UNIT 18 ORTHOGONAL TRANSFORMATIONS

Structure **Page No**

18.1 INTRODUCTION

You must have studied the linear transformation in your earlier studies. Now in this unit, we shall focus on a particular linear transformation called orthogonal transformation. In Section 18.2, we shall define the orthogonal transformation through a matrix. Then in Section 18.3, we shall construct the orthogonal matrices by using the orthogonality in a matrix. At the end, in the Section 18.4, we shall apply the orthogonal transformation in multivariate normal distribution.

Objectives

After reading this unit, you should be able to:

- define orthogonal transformation through orthogonal matrix;
- learn the properties of orthogonal matrix;
- use the orthogonal transformation for multivariate normal distribution.

18.2 ORTHOGONAL TRANSFORMATION

Let us recall first the linear transformation. For this consider a linear transformation $T : \Omega_X \to \Omega_Y$ which can be expressed through a matrix A.

That is if we transform X to Y by the transformation T, we can write $Y = T(X) = AX$. Consider that $\Omega_X = \Omega_Y = R^n$ and $Y = AX$ is a linear transformation. If the matrix A happens to be the orthogonal matrix, the transformation is called an orthogonal transformation. It may be noted here that orthogonal matrix is always a square matrix.

Before we proceed let us start with the orthogonal matrices first.

An orthogonal matrix is the real specialization of a unitary matrix, and thus always a normal matrix. Although we consider only real matrices here, the definition can be used for matrices with entries from any field. However, orthogonal matrices arise naturally from inner products, and for matrices of complex numbers that leads instead to the unitary requirement. Orthogonal matrices preserve inner product.

To see the inner product connection, consider a vector X in an *n*-dimensional real inner product space. Written with respect to an orthonormal basis, the squared length of X is $X'X$. If a linear transformation, in matrix form Y = M X, preserves vector lengths, then

$$Y'Y = (MX')(MX) = X'(M'M)X.$$

And it should be equal to $X'X$. This can only happen if $M'M = I$ where I is an identity matrix. Thus, we see that $M' = M^{-1}$ for orthogonal matrix M. You may also note that the columns of an orthogonal matrix are an orthonormal set of vectors. Similarly, the rows of an orthogonal matrix are an orthonormal set of rows. As noted above, in order for a matrix to be orthogonal matrix it must be a square matrix. So a matrix that is not square, but does have orthonormal columns will not be orthogonal. To make the point clear let us consider the following example:

Example 1: Consider the following matrices A and B:

$$A = \begin{bmatrix} 2 & 1 & 3 \\ -1 & 0 & 7 \\ 0 & -1 & -1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -5 & 2 \\ -3 & 1 & -4 \\ 1 & -1 & 1 \end{bmatrix}$$

We know that every matrix with linearly independent columns can be factored as QR where Q is a matrix with orthonormal columns and *R* is an invertible upper triangular matrix.

Thus, A and B can be decomposed as

$$A = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & \frac{2}{\sqrt{6}} \\ 0 & -\frac{5}{\sqrt{30}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \sqrt{5} & \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ 0 & \frac{6}{\sqrt{30}} & \frac{22}{\sqrt{30}} \\ 0 & 0 & \frac{16}{\sqrt{6}} \end{bmatrix} = Q_1 \; R_1 \text{ (say)}$$

$$B = \begin{bmatrix} \frac{2}{\sqrt{15}} & \frac{7}{\sqrt{4485}} & -\frac{118}{\sqrt{20930}} \\ \frac{1}{\sqrt{15}} & -\frac{64}{\sqrt{4485}} & -\frac{11}{\sqrt{20930}} \\ -\frac{3}{\sqrt{15}} & -\frac{18}{\sqrt{4485}} & -\frac{81}{\sqrt{20930}} \\ \frac{1}{\sqrt{15}} & -\frac{4}{\sqrt{4485}} & -\frac{18}{20930} \end{bmatrix} \begin{bmatrix} \sqrt{15} & -\frac{11}{\sqrt{15}} & \frac{17}{\sqrt{15}} \\ 0 & \frac{299}{\sqrt{4485}} & -\frac{53}{\sqrt{4485}} \\ 0 & 0 & \frac{210}{\sqrt{20930}} \end{bmatrix} = Q_2R_2 \text{ (say)}$$

For the matrix A, the matrix Q, given as

$$Q_1 = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & \frac{2}{\sqrt{6}} \\ 0 & -\frac{5}{\sqrt{30}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

by construction has orthonormal columns and since it is a square matrix it is an orthogonal matrix. But for the matrix B, the matrix

$$Q_2 = \begin{bmatrix} \frac{2}{\sqrt{15}} & \frac{7}{\sqrt{4485}} & -\frac{118}{\sqrt{20930}} \\ \frac{1}{\sqrt{15}} & -\frac{64}{\sqrt{4485}} & -\frac{11}{\sqrt{20930}} \\ -\frac{3}{\sqrt{15}} & -\frac{18}{\sqrt{4485}} & -\frac{81}{\sqrt{20930}} \\ \frac{1}{\sqrt{15}} & -\frac{4}{\sqrt{4485}} & -\frac{18}{\sqrt{20930}} \end{bmatrix}$$

Again by construction has orthonormal columns but it is not a square matrix therefore it is not an orthogonal matrix.

Now, try an exercise.

E1) Find value(s) for a, b, and c for which the following matrix A will be

orthogonal $A = \begin{bmatrix} 0 & -\frac{2}{3} & a \\ \frac{1}{\sqrt{5}} & \frac{2}{3} & b \\ -\frac{2}{\sqrt{5}} & \frac{1}{3} & c \end{bmatrix}$.

You may also deduce from the above discussions that finite-dimensional linear isometries-rotations, reflections, and their combinations-produce orthogonal matrices. The converse is also true: orthogonal matrices imply orthogonal transformations. However, linear algebra includes orthogonal transformations between spaces which may be neither finite-dimensional nor of the same dimension, and these have no orthogonal matrix equivalent.

Orthogonal matrices are important for a number of reasons, both theoretical and practical. The $n \times n$ orthogonal matrices form a group, the orthogonal group denoted by O(n), which, with its subgroups, is widely used in mathematics and the physical sciences. For example, the point group of a molecule is a subgroup of O(3). Because floating point versions of orthogonal matrices have advantageous properties, they are key to many algorithms in numerical linear algebra, such as QR decomposition. As another example, with appropriate normalization the discrete cosine transform (used in MP3 compression) is represented by an orthogonal matrix. Below are a few examples of small orthogonal matrices and possible interpretations:

Example 2:

S. No.	**Matrix**	**Interpretations**
1.	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	Identity transformation
2.	$\begin{bmatrix} 0.96 & -0.28 \\ 0.28 & 0.96 \end{bmatrix}$	Rotation by 16.26^0
3.	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	Reflection across x-axis

Now, try an exercise.

E2) Check whether the following matrix A has roto inversing or reflection across x-axis or y-axis.

$$A = \begin{bmatrix} 0 & -0.80 & -0.60 \\ -0.80 & -0.36 & 0.48 \\ 0.60 & 0.48 & -0.64 \end{bmatrix}.$$

Now, in the following section we shall discuss the methods of construction of orthogonal transformation.

18.3 CONSTRUCTION OF ORTHOGONAL TRANSFORMATION

In this section, we will discuss briefly the method of construction of orthogonal transformations. We have seen in the previous section that an orthogonal transformation Y = AX can be constructed through an orthogonal matrix A. Now, we will see how we can develop an orthogonal matrix of A of given dimension.

Let us start with the matrix of order 1 that is one dimensional.

The simplest orthogonal matrices are the 1×1 matrices namely $[1]$ and $[-1]$ which we can be interpreted as the identity and a reflection of the real line across the origin.

Consider the case of two dimensional arrays that is the matrix of order 2×2. The 2×2 matrices have the form

$$\begin{bmatrix} p & t \\ q & u \end{bmatrix}.$$

The orthogonality demands to satisfy the following three equations:

$$1 = p^2 + q^2$$
$$1 = t^2 + u^2$$
$$0 = pt + qu.$$

Since there are four unknowns and three equations, more than one solution exists. Let us consider that $p = \cos\theta,\ q = \sin\theta$; then either $t = -q,\ u = p$ or $t = q,\ u = -p$. We can interpret the first case as a rotation by θ (where $\theta = 0$ is the identity), and the second as a reflection across a line at an angle of $\theta/2$. Thus, we see that

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

is an orthogonal transformation which is a rotational transformation. Similarly,

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

is an orthogonal transformation which is a reflectional transformation.

The reflection at 45° exchanges X_1 and X_2; it is a permutation matrix, with a single 1 in each column and row and 0 otherwise, i.e.

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The identity is also a permutation matrix.

A reflection is its own inverse, which implies that a reflection matrix is symmetric (equal to its transpose) as well as orthogonal. The product of two rotation matrices is a rotation matrix, and the product of two reflection matrices is also a rotation matrix.

Now, let us discuss this for matrices of higher dimension.

Regardless of the dimension, it is always possible to classify orthogonal matrices as purely rotational or not, but for 3×3 matrices and larger the non-rotational matrices can be more complicated than reflections. For example,

$$\begin{bmatrix} -1 & 02 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

represents an *inversion* through the origin and a *rotoinversion* about the z-axis. Rotations also become more complicated; they can no longer be completely characterized by an angle, and may affect more than one planar subspace. While it is common to describe a 3×3 rotation matrix in terms of an axis and angle, the existence of an axis is an accidental property of this dimension that applies in no other. However, we have elementary building blocks for permutations, reflections, and rotations that apply in general.

Primitives

The most elementary permutation is a transposition, obtained from the identity matrix by exchanging two rows. Any $n\times n$ permutation matrix can be constructed as a product of at most $n-1$ transpositions.

A Householder reflection is constructed from a non-null vector $\mathbf{v}$. Here the numerator is a symmetric matrix while the denominator is a number, the squared magnitude of v. This is a reflection in the hyperplane perpendicular to v (negating any vector component parallel to v). If v is a unit vector, then $Q = I - 2\mathbf{v}\mathbf{v}^T$ suffices. A Householder reflection is typically used to simultaneously zero the lower part of a column. Any orthogonal matrix of size $n\times n$ can be constructed as a product of at most n such reflections.

A given rotation acts on a two-dimensional (planar) subspace spanned by two coordinate axes, rotating by a chosen angle. It is typically used to zero a single sub diagonal entry. Any rotation matrix of size $n\times n$ can be constructed as a product of at most $n(n-1)/2$ such rotations. In the case of 3×3 matrices, three such rotations suffice; and by fixing the sequence we can thus describe all 3×3 rotation matrices (though not uniquely) in terms of the three angles used, often called Euler angles.
A Jacobi rotation has the same form as a Givens rotation, but is used as a similarity transformation chosen to zero both off-diagonal entries of a 2×2 symmetric submatrix.

Matrix Properties of Orthogonality

A real square matrix is orthogonal if and only if its columns form an orthonormal basis of the Euclidean space $\mathbf{R}^n$ with the ordinary Euclidean dot product, which is the case if and only if its rows form an orthonormal basis of $\mathbf{R}^n$. It might be tempting to suppose a matrix with orthogonal (not orthonormal) columns would be called an orthogonal matrix, but such matrices have no special interest and no special name; they only satisfy $M^TM = D$, with D a diagonal matrix.

The determinant of any orthogonal matrix is $+1$ or -1. This follows from basic facts about determinants, as follows:

$$1 = \det(I) = \det(Q^tQ) = \det(Q^t)\det(Q) = (\det(Q))^2.$$

The converse is not true; having a determinant of $+1$ is no guarantee of orthogonality, even with orthogonal columns, as shown by the following counter example:

$$\begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

The inverse of every orthogonal matrix is again orthogonal, as is the matrix product of two orthogonal matrices.

Example 3: Construct an orthogonal matrix of order 3×3 such that the elements in the first row are equal.

Solution: Let us suppose that the elements in the first row are 'a'. In order that the matrix is orthogonal, sum of square of these must be one, i.e. $3a^2 = 1$. This shows that the matrix may be of the following forms:

$$\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ b & d & f \\ c & e & g \end{bmatrix} \text{ or } \begin{bmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ b & d & f \\ c & e & g \end{bmatrix} \text{ or } \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ b & d & f \\ c & e & g \end{bmatrix} \text{ and so on.}$$

Let us consider the first one. In order that it is an orthogonal matrix rest of the elements should fulfill the following conditions:

$$b + d + f = 0, \tag{1}$$

$$c + e + g = 0, \tag{2}$$

$$bc + de + fg = 0, \tag{3}$$

$$b^2 + d^2 + f^2 = 1, \tag{4}$$

$$c^2 + e^2 + g^2 = 1. \tag{5}$$

If we choose $b = 2/\sqrt{6}$, $d = -1/\sqrt{6}$ and $f = -1/\sqrt{6}$, Eqn. (1) and Eqn. (4) are satisfied. Eqn. (3) reduces to the following on substitution of these values:

$$2c - e - g = 0 \tag{6}$$

On adding it with Eqn. (2) we get $c = 0$ and hence $e = -g$. From Eqn. (5) we have $e = -g = \pm 1/\sqrt{2}$. Therefore all of the following will be an orthogonal matrix.

$$\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \text{ or } \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Similarly, you consider other forms and develop orthogonal matrices. Hence show that many more orthogonal matrices can be constructed.

So far, we have discussed the orthogonal matrices in detail. Now in the following section, let us focus on the orthogonal transformation in multivariate normal distribution.

18.4 USE OF ORTHOGONAL TRANSFORMATION IN MULTIVARIATE NORMAL DISTRIBUTIONS

Let X be a p-dimensional random vector such that $X = [X_1, X_2, \ldots, X_n]^t$ and the density function of X, i.e. the joint density function of $X_1, X_2, \ldots, X_p$ be $f(x_1, x_2, \ldots, x_p)$. Let us define a real valued function $Y = y(x)$ i.e.

$$Y_j = y_j(x_1, x_2, \ldots, x_p); \; j = 1, 2, \ldots, p. \tag{7}$$

If the transformation X to Y is one-to-one, the inverse transformation can be written as

$$X_i = x_i(y_1, y_2, \ldots, y_p). \quad (8)$$

Now, you know that the density function of random vector Y, i.e. the joint density of $Y_1, Y_2, \ldots, Y_p$ is obtained as

$$G(y_1, y_2, \ldots, y_p) = f(x_1(y_1, y_2, \ldots, y_p), x_2(y_1, y_2, \ldots, y_p) \ldots x_i(y_1, y_2, \ldots, y_p)).|J(y_1, y_2, \ldots, y_p)|. \quad (9)$$

where $|J|$ is the magnitude (positive value) of the Jacobian of the transformation defined below:

$$J(y_1, y_2, \ldots, y_p) = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \ldots & \ldots & \frac{\partial x_1}{\partial y_p} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \ldots & \ldots & \frac{\partial x_2}{\partial y_p} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \frac{\partial x_p}{\partial y_1} & \frac{\partial x_1}{\partial y_1} & \ldots & \ldots & \frac{\partial x_p}{\partial y_p} \end{vmatrix}. \quad (10)$$

Theorem 1: The Jacobian of an orthogonal transformation is always one.

Proof: Let us consider that a p-dimensional random vector X is transformed to p-dimensional random vector Y by an orthogonal transformation $Y = AX$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ a_{21} & a_{22} & \ldots & a_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ a_{p1} & a_{p2} & \ldots & a_{pp} \end{bmatrix}. \quad (11)$$

Thus in the expanded form the transformation is written as

$$\begin{aligned} Y_1 &= a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p \\ Y_2 &= a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ Y_p &= a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pp}X_p \end{aligned} \quad (12)$$

Now the inverse transformation can easily be obtained by recalling that $Y = AX$ is an orthogonal transformation and A is an orthogonal matrix. Thus, $A^tA = I$. Pre-multiplying by A^t in $Y = AX$ on both the sides, we have $A^tY = A^tAX = X$, therefore the inverse transformation in extended form can be written as

$$\begin{aligned} X_1 &= a_{11}Y_1 + a_{21}Y_2 + \cdots + a_{p1}Y_p \\ X_2 &= a_{12}Y_1 + a_{22}Y_2 + \cdots + a_{p2}Y_p \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ X_p &= a_{1p}Y_1 + a_{2p}Y_2 + \cdots + a_{pp}Y_p \end{aligned} \quad (13)$$

Now, using Eqn. (10), we get the Jacobian of the transformation as

$$J = \begin{vmatrix} a_{11} & a_{21} & \dots & \dots & a_{p1} \\ a_{12} & a_{22} & \dots & \dots & a_{p2} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{1p} & a_{2p} & \dots & \dots & a_{pp} \end{vmatrix} = |A^t| \tag{14}$$

But $A^tA = I$, therefore $|A^tA| = 1 \Rightarrow |A^t||A| = 1$. The magnitude of the determinant of A and A^t are same and hence magnitude of $|A^t|$ = magnitude of $|A| = 1$.

Remark: From the above theorem we see that if we have to use orthogonal transformation, there is in fact no need to calculate the Jacobian of the transformation. The density of the transformed new random variable can simply be obtained by replacing the old variables in terms of new variable as obtained from the inverse transformation. This ease of the orthogonal transformation gives the following interesting result which is popularly known as **Fisher's lemma.**

Theorem 2: Let X is p component random vector following normal distribution with mean vector 0 and variance covariance matrix I (now onwards it will be denoted as $X \sim N_p(0, I)$. Then every orthogonal transformation $Y = AX$ produces a new random vector Y having similar distribution, i.e. $Y \sim N_p(0, I)$.

Proof: Since $X \sim N_p(0, I)$, the density function of x is given as:

$$f(X) = (2\pi)^{-p/2} \exp(-X^tX) \tag{15}$$

Hence the density of $Y = AX$, where A is orthogonal matrix can be easily obtained, using the above theorem by replacing X by A^tY in (16), as

$$\begin{aligned} g(Y) &= (2\pi)^{-p/2} \exp(-(A^tY)^t(A^tY)) \\ &= (2\pi)^{-p/2} \exp(-Y^tA^tAY) \\ &= (2\pi)^{-p/2} \exp(-Y^tY) \qquad \text{(since } A^tA = I) \end{aligned} \tag{16}$$

Remark: The Eqn. (15) can be written as

$$f(X_1, X_2, \dots, X_p) = (2\pi)^{-p/2} \exp\left[-\left(\sum_{i=1}^{p} X_i^2\right) = \prod_{i=1}^{p} \frac{1}{2\pi} \exp(X_i^2)\right] \tag{17}$$

This means that X_i's are independently and identically distributed (I.I.D.) standard normal variables. Thus, the theorem can also be stated as "Orthogonal transformation on a set of I.I.D. standard normal variables gives a new set of I.I.D. standard normal variables".

Let us consider another application of orthogonal transformation in studying the effect of independent linear constraints on chi-square distribution. Let X be an n-dimensional random vector such that $X \sim N_n(0, I)$, i.e. components X_i's are I.I.D. standard normal variables. Therefore you know that $X^tX = \sum_{i=1}^{n} X_i^2$ is distributed as chi-square with n degrees of freedom. However, if the variables are subject to linear constraints, the degrees of freedom of X^tX is reduced.

Theorem 3: If $X \sim N_n(0, I)$ where X is subjected to m linearly independent constraints $BX = 0$, the distribution of X^tX is chi-square with $n - m$ degrees of freedom.

Proof: Let us consider the expanded form of the constraints given below:

$$\begin{aligned} &b_{11}X_1 + b_{12}X_2 + \cdots + b_{1n}X_n = 0 \\ &b_{21}X_1 + b_{22}X_2 + \cdots + b_{2n}X_n = 0 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &b_{m1}X_1 + b_{m2}X_2 + \cdots + b_{mn}X_n = 0 \end{aligned} \qquad (18)$$

Consider a transformation $Y = AX$ such that jth element in the first row of the matrix A is

$$a_{1j} = \frac{b_{1j}}{\sqrt{\sum_{i=1}^{n} b_{1j}^2}} \text{ for } j = 1, 2, \ldots, n. \qquad (19)$$

Now, we choose rest $n(n-1)$ of the elements in A such that it is an orthogonal matrix, i.e. we have $n(n-1))$ unknowns to be obtained by solving $(n-1)(n-2)/2$ equations for $n(n-1)$ unknowns. Since the number of unknowns is more than the number of equations; more than one solution exists. Further,

$$Y_1 = \sum_{j=1}^{n} a_{1j}X_j = \frac{\sum_{j=1}^{n} b_{1j}X_j}{\sqrt{\sum_{j=1}^{n} b_{1j}^2}} = 0, \qquad (20)$$

because of the first linear constraints. It may also be noted that $Y^tY = (AX)^t(AX) = X^tA^tAX = X^tX$, i.e.

$$\sum_{i=1}^{n} Y_i^2 = \sum_{i=1}^{n} X_i^2 = \sum_{i=2}^{n} Y_i^2 \qquad (Y_1 = 0). \qquad (21)$$

In terms of Y_i's the constraints (18) reduces to $0 = BX = BA^tY = CY$, i.e.

$$\begin{bmatrix} b_{11} & b_{12} & \cdots & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{m1} & b_{m2} & \cdots & \cdots & b_{mn} \end{bmatrix} \begin{bmatrix} \frac{b_{11}}{\sqrt{\sum_{i=1}^{n} b_{1j}^2}} & \frac{b_{12}}{\sqrt{\sum_{i=1}^{n} b_{1j}^2}} & \cdots & \cdots & \frac{b_{1n}}{\sqrt{\sum_{i=1}^{n} b_{1j}^2}} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{bmatrix}^t \begin{bmatrix} 0 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

or,

$$\begin{bmatrix} b_{11} & b_{12} & \cdots & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{m1} & b_{m2} & \cdots & \cdots & b_{mn} \end{bmatrix} \begin{bmatrix} \frac{b_{11}}{\sqrt{\sum_{i=1}^{n} b_{1j}^2}} & a_{21} & \cdots & \cdots & a_{n1} \\ \frac{b_{12}}{\sqrt{\sum_{i=1}^{n} b_{1j}^2}} & a_{22} & \cdots & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{b_{1n}}{\sqrt{\sum_{i=1}^{n} b_{1j}^2}} & a_{2n} & \cdots & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} 0 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

or,

$$\begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ c_{21} & c_{22} & \cdots & \cdots & c_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{m1} & c_{m2} & \cdots & \cdots & c_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ Y_2 \\ \vdots \\ \vdots \\ Y_n \end{bmatrix} = 0$$

or,

$$\begin{aligned} c_{22}Y_2 + c_{23}Y_3 + \cdots + c_{2n}Y_n &= 0 \\ c_{32}Y_2 + c_{33}Y_3 + \cdots + c_{3n}Y_n &= 0 \\ &\cdots\cdots \\ &\cdots\cdots \\ c_{m2}Y_2 + c_{m3}Y_3 + \cdots + c_{mn}Y_n &= 0 \end{aligned} \qquad (22)$$

Therefore, the original problem of finding the distribution of $\sum_{i=1}^{n} X_i^2$ (sum of n standard normal variables) subject to the constraint $BX = 0$, now reduces to finding the distribution of $\sum_{i=2}^{n} Y_i^2$ (sum of $(n-1)$ standard normal variables) subject to $(m-1)$ linear constraints. Needless to mention that the transformation X to Y be orthogonal therefore Y_i's are also standard normal variables (as per Theorem (2)).

Now we repeat the above process through transforming Y to Z using an orthogonal transformation $Z = A^* Y$ (where $A^* = ((a_{ij}))$; i, J = 2, 3, ...n) and choosing

$$a_{2j} = \frac{c_{2j}}{\sqrt{\sum_{j=2}^{n} c_{2j}^2}}.$$

We see that original problem of finding the distribution of sum of n standard normal variables subject to m linear constraints, now reduces to finding the distribution of sum of $(n-2)$ standard normal variables subject to $(m-2)$ linear constraints. Repeating this process for m times we conclude that the problem finally reduces to find the distribution of sum of square of $(n-m)$ standard normal variables free from any constraint thus follows chi-square with $(n-m)$ degrees of freedom.

Now, we shall summarize the unit.

18.5 SUMMARY

In this unit, we have covered the following points.

1. We introduced the concepts transformation particularly orthogonal transformation. Also the advantages of orthogonal transformation were discussed.
2. We generated orthogonal matrices of our interest for manipulating the problems in finding the distributions. The examples given in the unit were deliberately taken to be of low dimension; however they will learn the procedure and may use it for higher dimension also.
3. The orthogonal transformation of multivariate normal distribution was discussed in detail.

18.6 SOLUTIONS/ANSWERS

E1) You see that the columns of the matrix M are

$$M_1 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix} \quad M_2 = \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} \quad M_3 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

You can verify that M_1 and M_2 are orthonormal vectors possess the properties that $M_1.M_1 = 1$, $M_2.M_2 = 1$ and $M_1.M_2 = 0$, i.e. these fulfill the requirement of columns of an orthogonal matrix. Therefore, we need to find a, b and c such that $M_3.M_3 = 1$, $M_1.M_3 = 0$ and $M_2.M_3 = 0$, i.e.

$$M_1.M_3 = \frac{b}{\sqrt{5}} - \frac{2c}{\sqrt{5}} = 0$$

$$M_2.M_3 = -\frac{2}{3}a + \frac{2}{3}b + \frac{1}{3}c = 0$$

From the first dot product we can see that $b = 2c$. Substituting this into the second dot product, we get $5c = 2a$. Combining these results we have the values of b and c in terms of a as $b = 4a/5$ and $c = 2a/5$. Now, in order that the third column be orthonormal to make the matrix M orthogonal, we must have $M_3.M_3 = 1$.

Since $M_3 = \begin{bmatrix} a \\ \frac{4}{5}a \\ \frac{2}{5}a \end{bmatrix}$ $M_3.M_3 = a^2 + \frac{16}{25}a^2 + \frac{4}{25}a^2 = 1$

This gives us two possible values of a namely $a = \pm\frac{5}{\sqrt{45}}$ that we can use and this in turn means that we could used either of the following two vectors

$$M_3 = \begin{bmatrix} \frac{5}{\sqrt{45}} \\ \frac{4}{\sqrt{45}} \\ \frac{2}{\sqrt{45}} \end{bmatrix} \text{ or, } M_3 = \begin{bmatrix} -\frac{5}{\sqrt{45}} \\ -\frac{4}{\sqrt{45}} \\ -\frac{2}{\sqrt{45}} \end{bmatrix}$$

E2)

S.No.	Matrix	Interpretations
1.	$\begin{bmatrix} 0 & -0.80 & -0.60 \\ -0.80 & -0.36 & 0.48 \\ 0.60 & 0.48 & -0.64 \end{bmatrix}$	Rotoinversion: axis (0, -3/5, 4/5), angle 90^0

18.7 PRACTICAL ASSIGNMENT

1) Write a program in 'C' language to show that the matrix A is orthogonal.

$$A = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$$

Also, extend the program to check that the columns of A form an orthonormal basis of R^4.

2) Write a program in 'C' language to develop an orthogonal matrix of order 3×3 such that the elements in the first row are equal.