y = 2^{\frac{p(p+1)}{2}}|A+X|_+^{-(p+1)}\left\{\prod_{j=1}^{p}|a_{jj}|^{j}\right\}\mathrm{d}X \tag{7.2.22}$$

where $|\cdot|_+$ denotes that the absolute value is taken.

7.2.10. State and prove the corresponding results in Exercise 7.2.9 for upper trinagular matrices.

7.3. Transformations Involving Orthonormal Matrices

Here we will consider a few matrix transformations involvig orthonormal and semiorthonormal matrices. Some basic results that we need later on are discussed here. For more on these types and various other types of transformations see Mathai (1997). Since the proofs in many of the results are too involved and beyond the scope of this School we will not go into the details of the proofs.

Theorem 7.3.1. *Let X be a $p\times n, n\geq p$, matrix of rank p and of functionally independent real variables, and let $X = TU_1'$ where T is $p\times p$ lower triangular with* *distinct nonzero diagonal elements and* U_1' *a unique* $n \times p$ *semiorthonormal matrix,* $U_1'U_1 = I_p$*, all are of functionally independent real variables. Then*

$$X = TU_1' \Rightarrow \mathrm{d}X = \left\{\prod_{j=1}^{p} |t_{jj}|^{n-j}\right\} \mathrm{d}T \wedge U_1'(\mathrm{d}U_1) \tag{7.3.1}$$

where

$$\int \wedge U_1'(\mathrm{d}U_1) = \frac{2^p \pi^{\frac{pn}{2}}}{\Gamma_p\left(\frac{n}{2}\right)}. \tag{7.3.2}$$

(see equation (7.2.3) for $\Gamma_p(\cdot)$ *).*

Proof 7.3.1. For proving the main part of the theorem take the differentials on both sides of $X = TU_1'$ and then take the wedge product of the differentials systematically. Since it involves many steps the proof of the main part is not given here. The second part can be proved without much difficulty. Consider the X the $p \times n, n \geq p$ real matrix. Observe that

$$\mathrm{tr}(XX') = \sum_{ij} x_{ij}^2,$$

that is, the sum of squares of all elements in $X = (x_{ij})$ and there are np terms in $\sum_{ij} x_{ij}^2$. Now consider the integral

$$\begin{aligned}\int_X \mathrm{e}^{-\mathrm{tr}(X)} \mathrm{d}X &= \prod_{ij} \int_{-\infty}^{\infty} \mathrm{e}^{-x_{ij}^2} \mathrm{d}x_{ij} \\ &= \pi^{np/2}\end{aligned}$$

since each integral over x_{ij} gives $\sqrt{\pi}$. Now let us evaluate the same integral by using Theorem 7.3.1. Consider the same transformation as in Theorem 7.3.1, $X = TU_1'$. Then

$$\begin{aligned}\pi^{np/2} &= \int_X \mathrm{e}^{-\mathrm{tr}(X)} \mathrm{d}X \\ &= \int_T \left\{\prod_{j=1}^{p} |t_{jj}|^{n-j}\right\} \mathrm{e}^{-(\sum_{i \geq j} t_{ij}^2)} \mathrm{d}T \\ &\times \int \wedge U_1'(\mathrm{d}U_1).\end{aligned}$$

But for $0 < t_{jj} < \infty, -\infty < t_{ij} < \infty,\ i > j$ and U_1 unrestricted semiorthonormal, we have

$$\int_T \left\{\prod_{j=1}^{p} |t_{jj}|^{n-j}\right\} \mathrm{e}^{-(\sum_{i\geq j} t_{ij}^2)} \mathrm{d}T = 2^{-p}\Gamma_p\left(\frac{n}{2}\right) \tag{7.3.3}$$

observing that for $j = 1, ..., p$ the p integrals

$$\int_0^{\infty} |t_{jj}|^{n-j} \mathrm{e}^{-t_{jj}^2} \mathrm{d}t_{jj} = 2^{-1}\Gamma\left(\frac{n}{2} - \frac{j-1}{2}\right),\ n > j-1, \tag{7.3.4}$$

and each of the $p(p-1)/2$ integrals

$$\int_{-\infty}^{\infty} \mathrm{e}^{-t_{ij}^2} \mathrm{d}t_{ij} = \sqrt{\pi},\ i > j. \tag{7.3.5}$$

Now, substituting these the result in (7.3.2) is established.

Remark 7.3.1. For the transformation $X = TU_1'$ to be unique, either one can take T with the diagonal elements $t_{jj} > 0,\ j = 1, ..., p$ and U_1 unrestricted semiorthonormal matrix or $-\infty < t_{jj} < \infty$ and U_1 a unique semiorthonormal matrix.

From the outline of the proof of Theorem 7.3.1 we have the follwing result:

$$\int_{V_{p,n}} \wedge U_1'(\mathrm{d}U_1) = \frac{2^p \pi^{pn/2}}{\Gamma_p(\frac{n}{2})} \tag{7.3.6}$$

where $V_{p,n}$ is the *Stiefel manifold*, or the set of semiorthonormal matrices of the type U_1, $n \times p$, such that $U_1'U_1 = I_p$ where I_p is an identity matrix of order p. For $n = p$ the Stiefel manifold becomes the *full orthogonal group*, denoted by O_p. Then we have for, $n = p$,

$$\int_{O_p} \wedge U_1'(\mathrm{d}U_1) = \frac{2^p \pi^{p^2}}{\Gamma_p(\frac{n}{2})}. \tag{7.3.7}$$

Following through the same steps as in Theorem 7.3.1 we can have the following theorem involving an upper triangular matrix T_1.

Theorem 7.3.2. *Let X_1 be an $n\times p$, $n \geq p$, matrix of rank p and of functionally independent real variables and let $X_1 = U_1 T_1$ where T_1 is a real $p \times p$ upper* *triangular matrix with distinct nonzero diagonal elements and* U_1 *is a unique real* $n \times p$ *semiorthonormal matrix, that is,* $U_1'U_1 = I_p$*. Then, ignoring the sign,*

$$X_1 = U_1 T_1 \Rightarrow \mathrm{d}X_1 = \left\{ \prod_{j=1}^{p} |t_{jj}|^{n-j} \right\} \mathrm{d}T_1 \ \wedge U_1'(\mathrm{d}U_1) \qquad (7.3.8)$$

As a corollary to Theorem 7.3.1 or independently we can prove the following result:

Corollary 7.3.1. *Let* X_1, T, U_1 *be as defined in Theorem 7.3.1 with the diagonal elements of* T *positive, that is,* $t_{jj} > 0, \ j = 1, ..., p$ *and* U_1 *an arbitrary semiorthonormal matrix, and let* $A = XX'$*, which implies,* $A = TT'$ *also. Then*

$$A = XX' \Rightarrow \qquad (7.3.9)$$

$$\mathrm{d}A = 2^p \left\{ \prod_{j=1}^{p} t_{jj}^{p+1-j} \right\} \mathrm{d}T \qquad (7.3.10)$$

$$\Rightarrow$$

$$\mathrm{d}T = 2^{-p} \left\{ \prod_{j=1}^{p} t_{jj}^{-p-1-j} \right\} \mathrm{d}A. \qquad (7.3.11)$$

In practical applications we would like to have $\mathrm{d}X$ *in terms of* $\mathrm{d}A$ *or vice versa after integrating out* $\wedge U_1'(\mathrm{d}U_1)$ *over the Stiefel manifold* $V_{p,n}$*. Hence we have the following corollary*

Corollary 7.3.2. *Let* X_1, T, U_1 *be as defined in Theorem 7.3.1 with* $t_{jj} > 0, \ j = 1, ..., p$ *and let* $S = XX' = TT'$*. Then, after integrating out* $\wedge U_1'(\mathrm{d}U_1)$ *we have*

$$X = TU_1' \textit{ and } S = XX' = TT' \Rightarrow \tag{7.3.12}$$

$$\mathrm{d}X = \left\{\prod_{j=1}^{p}(t_{jj}^2)^{\frac{n}{2}-\frac{j}{2}}\right\}\mathrm{d}T \wedge U_1'(\mathrm{d}U_1) \tag{7.3.13}$$

$$\int_{V_{p,n}} \wedge U_1'(\mathrm{d}U_1) = \frac{2^p\pi^{np/2}}{\Gamma_p(\frac{n}{2})} \tag{7.3.14}$$

$$\mathrm{d}S = 2^p\left\{\prod_{j=1}^{p}(t_{jj}^2)^{\frac{p+1}{2}-\frac{j}{2}}\right\}\mathrm{d}T \tag{7.3.15}$$

$$|S| = \prod_{j=1}^{p} t_{jj}^2 \tag{7.3.16}$$

and, finally,

$$\mathrm{d}X = |S|^{\frac{n}{2}-\frac{p+1}{2}}\mathrm{d}S. \tag{7.3.17}$$

Example 7.3.1. Let X be a $p \times n, n \geq p$ random matrix having an np-variate real Gaussian density

$$f(X) = \frac{\mathrm{e}^{-\frac{1}{2}\mathrm{tr}(V^{-1}XX')}}{(2\pi)^{\frac{np}{2}}|V|^{\frac{n}{2}}},\ V = V' > 0.$$

Evaluate the density of $S = XX'$.

Solution 7.3.1. Consider the transformation as in Theorem 7.3.1, $X = TU_1'$ where T is a $p \times p$ lower triangular matrix with positive diagonal elements and U_1 is a arbitrary $n \times p$ semiorthonormal matrix, $U_1'U_1 = I_p$. Then

$$\mathrm{d}X = \left\{\prod_{j=1}^{p} t_{jj}^{n-j}\right\}\mathrm{d}T \wedge U_1(\mathrm{d}U_1).$$

Integrating out $\wedge U_1(\mathrm{d}U_1)$ we have the marginal density of T, denoted by $g(T)$. That is,

$$g(T)\mathrm{d}T = \frac{2^p\mathrm{e}^{-\frac{1}{2}\mathrm{tr}(V^{-1}TT')}}{2^{np/2}|V|^{\frac{n}{2}}}\left\{\prod_{j=1}^{p} t_{jj}^{n-j}\right\}\mathrm{d}T.$$

Now substituting from Corollary 7.3.2, S and dS in terms of T and dT we have the density of S, denoted by, $h(S)$, given by

$$h(S) = C_1|S|^{\frac{n}{2}-\frac{p+1}{2}}\mathrm{e}^{-\frac{1}{2}V^{-1}S},\ S = S' > 0.$$

Since the total integral, $\int_S h(S)\mathrm{d}S = 1$, we have

$$C_1 = [2^{np/2}\Gamma_p\left(\frac{n}{2}\right)|V|^{n/2}]^{-1}.$$

Exercises 7.3.

7.3.1. Let $X = (x_{ij})$, $W = (w_{ij})$ be $p \times p$ lower triangular matrices of distinct real variables such that $x_{jj} > 0$, $w_{jj} > 0$ and $\sum_{k=1}^{j} w_{jk}^2 = 1$, $j = 1, ..., p$. Let $D = \mathrm{diag}(\lambda_1, ..., \lambda_p)$, $\lambda_j > 0$, $j = 1, ..., p$ be real positive and distinct. Let $D^{\frac{1}{2}} = \mathrm{diag}(\lambda_1^{\frac{1}{2}}, ..., \lambda_p^{\frac{1}{2}})$. Then show that

$$X = DW \Rightarrow \mathrm{d}X = \left\{\prod_{j=1}^{p} \lambda_j^{j-1} w_{jj}^{-1}\right\} \mathrm{d}D \wedge \mathrm{d}W. \tag{7.3.18}$$

7.3.2. Let X, D, W be as defined in Exercise 7.3.1 then show that

$$X = D^{\frac{1}{2}}W \Rightarrow \mathrm{d}X = \left\{2^{-p}\prod_{j=1}^{p}(\lambda_j^{\frac{1}{2}})^{j-2} w_{jj}^{-1}\right\} \mathrm{d}D \wedge \mathrm{d}W. \tag{7.3.19}$$

7.3.3. Let X, D, W be as defined in Exercise 7.3.1 then show that

$$X = D^{\frac{1}{2}}WW'D^{\frac{1}{2}} \Rightarrow \mathrm{d}X = \left\{\prod_{j=1}^{p} \lambda_j^{\frac{p-1}{2}} w_{jj}^{p-j}\right\} \mathrm{d}D \wedge \mathrm{d}W, \tag{7.3.20}$$

and

$$X = W'DW \Rightarrow \mathrm{d}X = \left\{\prod_{j=1}^{p}(\lambda_j w_{jj})^{j-1}\right\} \mathrm{d}D \wedge \mathrm{d}W. \tag{7.3.21}$$

7.3.4. Let X, D, W be as defined in Exercise 7.3.1 then show that

$$X = WD \Rightarrow \mathrm{d}X = \left\{\prod_{j=1}^{p} \lambda_j^{p-j} w_{jj}^{-1}\right\} \mathrm{d}D \wedge \mathrm{d}W, \tag{7.3.22}$$

$$X = WD^{\frac{1}{2}} \Rightarrow \mathrm{d}X = \left\{2^{-p} \prod_{j=1}^{p} (\lambda_j^{\frac{1}{2}})^{p-j-1} w_{jj}^{-1}\right\} \mathrm{d}D \wedge \mathrm{d}W, \tag{7.3.23}$$

$$X = WDW' \Rightarrow \mathrm{d}X = \left\{\prod_{j=1}^{p} (\lambda_j w_{jj})^{p-j}\right\} \mathrm{d}D \wedge \mathrm{d}W, \tag{7.3.24}$$

and

$$X = D^{\frac{1}{2}} W' W D^{\frac{1}{2}} \Rightarrow \mathrm{d}X = \left\{\prod_{j=1}^{p} \lambda_j^{\frac{p-1}{2}} w_{jj}^{j-1}\right\} \mathrm{d}D \wedge \mathrm{d}W. \tag{7.3.25}$$

7.3.5. Let X, T, U be $p \times p$ matrices of functionally independent real variables where all the principal minors of X are nonzero, T is lower triangular and U is lower triangular with unit diagonal elements. Then, ignoring the sign, show that

$$X = TU' \Rightarrow \mathrm{d}X = \left\{\prod_{j=1}^{p} |t_{jj}|^{p-j}\right\} \mathrm{d}T \wedge \mathrm{d}U, \tag{7.3.26}$$

and

$$X = T'U \Rightarrow \mathrm{d}X = \left\{\prod_{j=1}^{p} |t_{jj}|^{j-1}\right\} \mathrm{d}T \wedge \mathrm{d}U. \tag{7.3.27}$$

7.3.6. Let X be a $p \times p$ symmetric matrix of functionally independent real variables and with distinct and nonzero eigenvalues $\lambda_1 > \lambda_2 > ... > \lambda_p$ and let $D = \mathrm{diag}(\lambda_1, ..., \lambda_p)$, $\lambda_j \neq 0$, $j = 1, ..., p$. Let U be a unique $p \times p$ orthonormal matrix $U'U = I = UU'$ such that $X = UDU'$. Then, ignoring the sign, show that

$$\mathrm{d}X = \left\{\prod_{i=1}^{p-1} \prod_{j=i+1}^{p} |\lambda_i - \lambda_j|\right\} \mathrm{d}D \wedge U'(\mathrm{d}U). \tag{7.3.28}$$

7.3.7. For a 3×3 matrix X such that $X = X' > 0$ and $I - X > 0$ show that

$$\int_X \mathrm{d}X = \frac{\pi^2}{90}.$$

7.3.8. For a $p \times p$ matrix X of p^2 functionally independent real variables with positive eigenvalues, show that

$$\int_Y |Y'Y|^{\alpha - \frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(Y'Y)} \mathrm{d}Y = 2^{-p} \frac{\Gamma_p(\alpha - \frac{1}{2})\pi^{\frac{p^2}{2}}}{\Gamma_p(\frac{p}{2})}.$$

7.3.9. Let X be a $p \times p$ matrix of p^2 functionally independent real variables. Let $D = \mathrm{diag}(\mu_1, ..., \mu_p), \mu_1 > \mu_2 > ... > \mu_p$, μ_j real for $j = 1, ..., p$. Let U and V be orthonormal matrices such that

$$X = UDV' \Rightarrow$$
$$\mathrm{d}X = \left\{\prod_{i<j} |\mu_i^2 - \mu_j^2|\right\} \mathrm{d}D \wedge \mathrm{d}G \wedge \mathrm{d}H \tag{7.3.29}$$

where $(\mathrm{d}G) = U'(\mathrm{d}U)$ and $(\mathrm{d}H) = (\mathrm{d}V')V$, and the μ_j's are known as the singular values of X.

7.3.10. Let $\lambda_1 > ... > \lambda_p > 0$ be real variables and $D = \mathrm{diag}(\lambda_1, ..., \lambda_p)$. Show that

$$\int_D \mathrm{e}^{-\mathrm{tr}(D^2)} \left\{\prod_{i<j} |\lambda_i^2 - \lambda_j^2|\right\} \mathrm{d}D = \frac{[\Gamma_p(\frac{p}{2})]^2}{2^p \pi^{\frac{p^2}{2}}}.$$

7.4. Special Functions of Matrix Argument

Real scalar functions of matrix argument, when the matrices are real, will be dealt with in this chapter. It is difficult to develop a theory of functions of matrix argument for general matrices. The notations that we have used in Section 7.1 will be used in the present chapter also. A discussion of scalar functions of matrix argument when the elements of the matrices are in the complex domain may be seen from Mathai (1997).

7.4.1. Real matrix-variate scalar functions

When dealing with matrices it is often difficult to define uniquely fractional powers such as square roots even when the matrices are real square or even symmetric. For example

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, A_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, A_4 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

all give $A_1^2 = A_2^2 = A_3^2 = A_4^2 = I_2$ where I_2 is a 2×2 identity matrix. Thus, even for I_2, which is a nice, square, symmetric, positive definite matrix there are many matrices which qualify to be square roots of I_2. But if we confine to the class of positive definite matrices, when real, then for the square root of I_2 there is only one candidate, namely, $A_1 = I_2$ itself. Hence the development of the theory of scalar functions of matrix argument is confined to positive definite matrices, when real, and hermitian positive definite matrices when in the complex domain.

7.4.2. Real matrix-variate gamma

In the real scalar case the integral representation for a gamma function is the following:

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} \mathrm{e}^{-x} \mathrm{d}x, \;\; \Re(\alpha) > 0. \tag{7.4.1}$$

Let X be a $p \times p$ real symmetric positive definite matrix and consider the integral

$$\Gamma_p(\alpha) = \int_{X=X'>0} |X|^{\alpha - \frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(X)} \mathrm{d}X \tag{7.4.2}$$

where, when $p = 1$ the equation (7.4.2) reduces to (7.4.1). We have already evaluated this integral in earlier as an exercise to the basic nonlinear matrix transformatin $X = TT'$ where T is a lower triangular matrix with positive diagonal elements. Hence the derivation will not be repeated here. We will call $\Gamma_p(\alpha)$ the real matrix-variate gamma. Observe that for $p = 1$, $\Gamma_p(\alpha)$ reduces to $\Gamma(\alpha)$.

7.4.3. Real matrix-variate gamma density

With the help of (7.4.2) we can create the real matrix-variate gamma density as follows, where X is a $p \times p$ real symmetric positive definite matrix:

$$f(X) = \begin{cases} \frac{1 \times 1^{\alpha - \frac{p+1}{2}}}{\Gamma_p(\alpha)} \mathrm{e}^{-\mathrm{tr}(X)}, X = X^{'} > 0, \ \Re(\alpha) > \frac{p-1}{2} \\ 0, \ \text{elsewhere} . \end{cases} \tag{7.4.3}$$

If another parameter matrix is to be introduced then we obtain a gamma density with parameters $(\alpha, B), B = B^{'} > 0$, as follows:

$$f_1(X) = \begin{cases} \frac{|B|^{\alpha}}{\Gamma_p(\alpha)} |X|^{\alpha - \frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(BX)}, X = X' > 0, B = B' > 0, \Re(\alpha) > \frac{p-1}{2} \\ 0, \text{elsewhere}. \end{cases} \tag{7.4.4}$$

Remark 7.4.1. In $f_1(X)$ if B is replaced by $\frac{1}{2}V^{-1}, V = V' > 0$ and α is replaced by $\frac{n}{2}, n = p, p+1, ...$ then we have the most important density in multivariate statistical analysis known as the nonsingular Wishart density.

As in the scalar case, two matrix random variables X and Y are said to be independently distributed if the joint density of X and Y is the product of their marginal densities. We will examine the densities of some functions of independently distributed matrix random variables. To this end we will introduce a few more functions.

Definition 7.4.1. A real matrix-variate beta function, denoted by $B_p(\alpha,\beta)$, is defined as

$$B_p(\alpha,\beta) = \frac{\Gamma_p(\alpha)\Gamma_p(\beta)}{\Gamma_p(\alpha+\beta)}, \ \Re(\alpha) > \frac{p-1}{2}, \Re(\beta) > \frac{p-1}{2}. \tag{7.4.5}$$

The quantity in (7.4.5), analogous to the scalar case ($p = 1$), is the real matrix-variate beta function. Let us try to obtain an integral representation for the real matrix-variate beta function of (7.4.5). Consider

$$\Gamma_p(\alpha)\Gamma_p(\beta) = \left[\int_{X=X'>0} |X|^{\alpha - \frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(X)} \mathrm{d}X \right] \times \left[\int_{Y=Y'>0} |Y|^{\beta - \frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(Y)} \mathrm{d}Y \right]$$

where both X and Y are $p \times p$ matrices.

$$= \int \int |X|^{\alpha - \frac{p+1}{2}} |Y|^{\beta - \frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(X+Y)} \mathrm{d}X \wedge \mathrm{d}Y.$$

Put $U = X + Y$ for a fixed X. Then

$$Y = U - X \Rightarrow |Y| = |U - X| = |U||I - U^{-\frac{1}{2}}XU^{-\frac{1}{2}}|$$

where, for convenience, $U^{\frac{1}{2}}$ is the symmetric positive definite square root of U. Observe that when two matrices A and B are nonsingular where AB and BA are defined, even if they do not commute,

$$|I - AB| = |I - BA|$$

and if $A = A' > 0$ and $B = B' > 0$ then

$$|I - AB| = |I - A^{\frac{1}{2}}BA^{\frac{1}{2}}| = |I - B^{\frac{1}{2}}AB^{\frac{1}{2}}|.$$

Now,

$$\Gamma_p(\alpha)\Gamma_p(\beta) = \int_U \int_X |U|^{\beta-\frac{p+1}{2}}|X|^{\alpha-\frac{p+1}{2}}|I - U^{-\frac{1}{2}}XU^{-\frac{1}{2}}|^{\beta-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(U)}\mathrm{d}U \wedge \mathrm{d}X.$$

Let $Z = U^{-\frac{1}{2}}XU^{-\frac{1}{2}}$ for fixed U. Then $\mathrm{d}X = |U|^{\frac{p+1}{2}}\mathrm{d}Z$ by using Theorem 7.1.5. Now,

$$\Gamma_p(\alpha)\Gamma_p(\beta) = \int_Z |Z|^{\alpha-\frac{p+1}{2}}|I - Z|^{\beta-\frac{p+1}{2}}\mathrm{d}Z \int_{U=U'>0} |U|^{\alpha+\beta-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(U)}\mathrm{d}U.$$

Evaluation of the U-integral by using (7.4.2) yields $\Gamma_p(\alpha + \beta)$. Then we have

$$B_p(\alpha, \beta) = \frac{\Gamma_p(\alpha)\Gamma_p(\beta)}{\Gamma_p(\alpha + \beta)} = \int_Z |Z|^{\alpha-\frac{p+1}{2}}|I - Z|^{\beta-\frac{p+1}{2}}\mathrm{d}Z.$$

Since the integral has to remain non-negative we have $Z = Z' > 0, I - Z > 0$. Therefore, one representation of a real matrix-variate beta function is the following, which is also called the type-1 beta integral.

$$B_p(\alpha, \beta) = \int_{0<Z=Z'<I} |Z|^{\alpha-\frac{p+1}{2}}|I - Z|^{\beta-\frac{p+1}{2}}\mathrm{d}Z, \Re(\alpha) > \frac{p-1}{2},\ \Re(\beta) > \frac{p-1}{2}. \quad (7.4.6)$$

By making the transformation $V = I - Z$ note that α and β can be interchanged in the integral which also shows that $B_p(\alpha, \beta) = B_p(\beta, \alpha)$ in the integral representation also.

Let us make the following transformation in (7.4.6).

$$W = (I - Z)^{-\frac{1}{2}} Z (I - Z)^{-\frac{1}{2}} \Rightarrow W = (Z^{-1} - I)^{-1} \Rightarrow W^{-1} = Z^{-1} - I$$

$$\Rightarrow |W|^{-(p+1)} \mathrm{d}W = |Z|^{-(p+1)} \mathrm{d}Z \Rightarrow \mathrm{d}Z = |I + W|^{-(p+1)} \mathrm{d}W.$$

Under this transformation the integral in (7.4.6) becomes the following: Observe that

$$|Z| = |W||I + W|^{-1}, |I - Z| = |I + W|^{-1}.$$

$$B_p(\alpha,\beta) = \int_{W=W'>0} |W|^{\alpha - \frac{p+1}{2}} |I + W|^{-(\alpha+\beta)} \mathrm{d}W, \tag{7.4.7}$$

for $\Re(\alpha) > \frac{p-1}{2},\ \Re(\beta) > \frac{p-1}{2}$.

The representation in (7.4.7) is known as the *type-2 integral* for a real matrix-variate beta function. With the transformation $V = W^{-1}$ the parameters α and β in (7.4.7) will be interchanged. With the help of the type-1 and type-2 integral representations one can define the type-1 and type-2 beta densities in the real matrix-variate case.

Definition 7.4.2. **Real matrix-variate type-1 beta density for the $p \times p$ real symmetric positive definite matrix X such that $X = X' > 0,\ I - X > 0$.**

$$f_2(X) = \begin{cases} \frac{1}{B_p(\alpha,\beta)} |X|^{\alpha - \frac{p+1}{2}} |I - X|^{\beta - \frac{p+1}{2}} = 0 < X = X' < I,\ \Re(\alpha) > \frac{p-1}{2},\ \Re(\beta) > \frac{p-1}{2}, \\ 0,\ \text{elsewhere.} \end{cases} \tag{7.4.8}$$

Definition 7.4.3. **Real matrix-variate type-2 beta density for the $p \times p$ real symmetric positive definite matrix X.**

$$f_3(X) = \begin{cases} \frac{\Gamma_p(\alpha+\beta)}{\Gamma_p(\alpha)\Gamma_p(\beta)} |X|^{\alpha - \frac{p+1}{2}} |I + X|^{-(\alpha+\beta)}, X = X' > 0, \\ \qquad \Re(\alpha) > \frac{p-1}{2},\quad \Re(\beta) > \frac{p-1}{2} \\ 0,\ \text{elsewhere.} \end{cases} \tag{7.4.9}$$

Example 7.4.1. Let X_1, X_2 be $p \times p$ matrix random variables, independently distributed as (7.4.3) with parameters α_1 and α_2 respectively. Let $U = X_1 + X_2, V = (X_1 + X_2)^{-\frac{1}{2}} X_1 (X_1 + X_2)^{-\frac{1}{2}}, W = X_2^{-\frac{1}{2}} X_1 X_2^{-\frac{1}{2}}$. Evaluate the densities of U, V and W.

Solutions 7.4.1. The joint density of X_1 and X_2, denoted by $f(X_1,X_2)$, is available as the product of the marginal densities due to independence. That is,

$$f(X_1,X_2)=\frac{|X_1|^{\alpha_1-\frac{p+1}{2}}|X_2|^{\alpha_2-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(X_1+X_2)}}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)},\ X_1=X_1'>0,\ X_2=X_2'>0,$$
$$\Re(\alpha_1)>\frac{p-1}{2},\quad \Re(\alpha_2)>\frac{p-1}{2}. \tag{7.4.10}$$

$$U=X_1+X_2\Rightarrow|X_2|=|U-X_1|=|U||I-U^{-\frac{1}{2}}X_1U^{-\frac{1}{2}}|.$$

Then the joint density of $(U,U_1)=(X_1+X_2,X_1)$, the Jacobian is unity, is available as

$$f_1(U,U_1)=\frac{1}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)}|U_1|^{\alpha_1-\frac{p+1}{2}}|U|^{\alpha_2-\frac{p+1}{2}}|I-U^{-\frac{1}{2}}U_1U^{-\frac{1}{2}}|^{\alpha_2-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(U)}.$$

Put $U_2=U^{-\frac{1}{2}}U_1U^{-\frac{1}{2}}\Rightarrow \mathrm{d}U_1=|U|^{\frac{p+1}{2}}\mathrm{d}U_2$ for fixed U. Then the joint density of U and $U_2=V$ is available as the following:

$$f_2(U,V)=\frac{1}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)}|U|^{\alpha_1+\alpha_2-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(U)}|V|^{\alpha_1-\frac{p+1}{2}}|I-V|^{\alpha_2-\frac{p+1}{2}}.$$

Since $f_2(U,V)$ is a product of two functions of U and V, $U=U'>0, V=V'>0$, $I-V>0$ we see that U and V are independently distributed. The densities of U and V, denoted by $g_1(U)$, $g_2(V)$ are the following:

$$g_1(U)=c_1|U|^{\alpha_1+\alpha_2-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(U)}, U=U'>0$$

and

$$g_2(V)=c_2|V|^{\alpha_1-\frac{p+1}{2}}|I-V|^{\alpha_2-\frac{p+1}{2}},\ V=V'>0,\ I-V>0,$$

where c_1 and c_2 are the normalizing constants. But from the gamma density and type-1 beta density note that

$$c_1=\frac{1}{\Gamma_p(\alpha_1+\alpha_2)},\ c_2=\frac{\Gamma_p(\alpha_1+\alpha_2)}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)},\quad \Re(\alpha_1)>0,\quad \Re(\alpha_2)>0.$$

Hence U is gamma distributed with the parameter $(\alpha_1+\alpha_2)$ and V is type-1 beta distributed with the parameters α_1 and α_2 and further that U and V are independently distributed. For obtaining the density of $W=X_2^{-\frac{1}{2}}X_1X_2^{-\frac{1}{2}}$ start with (7.4.10). Change

(X_1, X_2) to (X_1, W) for fixed X_2. Then $\mathrm{d}X_1 = |X_2|^{\frac{p+1}{2}}\mathrm{d}W$. The joint density of X_2 and W, denoted by $f_{w,x_2}(W, X_2)$, is the following, observing that

$$\begin{aligned}\mathrm{tr}(X_1 + X_2) &= \mathrm{tr}[X_2^{\frac{1}{2}}(I + X_2^{-\frac{1}{2}}X_1X_2^{-\frac{1}{2}})X_2^{\frac{1}{2}}] \\ &= \mathrm{tr}[X_2^{\frac{1}{2}}(I + W)X_2^{\frac{1}{2}}] = \mathrm{tr}[(I + W)X_2] \\ &= \mathrm{tr}[(I + W)^{\frac{1}{2}}X_2(I + W)^{\frac{1}{2}}]\end{aligned}$$

by using the fact that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ for any two matrices where AB and BA are defined.

$$f_{w,x_2}(W, X_2) = \frac{1}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)}|W|^{\alpha_1-\frac{p+1}{2}}|X_2|^{\alpha_1+\alpha_2-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}[(I+W)^{\frac{1}{2}}X_2(I+W)^{\frac{1}{2}}]}.$$

Hence the marginal density of W, denoted by $g_w(W)$, is available by integrating out X_2 from $f_{w,x_2}(W, X_2)$. That is,

$$g_w(W) = \frac{1}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)}|W|^{\alpha_1-\frac{p+1}{2}}\int_{X_2=X_2'>0}|X_2|^{\alpha_1+\alpha_2-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}[(I+W)^{\frac{1}{2}}X_2(I+W)^{\frac{1}{2}}]}\mathrm{d}X_2.$$

Put $X_3 = (I + W)^{\frac{1}{2}}X_2(I + W)^{\frac{1}{2}}$ for fixed W, then $\mathrm{d}X_3 = |I + W|^{\frac{p+1}{2}}\mathrm{d}X_2$.
Then the integral becomes

$$\begin{aligned}&\int_{X_2=X_2'>0}|X_2|^{\alpha_1+\alpha_2-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}[(I+W)^{\frac{1}{2}}X_2(I+W)^{\frac{1}{2}}]}\mathrm{d}X_2 \\ &\qquad = \Gamma_p(\alpha_1 + \alpha_2)|I + W|^{-(\alpha_1+\alpha_2)}.\end{aligned}$$

Hence,

$$g_w(W) = \begin{cases}\frac{\Gamma_p(\alpha_1+\alpha_2)}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)}|W|^{\alpha_1-\frac{p+1}{2}}|I + W|^{-(\alpha_1+\alpha_2)}, W = W' > 0, \\ \qquad \Re(\alpha_1) > \frac{p-1}{2}, \quad \Re(\alpha_2) > \frac{p-1}{2} \\ 0, \text{elsewhere},\end{cases}$$

which is a type-2 beta density with the parameters α_1 and α_2. Thus, W is real matrix-variate type-2 beta distributed.

Exercises 7.4.

7.4.1. For a real $p\times p$ matrix X such that $X=X'>0, 0<X<I$ show that

$$\int_X \mathrm{d}X=\frac{[\Gamma_p(\frac{p+1}{2})]^2}{\Gamma_p(p+1)}.$$

7.4.2. Let X be a 2×2 real symmetric positive definite matrix with eigenvalues in $(0,1)$. Then show that

$$\int_{0<X<I}|X|^{\alpha}\mathrm{d}X=\frac{\pi}{(\alpha+1)(\alpha+2)(2\alpha+3)}.$$

7.4.3. For a 2×2 real positive definite matrix X show that

$$\int_X |I+X|^{-3}\mathrm{d}X=\frac{\pi}{6}.$$

7.4.4. For a 4×4 real positive definite matrix X such that $0<X<I$, show that

$$\int_X \mathrm{d}X=\frac{2\pi^4}{7!5}.$$

7.4.5. If the $p\times p$ real positive definite matrix random variable X is distributed as a real matrix-variate type-1 beta (having a type-1 beta density), evaluate the density of $Y=A^{\frac{1}{2}}XA^{\frac{1}{2}}$ where the constant matrix $A=A'>0$.

7.5. The Laplace Transform in the Matrix Case

If $f(x_1,\cdots,x_k)$ is a scalar function of the real scalar variables $x_1,\cdots,x_k$ then the Laplace transform of f, with the parameters $t_1,\cdots,t_k$, is given by

$$L_f(t_1,\cdots,t_k)=\int_0^{\infty}\cdots\int_0^{\infty}\mathrm{e}^{-t_1x_1-\cdots-t_kx_k}f(x_1,\cdots,x_k)\mathrm{d}x_1\wedge\cdots\wedge\mathrm{d}x_k. \tag{7.5.1}$$

If $f(X)$ is a real scalar function of the $p\times p$ real symmetric positive definite matrix X then the Laplace transform of $f(X)$ should be consistent with (7.5.1). When $X=X'$ there are only $p(p+1)/2$ distinct elements, either $x_{ij}'s, i\le j$ or $x_{ij}'s, i\ge j$. Hence what is needed is a linear function of all these variables. That is, in the exponent we should have the linear function $t_{11}x_{11}+(t_{21}x_{21}+t_{22}x_{22})+\cdots+(t_{p1}x_{p1}+\cdots+t_{pp}x_{pp})$. Even if we take a symmetric matrix $T=(t_{ij})=T'$ then the trace of TX,

$$\mathrm{tr}(TX)=t_{11}x_{11}+\cdots+t_{pp}x_{pp}+2\sum_{i<j=1}^{p}t_{ij}x_{ij}.$$

Hence if we take a symmetric matrix of parameters t_{ij}'s such that

$$T^*=\begin{bmatrix} t_{11} & \frac{1}{2}t_{12} & \cdots & \frac{1}{2}t_{1p}\\ \frac{1}{2}t_{21} & t_{22} & \cdots & \frac{1}{2}t_{2p}\\ \vdots & & & \\ \frac{1}{2}t_{p1} & \frac{1}{2}t_{p2} & \cdots & t_{pp}\end{bmatrix},\quad T^*=(t^*_{ij})\Rightarrow t^*_{jj}=t_{jj},t^*_{ij}=\frac{1}{2}t_{ij},\quad i\neq j$$

then

$$\mathrm{tr}(T^*X)=t_{11}x_{11}+\cdots+t_{pp}x_{pp}+\sum_{i=1}^{p}\sum_{j=1,i>j}^{p}t_{ij}x_{ij}.$$

Hence the Laplace transform in the matrix case, for real symmetric positive definite matrix X, is defined with the parameter matrix T^*.

Definition 7.5.1. Laplace transform in the matrix case.

$$L_f(T^*)=\int_{X=X'>0}\mathrm{e}^{-\mathrm{tr}(T^*X)}f(X)\mathrm{d}X, \tag{7.5.2}$$

whenever the integral is convergent.

Example 7.5.1. Evaluate the Laplace transform for the two-parameter gamma density in (7.4.4).

Solution 7.5.1. Here,

$$f(X)=\frac{|B|^{\alpha}}{\Gamma_p(\alpha)}|X|^{\alpha-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(BX)},X=X'>0,B=B'>0,\quad \Re(\alpha)>\frac{p-1}{2}. \tag{7.5.3}$$

Hence the Laplace transform of f is the following:

$$L_f(T^*)=\frac{|B|^{\alpha}}{\Gamma_p(\alpha)}\int_{X=X'>0}|X|^{\alpha-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(T^*X)}\mathrm{e}^{-\mathrm{tr}(BX)}\mathrm{d}X.$$

Note that since T^*, B and X are $p\times p$,

$$\mathrm{tr}(T^*X)+\mathrm{tr}(BX)=\mathrm{tr}[(B+T^*)X].$$

Thus for the integral to converge the exponent has to remain positive definite. Then the condition $B + T^* > 0$ is sufficient. Let $(B + T^*)^{\frac{1}{2}}$ be the symmetric positive definite square root of $B + T^*$. Then

$$\mathrm{tr}[(B + T^*)X] = \mathrm{tr}[(B + T^*)^{\frac{1}{2}} X (B + T^*)^{\frac{1}{2}}],$$
$$(B + T^*)^{\frac{1}{2}} X (B + T^*)^{\frac{1}{2}} = Y \Rightarrow \mathrm{d}X = |B + T^*|^{-\frac{p+1}{2}} \mathrm{d}Y$$

and

$$|X|^{\alpha - \frac{p+1}{2}} \mathrm{d}X = |B + T^*|^{-\alpha} |Y|^{\alpha - \frac{p+1}{2}} \mathrm{d}Y.$$

Hence,

$$\begin{aligned} L_f(T^*) &= \frac{|B|^\alpha}{\Gamma_p(\alpha)} \int_{Y=Y'>0} |B + T^*|^{-\alpha} |Y|^{\alpha - \frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(Y)} \mathrm{d}Y \\ &= |B|^\alpha |B + T^*|^{-\alpha} = |I + B^{-1} T^*|^{-\alpha}. \end{aligned} \tag{7.5.4}$$

Thus for known B and arbitrary T^*, (7.5.4) will uniquely determine (7.5.3) through the uniqueness of the inverse Laplace transform. The conditions for the uniqueness will not be discussed here. For some results in this direction see Mathai (1993, 1997) and the references therein.

7.5.1. A convolution property for Laplace transforms

Let $f_1(X)$ and $f_2(X)$ be two real scalar functions of the real symmetric positive definite matrix X and let $g_1(T^*)$ and $g_2(T^*)$ be their Laplace transforms. Let

$$f_3(X) = \int_{0<S=S'<X} f_1(X - S) f_2(S) \mathrm{d}S. \tag{7.5.5}$$

Then $g_1 g_2$ is the Laplace transform of $f_3(X)$.

This result can be established from the definition itself.

$$L_{f_3}(T^*) = \int_{X=X'>0} \mathrm{e}^{-\mathrm{tr}(T^* X)} f_3(X) \mathrm{d}X$$

$$= \int_{x>0} \int_{S<X} \mathrm{e}^{-\mathrm{tr}(T^* X)} f_1(X - S) f_2(S) \mathrm{d}S \wedge \mathrm{d}X.$$

Note that $\{S < X, X > 0\}$ is also equivalent to $\{X > S, S > 0\}$. Hence we may interchange the integrals. Then

$$L_{f_3}(T^*) = \int_{S>0} f_2(S)\left[\int_{X>S} \mathrm{e}^{-\mathrm{tr}(T^*X)} f_1(X-S)\mathrm{d}X\right] \wedge \mathrm{d}S.$$

Put $X - S = Y \Rightarrow X = Y + S$ and then

$$\begin{aligned} L_{f_3}(T^*) &= \int_{S>0} \mathrm{e}^{-\mathrm{tr}(T^*S)} f_2(S)\left[\int_{Y>0} \mathrm{e}^{-\mathrm{tr}(T^*Y)} f_1(Y)\mathrm{d}Y\right] \wedge \mathrm{d}S \\ &= g_2(T^*)g_1(T^*). \end{aligned}$$

Example 7.5.2. Using the convolution property for the Laplace transform and an integral representation for the real matrix-variate beta function show that

$$B_p(\alpha,\beta) = \Gamma_p(\alpha)\Gamma_p(\beta)/\Gamma_p(\alpha+\beta).$$

Solution 7.5.2. Let us start with the integral representation

$$\begin{aligned} B_p(\alpha,\beta) &= \int_{0<X<I} |X|^{\alpha-\frac{p+1}{2}}|I-X|^{\beta-\frac{p+1}{2}}\mathrm{d}X, \\ &\qquad \Re(\alpha) > \frac{p-1}{2}, \Re(\beta) > \frac{p-1}{2}. \end{aligned}$$

Consider the integral

$$\begin{aligned} \int_{0<U<X} |U|^{\alpha-\frac{p+1}{2}}|X-U|^{\beta-\frac{p+1}{2}}\mathrm{d}U &= |X|^{\beta-\frac{p+1}{2}} \int_{0<U<X} |U|^{\alpha-\frac{p+1}{2}} \\ &\quad \times |I - X^{-\frac{1}{2}}UX^{-\frac{1}{2}}|^{\beta-\frac{p+1}{2}}\mathrm{d}U \\ &= |X|^{\alpha+\beta-\frac{p+1}{2}} \int_{0<Y<I} |Y|^{\alpha-\frac{p+1}{2}}|I-Y|^{\beta-\frac{p+1}{2}}\mathrm{d}Y, Y = X^{-\frac{1}{2}}UX^{-\frac{1}{2}}. \end{aligned}$$

Then

$$B_p(\alpha,\beta)|X|^{\alpha+\beta-\frac{p+1}{2}} = \int_{0<U<X} |U|^{\alpha-\frac{p+1}{2}}|X-U|^{\beta-\frac{p+1}{2}}\mathrm{d}U. \tag{7.5.6}$$

Take the Laplace transform on both sides to obtain the following:
On the left side,

$$B_p(\alpha,\beta)\int_{X>0}|X|^{\alpha+\beta-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(T^*X)}\mathrm{d}X = B_p(\alpha,\beta)|T^*|^{-(\alpha+\beta)}\Gamma_p(\alpha+\beta).$$

On the right side we get,

$$\int_{X>0}\mathrm{e}^{-\mathrm{tr}(T^*X)}\left[\int_{0<U<X}|U|^{\alpha-\frac{p+1}{2}}|X-U|^{\beta-\frac{p+1}{2}}\mathrm{d}U\right]\mathrm{d}X$$
$$= \Gamma_p(\alpha)\Gamma_p(\beta)|T^*|^{-(\alpha+\beta)} \text{ (by the convolution property in (7.5.5).)}$$

Hence

$$B_p(\alpha,\beta) = \Gamma_p(\alpha)\Gamma_p(\beta)/\Gamma_p(\alpha+\beta).$$

Example 7.5.3. Let $h(T^*)$ be the Laplace transform of $f(X)$, that is, $h(T^*) = L_f(T^*)$. Then show that the Laplace transform of $|X|^{-\frac{p+1}{2}}\Gamma_p(\frac{p+1}{2})f(X)$ is equivalent to $\int_{U>T^*}h(U)\mathrm{d}U$.

Solution: From (7.5.3) observe that for symmetric positive definite constant matrix B the following is an identity.

$$|B|^{-\alpha} = \frac{1}{\Gamma_p(\alpha)}\int_{X>0}|X|^{\alpha-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(BX)}\mathrm{d}X,\ \Re(\alpha) > \frac{p-1}{2}. \tag{7.5.7}$$

Then we can replace $|X|^{-\frac{p+1}{2}}\Gamma_p(\frac{p+1}{2})$ by an equivalent integral.

$$|X|^{-\frac{p+1}{2}}\Gamma_p\left(\frac{p+1}{2}\right) \equiv \int_{Y>0}|Y|^{\frac{p+1}{2}-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(XY)}\mathrm{d}Y = \int_{Y>0}\mathrm{e}^{-\mathrm{tr}(XY)}\mathrm{d}Y.$$

Then the Laplace transform of $|X|^{-\frac{p+1}{2}}\Gamma_p\left(\frac{p+1}{2}\right)f(X)$ is given by,

$$\int_{X>0}\mathrm{e}^{-\mathrm{tr}(T^*X)}f(X)[\int_{Y>0}\mathrm{e}^{-\mathrm{tr}(YX)}\mathrm{d}Y]\wedge\mathrm{d}X$$
$$= \int_{X>0}\int_{Y>0}\mathrm{e}^{-\mathrm{tr}[(T^*+Y)X]}f(X)\mathrm{d}Y\wedge\mathrm{d}X. \text{ (Put } T^*+Y=U \Rightarrow U>T^*)$$
$$= \int_{Y>0}h(T^*+Y)\mathrm{d}Y = \int_{U>T^*}h(U)\mathrm{d}U.$$

Example 7.5.4. For $X > B, B = B' > 0$ and $\nu > -1$ show that the Laplace transform of $|X-B|^\nu$ is $|T|^{-(\nu+\frac{p+1}{2})}\mathrm{e}^{-\mathrm{tr}(T^*B)}\Gamma_p(\nu+\frac{p+1}{2})$.

Solution 7.5.3. Laplace transform of $|X-B|^{\nu}$ with parameter matrix T^* is given by,

$$\int_{X>B}|X-B|^{\nu}\mathrm{e}^{-\mathrm{tr}(T^*X)}\mathrm{d}X = \mathrm{e}^{-\mathrm{tr}(BT^*)}\int_{Y>0}|Y|^{\nu}\mathrm{e}^{-\mathrm{tr}(T^*Y)}\mathrm{d}Y,\ Y=X-B$$
$$= \mathrm{e}^{-\mathrm{tr}(BT^*)}\ \Gamma_p\left(\nu+\frac{p+1}{2}\right)|T^*|^{-(\nu+\frac{p+1}{2})}$$

(by writing $\nu=\nu+\frac{p+1}{2}-\frac{p+1}{2}$) for $\nu+\frac{p+1}{2}>\frac{p-1}{2}\Rightarrow\nu>-1$.

Exercises 7.5.

7.5.1. By using the process in Example 7.5.3, or otherwise, show that the Laplace transform of $[\Gamma_p(\frac{p+1}{2})|X|^{-\frac{p+1}{2}}]^n f(X)$ can be written as

$$\int_{W_1>T^*}\int_{W_2>W_1}\dots\int_{W_n>W_{n-1}} h(W_n)\mathrm{d}W_1\wedge\cdots\wedge\mathrm{d}W_n$$

where $h(T^*)$ is the Laplace transform of $f(X)$.

7.5.2. Show that the Laplace transform of $|X|^n$ is $|T^*|^{-n-\frac{p+1}{2}}\Gamma_p(n+\frac{p+1}{2})$ for $n>-1$.

7.5.3. If the $p\times p$ real matrix random variable X has a type-1 beta density with parameters (α_1,α_2) then show that

(i) $\mathrm{U}=(I-X)^{-\frac{1}{2}}X(I-X)^{-\frac{1}{2}}\sim$ type-2 beta (α_1,α_2)

(ii) $\mathrm{V}=X^{-1}-I\sim$ type-2 beta (α_2,α_1)

where " $\sim$ " indicates "distributed as", and the parameters are given in the brackets.

7.5.4. If the $p\times p$ real symmetric positive definite matrix random variable X has a type-2 beta density with parameters α_1 and α_2 then show that

(i) $U=X^{-1}\sim$ type-2 beta (α_2,α_1)

(ii) $V=(I+X)^{-1}\sim$ type-1 beta (α_2,α_1)

(iii) $W=(I+X)^{-\frac{1}{2}}X(I+X)^{-\frac{1}{2}}\sim$ type-1 beta (α_1,α_2).

7.5.5. If the Laplace transform of $f(X)$ is $g(T^*)=L_{T^*}(f(X))$, where X is real symmetric positive definite and $p\times p$ then show that

$$\Delta^n g(T^*) = L_{T^*}(|X|^n f(X)), \quad \Delta = (-1)^p \left|\frac{\partial}{\partial T^*}\right|$$

where $|\frac{\partial}{\partial T^*}|$ means that first the partial derivatives with respect to t_{ij}'s for all i and j are taken, then written in the matrix form and then the determinant is taken, where $T^* = (t^*_{ij})$.

7.6. Hypergeometric Functions with Matrix Argument

There are essentially three approaches available in the literature for defining a hypergeometric function of matrix argument. One approach due to Bochner (1952) and Herz (1955) is through Laplace and inverse Laplace transforms. Under this approach, a hypergeometric function is defined as the function satisfying a pair of integral equations, and explicit forms are available for ${}_0F_0$ and ${}_1F_0$. Another approach is available from James (1961) and Constantine (1963) through a series form involving zonal polynomials. Theoretically, explicit forms are available for general parameters or for a general ${}_pF_q$ but due to the difficulty in computing higher order zonal polynomials, computations are feasible only for small values of p and q. For a detailed discussion of zonal polynomials see Mathai, Provost and Hayakawa (1995). The third approach is due to Mathai (1978, 1993) with the help of a generalized matrix transform or M-transform. Through this definition a hypergeometric function is defined as a class of functions satisfying a certain integral equation. This definition is the one most suited for studying various properties of hypergeometric functions. The series form is least suited for this purpose. All these definitions are introduced for symmetric functions in the sense that

$$f(X) = f(X') = f(QQ'X) = f(Q'XQ) = f(D),\ D = \text{diag}(\lambda_1, ..., \lambda_p).$$

If X is $p \times p$ and real symmetric then there exists an orthonormal matrix Q, that is, $QQ' = I, Q'Q = I$ such that $Q'XQ = \text{diag}(\lambda_1, ..., \lambda_p)$ where $\lambda_1, ..., \lambda_p$ are the eigenvalues of X. Thus, $f(X)$, a scalar function of the $p(p+1)/2$ functionally independent elements in X, is essentially a function of the p variables $\lambda_1, ..., \lambda_p$ when the function $f(X)$ is symmetric in the above sense.

7.6.1. Hypergeometric function through Laplace transform

Let ${}_rF_s(a_1, ..., a_r; b_1, ..., b_s; Z)$ be the hypergeometric function of the matrix argument Z to be defined, $Z = Z'$. Consider the following pair of Laplace and inverse Laplace transforms.

$$
\begin{aligned}
&{}_{r+1}F_s(a_1, ..., a_r, c; b_1, ..., b_s; -\wedge^{-1})|\wedge|^{-c} \\
&\quad = \frac{1}{\Gamma_p(c)} \int_{U=U'>0} \mathrm{e}^{-\mathrm{tr}(\wedge U)} {}_rF_s(a_1, ..., a_r; b_1, ..., b_s; -U)|U|^{c-\frac{p+1}{2}} \mathrm{d}U
\end{aligned} \tag{7.6.1}
$$

and

$$
\begin{aligned}
&{}_rF_{s+1}(a_1, ..., a_r; b_1, ..., b_r, c; -\wedge)|\wedge|^{c-\frac{p+1}{2}} \\
&\quad = \frac{\Gamma_p(c)}{(2\pi i)^{p(p+1)/2}} \int_{\Re(Z)=X>X_0} \mathrm{e}^{\mathrm{tr}(\wedge Z)} {}_rF_s(a_1, ..., a_r; b_1, ..., b_s; -Z^{-1})|Z|^{-c} \mathrm{d}Z
\end{aligned} \tag{7.6.2}
$$

where $Z = X + iY$, $i = \sqrt{-1}$, $X = X' > 0$, and X and Y belong to the class of symmetric matrices with the non-diagonal elements weighted by $\frac{1}{2}$. The function ${}_rF_s$ satisfying (7.6.1) and (7.6.2) can be shown to be unique under certain conditions and that function is defined as the hypergeometric function of matrix argument $\wedge$, according to this definition.

Then by taking ${}_0F_0(; ; -\wedge) = \mathrm{e}^{-\mathrm{tr}(\wedge)}$ and by using the convolution property of the Laplace transform and equations (7.6.1) and (7.6.2) one can systematically build up. The Bessel function ${}_0F_1$ for matrix argument is defined by Herz (1955). Thus we can go from ${}_0F_0$ to ${}_1F_0$ to ${}_0F_1$ to ${}_1F_1$ to ${}_2F_1$ and so on to a general ${}_pF_q$.

Example 7.6.1. Obtain an explicit form for ${}_1F_0$ from the above defi nition by using ${}_0F_0$ $(; ; -U) = \mathrm{e}^{-\mathrm{tr}(U)}$.

Solution 7.6.1. From (7.6.1)

$$
\begin{aligned}
&\frac{1}{\Gamma_p(c)} \int_{U=U'>0} |U|^{c-\frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(\wedge U)} {}_0F_0(; ; -U) \mathrm{d}U \\
&\quad = \frac{1}{\Gamma_p(c)} \int_{U>0} |U|^{c-\frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}[(I+\wedge)U]} \mathrm{d}U = |I + \wedge|^{-c},
\end{aligned}
$$

since

$$
{}_0F_0(; ; -U) = \mathrm{e}^{-\mathrm{tr}(U)}.
$$

But

$$|I + \wedge|^{-c} = |\wedge|^{-c}|I + \wedge^{-1}|^{-c}.$$

Then from (7.6.1)

$${}_1F_0(c; ; -\wedge^{-1}) = |I + \wedge^{-1}|^{-c}$$

which is an explicit representation.

7.6.2. Hypergeometric function through zonal polynomials

Zonal polynomials are certain symmetric functions in the eigenvalues of the $p \times p$ matrix Z. They are denoted by $C_K(Z)$ where K represents the partition of the positive integer k, $K = (k_1, ..., k_p)$ with $k_1 + \cdots + k_p = k$. When Z is 1×1 then $C_K(z) = z^k$. Thus, $C_K(Z)$ can be looked upon as a generalization of z^k in the scalar case. For details see Mathai, Provost and Hayakawa (1995). In terms of $C_K(Z)$ we have the representation for a

$${}_0F_0(; ; Z) = \mathrm{e}^{\mathrm{tr}(Z)} = \sum_{k=0}^{\infty} \frac{(\mathrm{tr}(Z))^k}{k!} = \sum_{k=0}^{\infty} \sum_{K} \frac{C_K(Z)}{k!}. \tag{7.6.3}$$

The binomial expansion will be the following:

$${}_1F_0(\alpha; ; Z) = \sum_{k=0}^{\infty} \sum_{K} \frac{(\alpha)_K C_K(Z)}{k!} = |I - Z|^{-\alpha}, \tag{7.6.4}$$

for $0 < Z < I$, where,

$$(\alpha)_K = \prod_{j=1}^{p} \left(\alpha - \frac{j-1}{2}\right)_{k_j}, K = (k_1, ..., k_p), k_1 + \cdots + k_p = k. \tag{7.6.5}$$

In terms of zonal polynomials a hypergeometric series is defined as follows:

$${}_pF_q(a_1, ..., a_p; b_1, ..., b_q; Z) = \sum_{k=0}^{\infty} \sum_{K} \frac{(a_1)_K \cdots (a_p)_K}{(b_1)_K \cdots (b_q)_K} \frac{C_K(Z)}{k!}. \tag{7.6.6}$$

For (7.6.6) to be defined, none of the denominator factors is equal to zero, $q \geq p$, or $q = p + 1$ and $0 < Z < I$. For other details see Constantine (1963). In order to study properties of a hypergeometric function with the help of (7.6.6) one needs the Laplace and inverse Laplace transforms of zonal polynomials. These are the following:

$$\int_{X=X'>0} |X|^{\alpha-\frac{p+1}{2}} e^{-\mathrm{tr}(XZ)} C_K(XT)\mathrm{d}X = |Z|^{-\alpha} C_K(TZ^{-1})\Gamma_p(\alpha, K) \qquad (7.6.7)$$

where

$$\Gamma_p(\alpha, K) = \pi^{p(p-1)/4} \prod_{j=1}^{p} \Gamma\left[\alpha + k_j - \frac{j-1}{2}\right] = \Gamma_p(\alpha)(\alpha)_K. \qquad (7.6.8)$$

$$\frac{1}{(2\pi i)^{p(p+1)/2}} \int_{\Re(Z)=X>X_0} e^{\mathrm{tr}(SZ)} |Z|^{-\alpha} C_K(Z)\mathrm{d}Z = \frac{1}{\Gamma_p(\alpha, K)} |S|^{\alpha-\frac{p+1}{2}} C_K(S), i = \sqrt{-1} \qquad (7.6.9)$$

for $Z = X + iY,\ X = X' > 0,\ X$ and Y are symmetric and the nondiagonal elements are weighted by $\frac{1}{2}$. If the non-diagonal elements are not weighted then the left side in (7.6.9) is to be multiplied by $2^{p(p-1)/2}$. Further,

$$\int_{0<X<I} |X|^{\alpha-\frac{p+1}{2}} |I - X|^{\beta-\frac{p+1}{2}} C_K(TX)\mathrm{d}X = \frac{\Gamma_p(\alpha, K)\Gamma_p(\beta)}{\Gamma_p(\alpha+\beta, K)} C_K(T)$$
$$\Re(\alpha) > \frac{p-1}{2},\ \Re(\beta) > \frac{p-1}{2}. \qquad (7.6.10)$$

Example 7.6.2. By using zonal polynomials establish the following results:

$$\begin{aligned} {}_2F_1(a, b; c; X) &= \frac{\Gamma_p(c)}{\Gamma_p(a)\Gamma_p(c-a)} \\ &\times \int_{0<\wedge<I} |\wedge|^{a-\frac{p+1}{2}} |I - \wedge|^{c-a-\frac{p+1}{2}} |I - \wedge X|^{-b} \mathrm{d}\wedge \qquad (7.6.11)\end{aligned}$$

for $\Re(a) > \frac{p-1}{2},\ \Re(c-a) > \frac{p-1}{2}$.

Solution 7.6.2. Expanding $|I - \wedge X|^{-b}$ in terms of zonal polynomials and then integrating term by term the right side reduces to the following:

$$|I - \wedge X|^{-b} = \sum_{k=0}^{\infty} \sum_{K} (b)_K \frac{C_K(\wedge X)}{k!} \quad \text{for } 0 < \wedge X < I$$

and

$$\int_{0<\wedge<I} |\wedge|^{a-\frac{p+1}{2}}|I-\wedge|^{c-a-\frac{p+1}{2}}C_K(\wedge X)\mathrm{d}\wedge = \frac{\Gamma_p(a,K)\Gamma_p(c-a)}{\Gamma_p(c,K)}C_K(X)$$

by using (7.6.10). But

$$\frac{\Gamma_p(a,K)\Gamma_p(c-a)}{\Gamma_p(c,K)} = \frac{\Gamma_p(a)\Gamma_p(c-a)}{\Gamma_p(c)}\frac{(a)_K}{(c)_K}.$$

Substituting these back, the right side becomes

$$\sum_{k=0}^{\infty}\sum_K \frac{(a)_K(b)_K}{(c)_K}\frac{C_K(X)}{k!} = {}_2F_1(a,b;c;X).$$

This establishes the result.

Example 7.6.3. Establish the result

$$ {}_2F_1(a,b;c;I) = \frac{\Gamma_p(c)\Gamma_p(c-a-b)}{\Gamma_p(c-a)\Gamma_p(c-b)} \tag{7.6.12}$$

for $\Re(c-a-b) > \frac{p-1}{2}$, $\Re(c-a) > \frac{p-1}{2}$, $\Re(c-b) > \frac{p-1}{2}$.

Solution 7.6.3. In (7.6.11) put $X = I$, combine the last factor on the right with the previous factor and integrate out with the help of a matrix-variate type-1 beta integral.

Uniqueness of the ${}_pF_q$ through zonal polynomials, as given in (7.6.6), is established by appealing to the uniqueness of the function defined through the Laplace and inverse Laplace transform pair in (7.6.1) and (7.6.2), and by showing that (7.6.6) satisfies (7.6.1) and (7.6.2).

The next definition, introduced by Mathai in a series of papers is through a special case of Weyl's fractional integral.

7.6.3. Hypergeometric functions through M-transforms

Consider the class of $p \times p$ real symmetric definite matrices and the null matrix O. Any member of this class will be either positive definite or negative definite or null. Let α be a complex parameter such that $\Re(\alpha) > \frac{p-1}{2}$. Let $f(S)$ be a scalar symmetric function in the sense $f(AB) = f(BA)$ for all A and B when AB and BA are defined. Then the M-transform of $f(S)$, denoted by $M_\alpha(f)$, is defined as

$$M_\alpha(f) = \int_{U=U'>0} |U|^{\alpha-\frac{p+1}{2}} f(U)\mathrm{d}U. \tag{7.6.13}$$

Some examples of symmetric functions are $\mathrm{e}^{\pm\mathrm{tr}(S)}$, $|I \pm S|^\beta$ for nonsingular $p \times p$ matrices A and B such that,

$$\mathrm{e}^{\pm\mathrm{tr}(AB)} = \mathrm{e}^{\pm\mathrm{tr}(BA)};\ |I \pm AB|^\beta = |I \pm BA|^\beta.$$

Is it possible to recover $f(U)$, a function of $p(p+1)/2$ elements in $U = (u_{ij})$ or a function of p eigenvalues of U, that is a function of p variables, from $M_\alpha(t)$ which is a function of one parameter α? In a normal course the answer is in the negative. But due to the properties that are seen, it is clear that there exists a set of sufficient conditions by which $M_\alpha(f)$ will uniquely determine $f(U)$. It is easy to note that the class of functions defined through (7.6.13) satisfy the pair of integral equations (7.6.1) and (7.6.2) defining the unique hypergeometric function.

A hypergeometric function through M-transform is defined as a class of functions ${}_rF_s^*$ satisfying the following equation:

$$\int_{X=X'>0} |X|^{\alpha-\frac{p+1}{2}} {}_rF_s{}^*(a_1, ..., a_p; b_1, ..., b_q; -X)\mathrm{d}X$$
$$= \frac{\{\prod_{j=1}^s \Gamma_p(b_j)\}}{\{\prod_{j=1}^r \Gamma_p(a_j)\}} \frac{\{\prod_{j=1}^r \Gamma_p(a_j - \rho)\}}{\{\prod_{j=1}^s \Gamma_p(b_j - \rho)\}} \Gamma_p(\rho) \tag{7.6.14}$$

where ρ is an arbitrary parameter such that the gammas exist.

Example 7.6.4. Re-establish the result

$$L_T(|X - B|^\nu) = \Gamma_p\left(\nu + \frac{p+1}{2}\right) |T|^{-(\nu+\frac{p+1}{2})} \mathrm{e}^{-\mathrm{tr}(TB)} \tag{7.6.15}$$

by using M-transforms.

Solution 7.6.4. We will show that the M-transforms on both sides of (7.6.15) are one and the same. Taking the M-transform of the left-side, with respect to the parameter ρ, we have,

$$\int_{T>0}|T|^{\rho-\frac{p+1}{2}}\{L_T(|X-B|^{\nu})\}\mathrm{d}T=\int_{T>0}|T|^{\rho-\frac{p+1}{2}}\left[\int_{X>B}|X-B|^{\nu}\mathrm{e}^{-\mathrm{tr}(TX)}\mathrm{d}X\right]\mathrm{d}T$$
$$=\int_{T>0}|T|^{\rho-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(TB)}\left[\int_{Y>0}|Y|^{\nu}\mathrm{e}^{-\mathrm{tr}(TY)}\mathrm{d}Y\right]\mathrm{d}T.$$

Noting that $\nu=\nu+\frac{p+1}{2}-\frac{p+1}{2}$ the Y-integral gives $|T|^{-\nu-\frac{p+1}{2}}\Gamma_p(\nu+\frac{p+1}{2})$. Then the T-integral gives

$$M_\rho(\text{left-side})=\Gamma_p\left(\nu+\frac{p+1}{2}\right)\Gamma_p\left(\rho-\nu-\frac{p+1}{2}\right)|B|^{-\rho+\nu+\frac{p+1}{2}}.$$

M-transform of the right side gives,

$$M_\rho(\text{right-side})=\int_{T>0}|T|^{\rho-\frac{p+1}{2}}\left\{\Gamma_p\left(\nu+\frac{p+1}{2}\right)|T|^{-(\nu+\frac{p+1}{2})}\mathrm{e}^{-\mathrm{tr}(TB)}\right\}\mathrm{d}T$$
$$=\Gamma_p\left(\nu+\frac{p+1}{2}\right)\Gamma_p\left(\rho-\nu-\frac{p+1}{2}\right)|B|^{-\rho+\nu+\frac{p+1}{2}}.$$

The two sides have the same M-transform.

Starting with ${}_0F_0(;;X)=\mathrm{e}^{\mathrm{tr}(X)}$, we can build up a general ${}_pF_q$ by using the M-transform and the convolution form for M-transforms, which will be stated next.

7.6.4. A convolution theorem for M-transforms

Let $f_1(U)$ and $f_2(U)$ be two symmetric scalar functions of the $p\times p$ real symmetric positive definite matrix U, with M-transforms $M_\rho(f_1)=g_1(\rho)$ and $M_\rho(f_2)=g_2(\rho)$ respectively. Let

$$f_3(S)=\int_{U>0}|U|^{\beta}f_1(U^{\frac{1}{2}}SU^{\frac{1}{2}})f_2(U)\mathrm{d}U \tag{7.6.16}$$

then the M-transform of f_3 is given by,

$$M_\rho(f_3)=g_1(\rho)g_2\left(\beta-\rho+\frac{p+1}{2}\right). \tag{7.6.17}$$

The result can be easily established from the definition itself by interchanging the integrals.

Example 7.6.5. Show that

$$ {}_1F_1(a;c;-\wedge) = \frac{\Gamma_p(c)}{\Gamma_p(a)\Gamma_p(c-a)} \int_{0<U<I} |U|^{a-\frac{p+1}{2}} |I-U|^{c-a-\frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(\wedge U)} \mathrm{d}U. \qquad (7.6.18) $$

Solution 7.6.5. We will establish this by showing that both sides have the same M-transforms. From the definition in (7.6.14) the M-transform of the left side with respect to the parameter ρ is given by the following:

$$ \begin{aligned} M_\rho(\text{left-side}) &= \int_{\wedge=\wedge'>0} |\wedge|^{\rho-\frac{p+1}{2}} \; {}_1F_1(a;c;-\wedge)\mathrm{d}\wedge \\ &= \left[\frac{\Gamma_p(a-\rho)}{\Gamma_p(c-\rho)}\Gamma_p(\rho)\right] \frac{\Gamma_p(c)}{\Gamma_p(a)}. \\ M_\rho(\text{right-side}) &= \int_{\wedge>0} |\wedge|^{\rho-\frac{p+1}{2}} \left\{ \frac{\Gamma_p(c)}{\Gamma_p(a)\Gamma_p(c-a)} \right. \\ &\quad \left. \times \int_{0<U<I} |U|^{a-\frac{p+1}{2}} |I-U|^{c-a-\frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(\wedge U)} \mathrm{d}U \right\} \mathrm{d}\wedge\,. \end{aligned} $$

Take,

$$ f_1(U) = \mathrm{e}^{-\mathrm{tr}(U)} \text{ and } f_2(U) = |U|^{a-\frac{p+1}{2}} |I-U|^{c-a-\frac{p+1}{2}}. $$

Then

$$ \begin{aligned} M_\rho(f_1) = g_1(\rho) &= \int_{U>0} |U|^{\rho-\frac{p+1}{2}} \mathrm{e}^{-\mathrm{tr}(U)} \mathrm{d}U = \Gamma_p(\rho), \Re(\rho) > \frac{p-1}{2}. \\ M_\rho(f_2) = g_2(\rho) &= \int_{U>0} |U|^{\rho-\frac{p+1}{2}} |U|^{a-\frac{p+1}{2}} |I-U|^{c-a-\frac{p+1}{2}} \mathrm{d}U \\ &= \frac{\Gamma_p(a+\rho-\frac{p+1}{2})\Gamma_p(c-a)}{\Gamma_p(c+\rho-\frac{p+1}{2})}, \Re(c-a) > \frac{p-1}{2}, \\ &\quad \Re(a+\rho) > p, \Re(c+\rho) > p. \end{aligned} $$

Taking f_3 in (7.6.16) as the second integral on the right above we have

$$ M_\rho(\text{right-side}) = \left\{\frac{\Gamma_p(c)}{\Gamma_p(a)}\right\} \Gamma_p(\rho) \frac{\Gamma_p(a-\rho)}{\Gamma_p(c-\rho)} = M_\rho(\text{left-side}). $$

Hence the result.

Almost all properties, analogous to the ones in the scalar case for hypergeometric functions, can be established by using the M-transform technique very easily. These can then be shown to be unique , if necessary, through the uniqueness of Laplace and inverse Laplace transform pair. Theories for functions of several matrix arguments, Dirichlet integrals, Dirichlet densities, their extensions, Appell's functions, Lauricella functions, and the like, are available. Then all these real cases are also extended to complex cases as well. For details see Mathai (1997). Problems involving scalar functions of matrix argument, real and complex cases, are still being worked out and applied in many areas such as statistical distribution theory, econometrics, quantum mechanics and engineering areas. Since the aim in this brief note is only to introduce the subject matter, more details will not be given here.

Exercises 7.6.

7.6.1. Show that for $\wedge = \wedge' > 0$ and $p \times p$,

$${}_1F_1(a;c;-\wedge) = \mathrm{e}^{-\mathrm{tr}(\wedge)}{}_1F_1(c-a;c;\wedge).$$

7.6.2. For $p \times p$ real symmetric positive definite matrices $\wedge$ and $\vee$ show that

$$\begin{aligned} {}_1F_1(a;c;-\wedge) &= \frac{\Gamma_p(c)}{\Gamma_p(a)\Gamma_p(c-a)}|\wedge|^{-(c-\frac{p+1}{2})}\int_{0<\vee<\wedge}\mathrm{e}^{-\mathrm{tr}(\vee)} \\ &\quad \times|\vee|^{a-\frac{p+1}{2}}|\wedge-\vee|^{c-a-\frac{p+1}{2}}\mathrm{d}V. \end{aligned}$$

7.6.3. Show that for ϵ a scalar and A a $p \times p$ matrix with p finite

$$\lim_{\epsilon\to 0}|I+\epsilon A|^{-\frac{1}{\epsilon}} = \lim_{\epsilon\to\infty}|I+\frac{A}{\epsilon}|^{-\epsilon} = \mathrm{e}^{-\mathrm{tr}(A)}.$$

7.6.4. Show that

$$\begin{aligned} \lim_{a\to\infty}{}_1F_1(a;c;-\frac{Z}{a}) &= \lim_{\epsilon\to 0}{}_1F_1\left(\frac{1}{\epsilon};c;-\epsilon Z\right) \\ &= {}_0F_1(;c;-Z). \end{aligned}$$

7.6.5. Show that

$${}_1F_1(a;c;-\wedge) = \frac{\Gamma_p(c)}{(2\pi i)^{\frac{p(p+1)}{2}}}\int_{\Re(Z)=X>X_0}\mathrm{e}^{\mathrm{tr}(Z)}|Z|^{-c}|I+\wedge Z^{-1}|^{-a}\mathrm{d}Z.$$

7.6.6. Show that

$$ {}_2F_1(a,b;c;X) = |I-X|^{-\beta}{}_2F_1(c-a,b;c;-X(I-X)^{-1}). $$

7.6.7. For $\Re(s) > \frac{p-1}{2}, \Re(b-s) > \frac{p-1}{2}, \Re(c-a-s) > \frac{p-1}{2}$, show that

$$ \int_{0<X<I} |X|^{s-\frac{p+1}{2}}|I-X|^{b-s-\frac{p+1}{2}}{}_2F_1(a,b;c;X)\mathrm{d}X $$

$$ = \frac{\Gamma_p(c)\Gamma_p(s)\Gamma_p(b-s)\Gamma_p(c-a-s)}{\Gamma_p(b)\Gamma_p(c-a)\Gamma_p(c-s)}. $$

7.6.8. Defining the Bessel function $A_r(S)$ with $p \times p$ real symmetric positive definite matrix argument S, as

$$ A_r(S) = \frac{1}{\Gamma_p(r+\frac{p+1}{2})}{}_0F_1(;r+\frac{p+1}{2};-S), \tag{7.6.19} $$

show that

$$ \int_{S>0} |S|^{\delta-\frac{p+1}{2}}A_r(S)\mathrm{e}^{-\mathrm{tr}(\wedge S)}\mathrm{d}S = \frac{\Gamma_p(\delta)}{\Gamma_p(r+\frac{p+1}{2})}|\wedge|^{-\delta}{}_1F_1\left(\delta;r+\frac{p+1}{2};-\wedge^{-1}\right). $$

7.6.9. If

$$ M(\alpha,\beta;A) = \int_{X=X'>0} |X|^{\alpha-\frac{p+1}{2}}|I+X|^{\beta-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(AX)}\mathrm{d}X, $$

$$ \Re(\alpha) > \frac{p-1}{2}, A = A^{'} > 0 $$

then show that

$$ \int_{X>0} |X+A|^{\nu}\mathrm{e}^{-\mathrm{tr}(TX)}\mathrm{d}X = |A|^{\nu+\frac{p+1}{2}}M\left(\frac{p+1}{2},\nu+\frac{p+1}{2};A^{\frac{1}{2}}TA^{\frac{1}{2}}\right). $$

7.6.10. If Whittaker function W is defined as

$$ \int_{Z>0} |Z|^{\mu-\frac{p+1}{2}}|I+Z|^{\nu-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(AZ)}\mathrm{d}Z $$

$$ = |A|^{-\frac{\mu+\nu}{2}}\Gamma_p(\mu)\mathrm{e}^{\frac{1}{2}\mathrm{tr}(A)}W_{\frac{1}{2}(\nu-\mu),\frac{1}{2}(\nu+\mu-\frac{(p+1)}{2})}(A) $$

then show that

$$\int_{X>U} |X+B|^{2\alpha-\frac{p+1}{2}}|X-U|^{2q-\frac{p+1}{2}}\mathrm{e}^{-\mathrm{tr}(MX)}\mathrm{d}X$$
$$= |U+B|^{\alpha+q-\frac{p+1}{2}}|M|^{-(\alpha+q)}\mathrm{e}^{\frac{1}{2}\mathrm{tr}[(B-U)M]}\Gamma_p(2q)$$
$$\times W_{(\alpha-q),(\alpha+q-\frac{(p+1)}{4})}(S), S=(U+B)^{\frac{1}{2}}M(U+B)^{\frac{1}{2}}.$$

7.7. A Pathway Model

As an application of a real scalar function of matrix argument, we will introduce a general real matrix-variate probability model, which covers almost all real matrix-variate densities used in multivariate statistical analysis. Through the density introudced here, a pathway is created to go from one functional form to another, to go from matrix-variate type-1 beta to matrix-variate type-2 beta to matrix-variate gamma to matrix-variate Gaussian densities.

7.7.1. The pathway density

Let $X=(x_{ij}), i=1,...,p, j=1,...,r, r\geq p$ be of rank p and of real scalar variables x_{ij}'s for all i and j, and having the density $f(X)$, where

$$f(X)=c|A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}|^{\alpha}|I-a(1-q)A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}|^{\frac{\beta}{1-q}} \tag{7.7.1}$$

for $A=A'>0$ and $p\times p, B=B'>0$ and $r\times r$ with $I-a(1-q)A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}>0$, A and B are free of the elements in X, a,β,q are scalar constants with $a>0,\beta>0$, and c is the normalizing constant. $A^{\frac{1}{2}}$ and $B^{\frac{1}{2}}$ denote the real positive definite square roots of A and B respectively.

For evaluating the normalizing constant c one can go through the following procedure: Let

$$Y=A^{\frac{1}{2}}XB^{\frac{1}{2}} \Rightarrow \mathrm{d}Y=A^{\frac{r}{2}}|B|^{\frac{p}{2}}\mathrm{d}X$$

by using Theorem 7.1.5. Let

$$U=YY' \Rightarrow \mathrm{d}Y=\frac{\pi^{\frac{rp}{2}}}{\Gamma_p\left(\frac{r}{2}\right)}|U|^{\frac{r}{2}-\frac{p+1}{2}}\mathrm{d}U$$

by using equation (7.3.17), where $\Gamma_P(\cdot)$ is the real matrix-variate gamma function. Let, for $q < 1$,

$$V = a(1-q)U \Rightarrow \mathrm{d}V = [a(1-q)]^{\frac{p(p+1)}{2}}\mathrm{d}U$$

from the same Theorem 7.1.5. If $f(X)$ is a density then the total integral is 1 and therefore,

$$1 = \int_X f(X)\mathrm{d}X = \frac{c}{|A|^{\frac{r}{2}}|B|^{\frac{p}{2}}}\int_Y |YY'|^{\alpha}|I - a(1-q)YY'|^{\frac{\beta}{1-q}}\mathrm{d}Y \tag{7.7.2}$$

$$= \frac{\pi^{\frac{rp}{2}}}{\Gamma_p\left(\frac{r}{2}\right)|A|^{\frac{r}{2}}|B|^{\frac{p}{2}}}\int_U |U|^{\alpha+\frac{r}{2}-\frac{p+1}{2}}|I - a(1-q)U|^{\frac{\beta}{1-q}}\mathrm{d}U. \tag{7.7.3}$$

Note 7.7.1. Note that from (7.7.2) and (7.7.3) we can also infer the densities of Y and U respectively.

At this stage we need to consider three cases.

Case (1): $q, < 1$. Then $a(1-q) > 0$. Make the transformation $V = a(1-q)U$, then

$$c^{-1} = \frac{\pi^{\frac{rp}{2}}}{\Gamma_p\left(\frac{r}{2}\right)|A|^{\frac{r}{2}}|B|^{\frac{p}{2}}[a(1-q)]^{p(\alpha+\frac{r}{2})}}\int_V |V|^{\alpha+\frac{r}{2}-\frac{p+1}{2}}|I - V|^{\frac{\beta}{1-q}}\mathrm{d}V. \tag{7.7.4}$$

The integral in (7.7.4) can be evaluated by using a real matrix-variate type-1 beta integral. Then we have

$$c^{-1} = \frac{\pi^{\frac{rp}{2}}}{\Gamma_p\left(\frac{r}{2}\right)|A|^{\frac{r}{2}}|B|^{\frac{p}{2}}[a(1-q)]^{p(\alpha+\frac{r}{2})}}\frac{\Gamma_p\left(\alpha+\frac{r}{2}\right)\Gamma_p\left(\frac{\beta}{1-q}+\frac{p+1}{2}\right)}{\Gamma_p\left(\alpha+\frac{r}{2}+\frac{\beta}{1-q}+\frac{p+1}{2}\right)} \tag{7.7.5}$$

for $\alpha + \frac{r}{2} > \frac{p-1}{2}$.

Note 7.7.2. In statistical problems usually the parameters are real and hence we will assume the parameters to be real here as well as in the discussions to follow. If α is in the complex domain then the condition will reduce to $\Re(\alpha) + \frac{r}{2} > \frac{p-1}{2}$.

Case (ii): $q > 1$.

In this case $1 - q = -(q-1)$ where $q - 1 > 0$. Then in (7.7.3) one factor in the integrand becomes

$$|I - a(1-q)U|^{\frac{\beta}{1-q}} = |I + a(q-1)U|^{-\frac{\beta}{q-1}} \tag{7.7.6}$$

and then making the substitution $V = a(q-1)U$ and then evaluating the integral by using a type-2 beta integral we have

$$c^{-1} = \frac{\pi^{\frac{rp}{2}}}{\Gamma_p\left(\frac{r}{2}\right)|A|^{\frac{r}{2}}|B|^{\frac{p}{2}}[a(q-1)]^{p(\alpha+\frac{r}{2})}}\int_V |V|^{\alpha+\frac{r}{2}-\frac{p+1}{2}}|I+V|^{-\frac{\beta}{q-1}}\mathrm{d}V.$$

Evalute the integral by using a real matrix-variate type-2 beta integral. We have the following:

$$c^{-1} = \frac{\pi^{\frac{rp}{2}}}{\Gamma_p\left(\frac{r}{2}\right)|A|^{\frac{r}{2}}|B|^{\frac{p}{2}}[a(q-1)]^{p(\alpha+\frac{r}{2})}}\frac{\Gamma_p\left(\alpha+\frac{r}{2}\right)\Gamma_p\left(\frac{\beta}{q-1}-\alpha-\frac{r}{2}\right)}{\Gamma_p\left(\frac{\beta}{q-1}\right)} \tag{7.7.7}$$

for $\alpha+\frac{r}{2} > \frac{p-1}{2}$, $\frac{\beta}{q-1}-\alpha-\frac{r}{2} > \frac{p-1}{2}$.

Case (iii): $q = 1$.

When q approaches 1 from the left or from the right it can be shown that the determinant containing q in (7.7.3) and (7.7.6) approaches an exponential form, which will be stated as a lemma:

Lemma 7.7.1.

$$\lim_{q\to 1}|I-a(1-q)U|^{\frac{\beta}{1-q}} = \mathrm{e}^{-a\beta\ \mathrm{tr}(U)}. \tag{7.7.8}$$

This lemma can be proved easily by observing that for any real symmetric matrix U there exists an orthonormal matrix Q such that $QQ' = I = Q'Q$, $Q'UQ = \mathrm{diag}(\lambda_1,...,\lambda_p)$ where λ_j's are the eigenvalues of U. Then

$$\begin{aligned}|I-a(1-q)U| &= |I-a(1-q)QQ'UQQ'|\\ &= |I-a(1-q)Q'UQ|\\ &= |I-a(1-q)\mathrm{diag}(\lambda_1,...,\lambda_p)|\\ &= \prod_{j=1}^{p}(1-a(1-q)\lambda_j).\end{aligned}$$

But

$$\lim_{q\to 1}(1-a(1-q)\lambda_j)^{\frac{\beta}{1-q}} = \mathrm{e}^{-a\beta\lambda_j}.$$

Then

$$\lim_{q\to 1}|I-a(1-q)U|^{\frac{\beta}{1-q}} = \mathrm{e}^{-a\beta\mathrm{tr}(U)}.$$

Hence in case (iii), for $q \to 1$, we have

$$
\begin{aligned}
c^{-1} &= \frac{\pi^{\frac{rp}{2}}}{\Gamma_p\left(\frac{r}{2}\right)|A|^{\frac{r}{2}}|B|^{\frac{p}{2}}} \int_U |U|^{\alpha+\frac{r}{2}-\frac{p+1}{2}} \mathrm{e}^{-a\beta \mathrm{tr}(U)} \mathrm{d}U \\
&= \frac{\pi^{\frac{rp}{2}}}{\Gamma_p\left(\frac{r}{2}\right)|A|^{\frac{r}{2}}|B|^{\frac{p}{2}}} \frac{\Gamma_p\left(\alpha+\frac{r}{2}\right)}{(a\beta)^{p(\alpha+\frac{r}{2})}},\ \alpha+\frac{r}{2} > \frac{p-1}{2} \qquad (7.7.9)
\end{aligned}
$$

by evaluating the integral with the help of a real matrix-variate gamma integral.

7.7.2. A general density

For X, A, B, a, β, q as defined in (7.7.1) the density $f(X)$ there has three different forms for three different situations of q. That is,

$$
\begin{aligned}
f(X) &= c_1|A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}|^{\alpha}|I - a(1-q)A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}|^{\frac{\beta}{1-q}}, \quad \text{for } q < 1 \qquad (7.7.10)\\
&= c_2|A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}|^{\alpha}|I + a(q-1)A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}|^{-\frac{\beta}{q-1}}, \quad \text{for } q > 1 \qquad (7.7.11)\\
&= c_3|A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}|^{\alpha} \exp\left\{-a\beta \mathrm{tr}[A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}]\right\}, \quad \text{for } q = 1 \qquad (7.7.12)
\end{aligned}
$$

where $c_1 = c$ for $q < 1, c_2 = c$ for $q > 1$ and $c_3 = c$ for $q = 1$, given in (7.7.5), (7.7.6) and (7.7.9) respectively.

Note 7.7.3. Observe that $f(X)$ maintains a generalized real matrix-variate type-1 beta form for $-\infty < q < 1$, $f(X)$ maintains a generalized real matrix-variate type-2 beta form for $1 < q < \infty$ and $f(X)$ keeps a generalized real matrix-variate gamma form when $q \to 1$.

Note 7.7.4. If a location parameter matrix is to be introduced then in $f(X)$, replace X by $X - M$ where M is a $p \times r$ constant matrix. All properties and derivations remain the same except that now X is located at M instead of at the origin O.

Remark 7.7.1. The parameter q in the density $f(X)$ can be taken as a pathway parameter. It defines a pathway from a generalized type-1 beta form to a type-2 beta form to a gamma form. Thus a wide variety of probability models are available from $f(X)$. If the experimenter needs a model with a thicker tail or thinner tail or the right and left tails cut off, all such models are available from $f(X)$ for various values of q. For $\alpha = 0$ one has the matrix-variate Gaussian form coming from $f(X)$.

7.7.3. Arbitray moments

Arbitrary moments of the determinant $|A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}|$ is available from the normalizing constant itself for various values of q. That is, denoting the expected values by E,

$$E|A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}|^h = \frac{1}{[a(1-q)]^{ph}}\frac{\Gamma_p\left(\alpha+h+\frac{r}{2}\right)}{\Gamma_p\left(\alpha+\frac{r}{2}\right)}\frac{\Gamma_p\left(\alpha+\frac{r}{2}+\frac{\beta}{1-q}+\frac{p+1}{2}\right)}{\Gamma_p\left(\alpha+h+\frac{r}{2}+\frac{\beta}{1-q}+\frac{p+1}{2}\right)} \quad (7.7.13)$$

$$\text{for } q<1,\ \alpha+h+\frac{r}{2}>\frac{p-1}{2}$$

$$= \frac{1}{[a(q-1)]^{ph}}\frac{\Gamma_p\left(\alpha+h+\frac{r}{2}\right)}{\Gamma_p\left(\alpha+\frac{r}{2}\right)}\frac{\Gamma_p\left(\frac{\beta}{q-1}-\alpha-h-\frac{r}{2}\right)}{\Gamma_p\left(\frac{\beta}{q-1}-\alpha-\frac{r}{2}\right)} \quad (7.7.14)$$

$$\text{for } q>1, \frac{\beta}{q-1}-\alpha-h-\frac{r}{2}>\frac{p-1}{2},\ \alpha+h+\frac{r}{2}>\frac{p-1}{2}$$

$$= \frac{1}{(a\beta)^{ph}}\frac{\Gamma_p\left(\alpha+h+\frac{r}{2}\right)}{\Gamma_p\left(\alpha+\frac{r}{2}\right)} \text{ for } q=1, \alpha+h+\frac{r}{2}>\frac{p-1}{2}. \quad (7.7.15)$$

7.7.4. Quadratic forms

The current theory in statistical leterature is based on a Gaussian or normal population and quadratic and bilinear forms in a simple random sample coming from such a normal population, or a quadratic form and bilinear forms in normal variables, independently distributed or jointly normally distributed. But from the structure of the density $f(X)$ in (7.7.1) it is evident that we can extend the theory to a much wider class. For $p=1, r>p$ the constant matrix A is a scalar quantity. For convenience let us take it as 1. Then we have

$$A^{\frac{1}{2}}XBX'A^{\frac{1}{2}} = (x_1,...,x_r)B\begin{bmatrix}x_1\\x_2\\\vdots\\x_r\end{bmatrix} = u(\text{ say }). \quad (7.7.16)$$

Here u is a real positive definite quadratic form in the first row of X, and this row is denoted by $(x_1,..,x_r)$. Now observe that the density of u is available as a special case in $f(X)$, from (7.7.10) for $q<1$, from (7.7.11) for $q>1$ and from (7.7.12) for $q=1$. [Write down the exact density in the three cases as an exercise].

7.7.5. Generalized quadratic form

For a general p, $U = A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}$ is the gneralized quadratic form in X where X has the density $f(X)$ in (7.7.1). The density of U is available from 7.7.3) in the following form, denoting it by $f_1(U)$. Then

$$f_1(U) = c^*|U|^{\alpha+\frac{r}{2}-\frac{p+1}{2}}|I-a(1-q)U|^{\frac{\beta}{1-q}} \tag{7.7.17}$$

where

$$c^* = \frac{[a(1-q)]^{p(\alpha+\frac{r}{2})}\Gamma_p\left(\alpha+\frac{r}{2}+\frac{\beta}{1-q}+\frac{p+1}{2}\right)}{\Gamma_p\left(\alpha+\frac{r}{2}\right)\Gamma_p\left(\frac{\beta}{1-q}+\frac{p+1}{2}\right)} \tag{7.7.18}$$

for $q < 1, \alpha + \frac{r}{2} > \frac{p-1}{2}$,

$$= \frac{[a(q-1)]^{p(\alpha+\frac{r}{2})}\Gamma_p\left(\frac{\beta}{q-1}\right)}{\Gamma_p\left(\alpha+\frac{r}{2}\right)\Gamma_p\left(\frac{\beta}{q-1}-\alpha-\frac{r}{2}\right)}, \tag{7.7.19}$$

for $q > 1, \alpha + \frac{r}{2} > \frac{p-1}{2}, \frac{\beta}{q-1} - \alpha - \frac{r}{2} > \frac{p-1}{2}$,

$$= \frac{(a\beta)^{p(\alpha+\frac{r}{2})}}{\Gamma_p\left(\alpha+\frac{r}{2}\right)}, \tag{7.7.20}$$

for $q = 1, \alpha + \frac{r}{2} > \frac{p-1}{2}$.

7.7.6. Applications to random volumes

Another connection to geometrical probability problems is established in Mathai (2005). This is coming from the fact that the columns of the $p \times r, r \geq p$ matrix of rank p can be considered to be p linearly independent points in a r-dimensional Euclidean space. Then the determinant of XX' represents the square of the volume content of the p-parallaletope generated by the convex hull of the p linearly independent points represented by X. If the points are random points in some sense, see for example a discussion of random points and random geometrical configurations from Mathai (1999), then we are dealing with a random volume in $|XX'|^{\frac{1}{2}}$. The distribution of this random volume is of interest in geometrical probability problems when the points have specified distributions. For problems of this type see Mathai (1999). Then the distributions of such random volumes will be based on the distribution of X where X has the very general density given in (7.7.1). Thus the existing theory in this area is extended to a very general class of basic distributions covered by the $f(X)$ of (7.7.1).

Exercises 7.7.

7.7.1. By using Stirling's approximation for gamma functions, namely,

$$\Gamma(z+a) \approx \sqrt{2\pi}\, z^{z+a-\frac{1}{2}}\, e^{-z} \tag{7.7.21}$$

for $|z| \to \infty$ and a a bounded quantity, show that the moment expressions in (7.7.13) and (7.7.14) reduce to the moment expression in (7.7.15).

7.7.2. By opening up $\Gamma_p(\cdot)$ in terms of gamma functions and by examining the structure of the gamma products in (7.7.13) show that for $q < 1$ we can write

$$E[|a(1-q)A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}|^h] = \prod_{j=1}^{p} E(x_j^h) \tag{7.7.22}$$

where x_j is a real scalar type-1 beta with the parameters

$$\left(\alpha + \frac{r}{2} - \frac{j-1}{2}, \frac{\beta}{1-q} + \frac{p+1}{2}\right),\ j = 1, ..., p.$$

7.7.3. By going through the procedure in Exercise 7.7.2 show that, for $q > 1$,

$$E[|a(q-1)A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}|^h] = \prod_{j=1}^{p} E(y_j^h) \tag{7.7.23}$$

where y_j is a real scalar type-2 beta random variable.

7.7.4. Let $q < 1, a(1-q) = 1, Y = XBX', Z = A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}$ where X has the density $f(X)$ of (7.7.1). Then show that Y has the non-standard real matrix-variate type-1 beta density and Z has standard type-1 beta density.

7.7.5. Let $q < 1, a(1-q) = 1, \alpha + \frac{r}{2} = \frac{p+1}{2}, \beta = 0, Z = A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}$ where X has the density $f(X)$ of (7.7.1). Then show that Z has a standard uniform density.

7.7.6. Let $q < 1, \alpha = 0, a(1-q) = 1, \frac{\beta}{1-q} = \frac{1}{2}(m-p-r-1)$. Then show that the $f(X)$ of (7.7.1) reduces to the inverted T density of Dickey.

7.7.7. Let $q > 1, a(q-1) = 1, Y = XBX', Z = A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}$. Then when X has the density in (7.7.1) show that Y has the non-standard matrix-variate type-2 beta density and Z has the standard type-2 beta density.

7.7.8. Let $q = 1, a = 1, \beta = 1, \alpha + \frac{r}{2} = \frac{n}{2}, Y = XBX', A = \frac{1}{2}V^{-1}$. Then show that Y has a Wishart density when X has the density in (7.7.1).

7.7.9. Let $q = 1, a = 1, \beta = 1, \alpha = 0$ in $f(X)$ of (7.7.1). Then show that $f(X)$ reduces to the real matrix-variate Gaussian density.

7.7.10. Let $q > 1, a(q-1) = 1, \alpha + \frac{r}{2} = \frac{p+1}{2}, \frac{\beta}{q-1} = 1, Y = A^{\frac{1}{2}}XBX'A^{\frac{1}{2}}$. Then if X has the density in (7.7.1) show that Y has a standard real matrix-variate Cauchy density.

References

Bochner, S. (1952). Bessel functions and modular relations of higher type and hyperbolic differential equations, *Com. Sem. Math. de l'Univ. de Lund, Tome supplémentaire dédié à Marcel Riesz*, 12-20

Constantine, A.G. (1963). Some noncentral distribution problems in multivariate analysis, *Annals of Mathematical Statistics*, **34**, 1270-1285.

Herz, C.S. (1955). Bessel functions of matrix argument, *Annals of Mathematics*, **61 (3)**, 474-523.

James, A.T. (1961). Zonal polynomials of the real positive definite matrices, *Annals of Mathematics*, **74** , 456-469.

Mathai, A.M. (1978). Some results on functions of matrix argument, *Math. Nachr.* **84** , 171-177.

Mathai, A.M. (1993). *A Handbook of Generalized Special Functions for Statistical and Physical Sciences*, Oxford University Press, Oxford.

Mathai, A.M. (1997). *Jacobians of Matrix Transformations and Functions of Matrix Argument*, World Scientific Publishing, New York.

Mathai, A.M. (1999). *An Introduction of Geometrical Probability: Distributional Aspects with Applications*, Gordon and Breach, New York.

Mathai, A.M., Provost, S.B. and Hayakawa, T. (1995). *Bilinear Forms and Zonal Polynomials*, Springer-Verlag Lecture Notes in Statistics, **102**, New York.

Mathai, A.M. (2005). A pathway to matrix-variate gamma and normal densities, *Linear Algebra and Its Applications*, **396**, 317-328.

Mathai, A.M. (1978). Some results on functions of matrix argument, *Math. Nachr.* **84** , 171-177.

Mathai, A.M. (2004). *Modules 1,2,3*, Centre for Mathematical Sciences, India.