UNIT 9 LINEAR SPACES - 2

Structure

9.0 Objectives
9.1 Introduction
9.2 Vectors and Matrices
9.3 Characteristic Value Problem
  9.3.1 Characteristic Equation
  9.3.2 Sum and Product of Roots
  9.3.3 Characteristic Vector
  9.3.4 Diagonalisation
9.4 Linear independence of Eigen Vectors
9.5 Quadratic Forms
9.6 Definiteness and Eigen Values
9.7 Vector Differentiation
  9.7.1 Vector Differentiation of a Linear Function
  9.7.2 Vector Differentiation of a Vector Functions
  9.7.3 Vector Differentiation of a Quadratic Form
  9.7.4 Vector Differentiation of a Bi-linear Form
9.8 Summary
9.9 Answer/Hint to Check Your Progress
9.10 References and Further Readings

9.0 INTRODUCTION

In this unit, we shall consider only vectors with a finite number of elements. Further, unless stated otherwise, we shall assume that each element in a vector is a real number. Thus, we are typically concerned with the ***n-tuples*** $(x_1,x_2,...,x_n)$ of real numbers. A vector all of whose elements are zero is called the *'null vector'*.

9.1 OBJECTIVES

After going through this unit, you should be able to:

- Present a system of linear equations as a matrix equation;
- explain the concepts of characteristic matrix, characteristic vector; characteristic root, eigenvector, eigenvalue and explain the characteristic value problem;
- identify the concept of quadratic form;
- present various types of quadratic form;
- discuss the relation of eigen values to definiteness of a quadratic form;
- explain the concept to vector differentiation; and
- describe the processes of various types of vector differentiation.

9.2 VECTORS AND MATRICES

A system of 'm' linear equations in 'n' variables (x_1,x_2, ...,x_n) can be arranged into the following format:

$a_{11}x_1+a_{12}x_2+...+a_{1n}x_n = d_1$

$a_{21}x_1+a_{22}x_2+...+a_{2n}x_n = d_2$

.....

$a_{m1}x_1+a_{m2}x_2+...+a_{mn}x_n = d_m$

There are three types of elements in the above equation system. The first is the set of coefficients a_{ij}; the second is the set of variables x_1, x_2, ..., x_n, and the last is the set of constant terms d_1, ..., d_m. From our earlier discussion in Unit 6 on Matrices, we know that if we arrange the three sets as three rectangular arrays and label them, respectively, as A, X and d (without subscripts), then the above system of equations can be written in the matrix form as:

AX = d, where,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}_{mxn} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{nx1} \quad d = \begin{bmatrix} d \\ d \\ \vdots \\ d_m \end{bmatrix}_{mx1}$$

Example: Consider the following linear- equations system

$6x_1 + 3x_2+x_3 = 22$

$x_1+4x_2-2x_3 = 12$

$4x_1-x_2+5x_3 = 10$

We can write, this system as the matrix equation

$$\begin{bmatrix} 6 & 3 & 1 \\ 1 & 4 & -2 \\ 4 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 22 \\ 12 \\ 10 \end{bmatrix}.$$

9.3 CHARACTERISTIC VALUE PROBLEM

A problem which arises frequently in application of linear algebra is that of finding values of a scalar parameter λ for which there exists vector x ≠ 0 (i.e., non-null vector) satisfying Ax= λx, where A is a given nth order square matrix. Such a problem is called a ***characteristic value ('Eigen Value', or proper value)*** **problem.** *If* x ≠ 0 satisfies the equation Ax= λx for a given vector x, then matrix A operating on vector x yields a vector which is a scalar multiple of x.

The scalar λ ***is called the characteristic value (or equivalently, characteristic root, Eigen value or proper value)*** and ***x is called the characteristic vector (or 'Eigen Vector')*** associated with A, if Ax= λx

9.3.1 Characteristic Equation

For given n, to solve the system of equations Ax= λx, for λ, it can be written as (A −λI) x touching 0 (A−λI x = 0), Where **I is a unit matrix of order n**. Now, if the system has to have a non-trivial solution for the variable vector X, then rank [A−λ1] < n , which is true if and only if

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0 \qquad \text{(i)}$$

The LHS in **(i)** gives us an n^{th} degree polynomial equation in λ. The equation (i) is called the ***'characteristic equation'*** of matrix A and the left-hand side is called the ***'characteristic polynomial'*** **of matrix A**.

Example: Characteristic polynomial for a square matrix of order 2.

$$\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21} = 0$$

$$\Rightarrow \lambda^2 - \lambda\,(a_{11}+a_{22}) + (a_{11}a_{22} - a_{12}a_{21}) = 0$$

which is a quadratic equation in λ.

Note: if A is of order n, then the equation | A−λI | = 0 has n roots which may be real or complex, distinct or multiple, and zero or non-zero.

9.3.2 Sum and Product of Roots

Consider the equation $\lambda^2 - \lambda(a_{11} + a_{22}) + (a_{11}a_{22} - a_{12}a_{21}) = 0$ and let us write it as $c_0\lambda^0 + c_1\lambda + c_2 = 0$ where, $c_0 = 1, c_1 = -(a_{11} + a_{22}), c_2 = a_{11}a_{22} - a_{12}a_{21}$

Applying the knowledge of determining roots of the quadratic equation, we can write,

$$\lambda_1\lambda_2 = \frac{c_2}{c_0} = a_{11}a_{22} - a_{21}a_{12} = |A|$$

In general, if A is a square matrix of order n, then the characteristic equation $|A - \lambda I| = 0$ can be written as $c_0\lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_{n-1}\lambda + c_n = 0$

where $c_0 = 1, c_1 = -\sum_{I=1}^{n} a_{ii}, and\ c_n = (-1)^n |A|$

c_2 = Sum of all principal minors of A of order 2.

$c_j = (-1)^j \sum$ all principal minor or order j

Thus, if characteristic roots are $\lambda_1, \lambda_2, \ldots\ldots\ldots, \lambda_n$,

then $\sum_{i=1}^{n} \lambda i = \sum_{i=1}^{n} a_{ii} = Trace\ (A) and\ \lambda_1, \lambda_2, \ldots . \lambda_n = \prod^{n}{}_{i=1} \lambda_i = |A|$

where Trace of a square matrix is the sum of its principal diagonal elements.

Example: Let, A = $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

The characteristic equation is

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 1 = 0$$

or, $4 - 4\lambda + \lambda^2 - 1 = 0$

or, $\lambda^2 - 4\lambda + 3 = 0$

$$\lambda = \frac{4 \pm \sqrt{16 - 12}}{2} = \frac{4 \pm 2}{2} = 3,1\ i.e. \lambda_1 = 3, \lambda_2 = 1$$

Example:

Find all the eigen values of A = $\begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 0 \\ 1 & 2 & 3 \end{bmatrix}$

Solution Characteristic equation of A is: $|A - \lambda I_3| = 0$

$$\Rightarrow \begin{vmatrix} \mathbf{2-\lambda} & \mathbf{1} & \mathbf{3} \\ \mathbf{1} & \mathbf{-1-\lambda} & \mathbf{0} \\ \mathbf{1} & \mathbf{2} & \mathbf{3-\lambda} \end{vmatrix} = \mathbf{0}$$

Expanding along C_1, we get

$$(2-\lambda)\begin{vmatrix}-(1+\lambda) & 0\\ 2 & (3-\lambda)\end{vmatrix} - \begin{vmatrix}1 & 3\\ 2 & 3-\lambda\end{vmatrix} + \begin{vmatrix}1 & 3\\ -1-\lambda & 0\end{vmatrix} = 0$$

$$\Rightarrow -(\mathbf{2}-\lambda)(1+\lambda)(3-\lambda) - (3-\lambda-6) + 3(1+\lambda) = 0$$

$$\Rightarrow \lambda^3 - 4\lambda^2 - 3\lambda = 0$$

$$\Rightarrow \lambda(\lambda^2 - 4\lambda - 3) = 0$$

$$\Rightarrow \lambda = 0 \quad or\ \lambda = \frac{1}{2}(4 \pm \sqrt{28} = \frac{1}{2}(2 \pm \sqrt{7})$$

Properties:

i) If A is singular square matrix, i.e., |A| = 0, then at least one characteristic root of A is zero.

ii) If A is symmetric all roots are real, i.e., a real symmetric matrix cannot have complex characteristic roots.

iii) If $\lambda_1, \lambda_{2,\ldots\ldots\ldots}, \lambda_n$ are the roots of a non-singular matrix A, then the roots of A^{-1} are $1/\lambda_1, 1/\lambda_2, \ldots\ldots 1/\lambda_n$

iv) The eigen values of a diagonal or a triangular matrix coincide with the diagonal elements.

v) If $\lambda_1, \lambda_{2,\ldots\ldots\ldots}, \lambda_n$ are the characteristic roots of A and k is an integer, then the roots of A^k are ${\lambda_1}^k, {\lambda_2}^k, \ldots\ldots {\lambda_n}^k$

9.3.3 Characteristic Vector

Let us now consider the determination of the value of **x** of the equation Ax = λ**x**. Note that is chosen in such a way that [A - λI] is singular, i.e., its rank is less than n. This ensures that [A- λ I]X=0 is a homogeneous system of equations having a non-trivial solution. But this solution may not be unique. We may make it unique through some additional restrictions. Further, one set of non-trivial solution x^i will correspond root of λ_i.

Note: In the above and following discussion the 'i' in x^i, is NOT used in the sense power of x, but as superfix, just like a suffix.

Example: Let us find the characteristics vector for the matrix $A = \begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}$

As we derived earlier, $\lambda_1 = 3, \lambda_2 = 1$.

$\therefore$ For $\lambda_1 = 3, A - \lambda_i I = \begin{bmatrix}2-3 & 1\\ 1 & 2-3\end{bmatrix} = \begin{bmatrix}-1 & 1\\ 1 & -1\end{bmatrix}$

To get $x^1 = \begin{bmatrix}x_1^1\\ x_2^1\end{bmatrix}$,

From $(A-\lambda_i I)\, x^1$, we have

$$\left.\begin{matrix} -x_1^1 + x_2^1 = 0 \\ x_1^1 - x_2^1 = 0 \end{matrix}\right\} \Rightarrow x_1^1 = x_2^1$$

So, the number of solutions (*one equation involving two variables*) is infinite.

Let us choose one by setting $x_1^1 = 1$, so that $x^1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Next, take $\lambda_2 = 1$, and proceeding in the same way, we get, $x_2^2 = -x_1^2$

Again, setting $x_2^2 = 1$, we get $x^2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

We could normalize x^1 and $x^2, i.e.,$ define these such that $\|x^1\| = \|x^2\|$.

This would give us $x^1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $x^2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

Note: x^1 and x^2 are orthogonal to each other.

Example:

Determine all the eigen values and eigen vectors of $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & -1 & -1 \\ -3 & 6 & 3 \end{bmatrix}$

Solution:

Characteristic equation of A is $|A - \lambda I_3| = 0$

$$\Rightarrow \begin{vmatrix} 1-\lambda & 0 & -1 \\ 1 & -1-\lambda & -1 \\ -3 & 6 & 3-\lambda \end{vmatrix} = 0$$

Expanding along R_1, we get $(1-\lambda)\begin{vmatrix} -(1+\lambda) & -1 \\ 6 & 3-\lambda \end{vmatrix} - \begin{vmatrix} 1 & -1-\lambda \\ -3 & 6 \end{vmatrix} = 0$

$$\Rightarrow -(1-\lambda)(1+\lambda)(3-\lambda) + 6 - [6 - 3(1+\lambda)] = 0$$

$$\Rightarrow -(\mathbf{1} - \lambda^2)(3-\lambda) + 6 - 3 + 3\lambda + 6 = 0$$

$$\Rightarrow -(\lambda^3 - 3\lambda^2 - \lambda + 3) - 3 + 3\lambda + 6 = 0$$

$$\Rightarrow \lambda(\lambda^3 - 3\lambda + 2) = 0 \Rightarrow \lambda = 0,1,2$$

To find eigen vector of 0.

We have to find $x = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

Such that Ax = 0x = 0

$\Rightarrow x - z = 0$

x + y - z = 0

-3x+6y+3z = 0

Solving we get x = z, y = 0.

∴ a required eigen vector of 0 is $\begin{bmatrix}1\\0\\1\end{bmatrix}$

To find eigen vector of 1.

$(A - \mathrm{I}_3) \times = 0$

$\Rightarrow -z = 0$

x-2y-z = 0

-3x+6y+2z = 0

$\Rightarrow x = 2y, z = 0$

∴ a required eigen vector of 1 is $\begin{bmatrix}2\\1\\0\end{bmatrix}$

To find eigen vector 2.

$(A - 2\mathrm{I}_3)\mathrm{X} = 0$

$\Rightarrow -x + 0y - z = 0$

$x - 3y - z = 0$

$-3x + 6y + z = 0$

$\Rightarrow x = -z, 3y = 2z.$

We may take the eigen vector as $\begin{bmatrix}3\\2\\-3\end{bmatrix}$

9.3.4 Diagonalisation

It is a general property of real symmetric matrices that their roots are always real and that the associated characteristic vectors are mutually orthogonal. This implies that such a matrix can always be diagonalized by P, the matrix of n characteristic vectors.

Continuing with the previous example, if we define a 2 x 2 matrix as

$$P \equiv (x^1 x^2) = \begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}$$

Then $P^T P = I$, i.e., P is an orthogonal matrix.

and $P^T AP = \begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}\begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}$

$= \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \Lambda$

i.e., $P^T AP = \Lambda$, where Λ is the diagonal matrix consisting of the characteristic values of A as its diagonal elements.

Example:

Diagonalize the matrix $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & -1 & -1 \\ -3 & 6 & 3 \end{bmatrix}$

Solution:

We have already seen that eigen vectors of A are

$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$

Let $P = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 0 & -3 \end{bmatrix}$

$\Rightarrow P^{-1} = \frac{1}{2}\begin{bmatrix} 3 & -6 & -1 \\ -2 & 6 & 2 \\ 1 & -2 & -1 \end{bmatrix}$

We now show that

$P^{-1}AP \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$

Where $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 1$.

We have $P^{-1}AP = (P^{-1}A)P$

$$= \frac{1}{2}\begin{bmatrix} 3 & -6 & -1 \\ -2 & 6 & 2 \\ 1 & -2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 1 & -1 & -1 \\ -3 & 6 & 3 \end{bmatrix} P$$

$$= \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 \\ -2 & 6 & 2 \\ 2 & -4 & -2 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 0 & -3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ -1 & 3 & 1 \\ 1 & -2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 0 & -3 \end{bmatrix}$$

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

Note: This is general procedure to diagonalize a square matrix.

9.4 LINEAR INDEPENDENCE OF EIGEN VECTORS

Another important property, in this context, is that, if $\lambda_1, \lambda_{2,\ldots\ldots,}\lambda_n$ *are distinct*, then the corresponding characteristic vectors viz. $\mathbf{x}_1$, $\mathbf{x}_2$,..., $\mathbf{x}_n$, of **A**, are *linearly independent*. The condition (for the roots to be distinct) is sufficient but not necessary. Even when some root, or more than one root, is repeated, it may be possible to find a set of n characteristic vectors, which are linearly independent.

Moreover, if two roots are distinct then the characteristic vectors corresponding to them will be orthogonal.

9.5 QUADRATIC FORMS

Consider the expression $a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2$ where a_{11}, a_{12} and a_{22} are given constants and x_1 and x_2 are real variables. Expression such as this--being homogeneous of degree 2 in x_1 and x_2-is called **quadratic form.**

Generally, on evaluation of the expression, its sign would depend not only on a's but also on the values of x_1 and x_2. But sometimes, given the values of a's, it is possible to determine the sign of the expression, whatever x_1 and x_2 may be.

For example, for $Q(x_1, x_2) = x_1^2 - 4x_1x_2+4x_2^2 = (x_1-2x_2)^2$,

$Q(x_1, x_2)$ is non-negative for all values of x and x_2.

Similarly, for $Q(x_1,x_2)= -10x_1^2+6x_1 x_2 - x_2^2 = - x_1^2 - (3x_1-x_2)^2$,

$Q(x_1, x_2)$ is non-positive for all values of (x_1, x_2).

The ***quadratic forms*** that have a definite sign for all values of (x_1, x_2) are said to ***be either positive definite or negative definite depending on the sign***.

Other quadratic equations, which could vary in sign, are said to be ***indefinite*** quadratic forms.

In general, with variables x_1, x_2 ,..., x_n, we may have a quadratic form as:

$$Q(x_1, x_2 , \ldots , x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, x_i\, x_j$$

Thus, Q(.) is ***positive definite quadratic form*** if $Q(x)>0 \; \forall \; \mathbf{X} \in R_{nn,} \; X \neq 0$

Or, positive semi-definite if $Q(x) \geq 0 \; \forall \; x \in R_n$

Identical definitions hold for ***negative definite and negative semi-definite quadratic forms.***

Consider the equation: $Q(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$

$$= a_{11}\left[(x_1 + \frac{a_{12}}{a_{11}}x_2)^2 + \left(\frac{a_{22}}{a_{11}} - \frac{a_{12}^2}{a_{11}^2}\right).x_2^2\right]$$

$$= a_{11}\left[(x_1 + \frac{a_{12}}{a_{11}}x_2)^2 + \left(\frac{a_{11}a_{22} - a_{12}^2}{a_{11}}\right)x_2^2\right]$$

Clearly, $Q(x_1, x_2) > 0$ for all $(x_1, x_2) \neq (0,0)$ *if* $a_{11} > 0$ and $\frac{a_{11}a_{22} - a_{12}^2}{a_{11}} > 0$

Thus, $a_{11} > 0, a_{11}a_{22} - a_{12}^2$ are sufficient conditions for the quadratic form Q to be positive definite.

These are *necessary* conditions also, as for $a_{11} < 0$, values of Q can be less than zero with high value of x_1, and low value of x_2, both in magnitudes. Also, for $(a_{11}a_{22} - a_{12}^2) < 0$, Q can be less than zero if $a_{11} < 0$.

Example: Write down the quadratic form :

$Q(x , y) = x^2 + 6xy + 10y^2$ (1)

in the symmetric matrix form. Show that $Q(x, y) > 0 \; \forall \; (x, y) \neq (0,0)$

Solution:

We can write (1) as:

$x^2 + 3xy + 3xy + y^2$

$= x(x+3y) + (3x+y)y$

$= [x+3y \quad 3x+y] \begin{bmatrix} x \\ y \end{bmatrix}$

$= [x \quad y] \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

We have $Q(x,y) = x^2 6xy + 9y^2 + y^2$

$= (x+3y)^2 + y^2$

Note that

$Q(x, y) = 0$

$\Leftrightarrow x + 3y = 0 \; and \; y = 0$

$\Leftrightarrow x = 0, y = 0$

Thus, $Q(x, y) > 0 \; \forall \; (x, y) \neq (0,0)$.

9.6 DEFINITENESS AND EIGEN VALUES

A necessary and sufficient condition for Q to be positive semi- definite is that |A| = 0 and $a_{11} \geq 0$, where A = $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$.

In terms of characteristic roots, the necessary and sufficient conditions for Q to be positive semi-definite are that [A- λI] = 0, $\lambda_1 \geq 0$, and $\lambda_2 \geq 0$.

For Q to be negative semi definite, the corresponding conditions are $a_{11} \leq 0$ and |A| = 0 and in terms of characteristic roots, $\lambda_1 \leq 0$, and $\lambda_2 \leq 0$.

9.7 VECTOR DIFFERENTIATION

In some maximisation and minimisation problems, expressions involving vectors and matrices must be differentiated. The derivatives of simple expressions involving vectors and matrices can be obtained directly from the definitions. The vector derivatives of linear functions, quadratic forms and bilinear forms are defined and illustrated in this section. Second and higher order derivatives are not discussed explicitly, but can be obtained by successive differentiation in the usual manner.

9.7.1 Vector Differentiation of a Linear Function

An expression of the form **a'x**, where **a** is nx1 matrix and **x** is nx1 vector, is called a linear function For the linear function:

$$\mathbf{a'x} = [a_1 \quad a_2 \quad \ldots \quad a_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_1x_1+a_2x_2+ \ldots +a_nx_n ,$$

The partial derivative of **a'x** with respect to the scalars x_i, i = 1, 2, … , n , are $\partial(\mathbf{a'x}) / \partial x_i = a_i$,

that is, the partial derivatives are the elements of the vector a. Thus, if the n partial derivatives are arranged as a vector a, the vector differentiation is given by $\frac{\partial(\mathbf{a'x})}{\partial x} = a$, where $\frac{\partial(\mathbf{a'x})}{\partial x}$ denotes the operation of differentiating **a'x** with respect to the elements of the vector **x**.

Check Your Progress -1

1) Define the terms: eigen value, eigen vector and characteristic equations.

.....

.....

.....

2) If a = $\begin{bmatrix} 2a \\ -a \\ 3a \\ a \end{bmatrix}$ and x = $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ then prove that $\frac{\partial}{\partial x}(a'x) = a$

.....

.....

.....

3) Write down the quadratic form :

$Q(x, y, z) = x^2 + 2y^2 - 7z^2 + 4xy + 8yz - 6zx$ (1)

in the symmetric matrix form.

.....

.....

.....

4) Write down the following quadratic forms in the symmetric matrix form.

i. $3x^2 + 6xy + 11y^2$

ii. $5x^2 + 20xy + 8y^2$

iii. $x^2 + 12xy + 5y^2$

.....

.....

.....

5) Write down the following quadratic forms in the symmetric matrix form.

i. $6x^2 + 3y^2 + 3z^2 - 4xy - 2yz + 4zx$

ii. $x^2 + 2y^2 - 6xy + 20xz$

iii. $3x^2 + 3y^2 + 4xy + 6zx + 10yz$

.....

.....

.....

9.7.2 Vector Differentiation of a Vector of Functions

Let **y** be an n dimensional column vector each element of which is a function of the m elements of the vector **x**=(x_1, x_2, ..., x_m), that is, $y_i = f_i(x_1, x_2, ..., x_m)$, i=1, 2, … n.

Then each y_i can be differentiated partially with respect to each x_j and these partial derivatives can be arranged in an m x n matrix as follows:

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_m} & \frac{\partial y_2}{\partial x_m} & & \frac{\partial y_n}{\partial x_m} \end{bmatrix}$$

Example:

If y = $\begin{bmatrix} x_1^2 + 2x_2^2 - 3x_1x_3 \\ 3x_1x_2^2 - x_2^2 + 4x_2x_3^2 \end{bmatrix}$,

find vector differentiation of y w.r.t the vector x = [x_1, x_2 , x_3].

Solution:

Here, the vector y = $\begin{bmatrix} y1 \\ y2 \end{bmatrix}$,

where $y_1 = x_1^2 + 2x_2^2 - 3x_1x_3$, and

$y_2 = 3x_1x_2^2 - x_2^2 + 4x_2x_3^2$.

Then $\frac{\partial y}{\partial x} = \begin{bmatrix} 2x_1 - 3x_3 & 3x_2^2 \\ 4x_2 & 6x_1x_2 - 2x_2 + 4x_3^2 \\ -3x_1 & 8x_2x_3 \end{bmatrix}$

9.7.3 Vector Differentiation of a Quadratic Form

An expression of the form $\mathbf{x^TAx}$, where **a** is an n x n symmetric matrix, is a quadratic form. The quadratic form $\mathbf{x^TAx}$ can be expanded as follows

$\mathbf{x^TAx}$ = $[x_1, x_2, ..., x_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

=

$a_{11}x_1^2+a_{12}x_1x_2+2a_{13}x_1x_3+\ldots+2a_{1n}x_1x_n+a_{22}x_2^2+2a_{23}x_2x_3+\ldots+2a_{2n}x_2x_n+\ldots+a_{nn}x_n^2$.

Taking partial derivatives with respect to the elements of **x**

$$\frac{\partial}{\partial x_1}(\mathrm{x^TAx}) = 2(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)$$

$$\frac{\partial}{\partial x_2}(\mathrm{x^TAx}) = 2(a_{12}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n)$$

$$\frac{\partial}{\partial x_n}(\mathrm{x^TAx}) = 2(a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{nn}x_n)$$

Apart from the factor 2, the right-hand side of the set of equations contains the elements of **Ax**, which is an n dimensional column vector. Alternatively, the right-hand side contains the elements of $\mathbf{x^T A}$, which is an n dimensional row vector. Thus, $\frac{\partial}{\partial x}(X^TAX) = 2AX$ and also, $\frac{\partial}{\partial x}(X^TAX) = 2X^TA$.

In practice, the choice between these two forms usually depends on the context in which differentiation occurs, since matrices can be equated only if they are of the same order and, specifically, a row vector cannot be set equal to a column vector.

9.7.4 Vector Differentiation of a Bi-linear Form

An expression f the form $\mathbf{x^T Bz}$, where $\mathbf{x^T}$is l × m, **B** is m × n, and **z** is n × l, is a bilinear form. The following derivatives for a bilinear form can be verified using the procedures above.

$\frac{\partial}{\partial \mathrm{x}}(\mathrm{x^T Bz}) = \mathrm{Bz}$ and $\frac{\partial}{\partial \mathrm{z}}(\mathrm{x^T Bz}) = \mathrm{B^T x}$.

Check Your Progress -2

1) If $\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\mathrm{A} = \begin{bmatrix} 3 & 1 & -2 \\ 1 & 0 & 3 \\ -2 & 3 & 2 \end{bmatrix}$ then show that $\frac{\partial}{\partial \mathrm{x}}(\mathrm{x^T Ax}) = 2\mathrm{xA^T}$

.....

.....

.....

2) If $\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\mathrm{B} = \begin{bmatrix} 2b_1 & b_1 + b_2 \\ b_2 + 2b_3 & 3b_2 \\ b_1 + b_3 & 4b_3 \end{bmatrix}$ $\mathrm{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$

then $\frac{\partial}{\partial \mathrm{x}}(\mathrm{x^T Bz}) = \mathrm{Bz}$ and $\frac{\partial}{\partial \mathrm{z}}(\mathrm{x^T Bz}) = B^T\mathrm{x}$

.....

.....

.....

9.8 SUMMERY

In this unit, first we discussed how to present a system of linear equations as a matrix equation. Then, we discussed the concepts of characteristic matrix, characteristic vector; characteristic root, eigen vector, eigen value etc. and explained the characteristic value problem. Next, this unit discussed the concept of quadratic form; and briefly discussed its various types. Next, we discussed the relation of eigen values to definiteness of a quadratic form.

Further, this unit introduced the concept of vector differentiation, and mentioned the processes of various types of vector differentiation, including that of a linear function, quadratic forms and bilinear forms.

9.9 ANSWER/HINT TO CHECK YOUR PROGRESS

Answer to Check Your Progress -1

1) See the text and answer
2) **[**Hint: $\mathbf{a^T x} = 2ax_1 - ax_2 + 3ax_3 + ax_4$, then find $\frac{\partial}{\partial x_1}(a^{\mathbf{T}}x) = 2a, \frac{\partial}{\partial x_2}(a^{\mathbf{T}}x) = -a, \frac{\partial}{\partial x_3}(a^{\mathbf{T}}x) = 3a, and\ \frac{\partial}{\partial x_4}(a^{\mathbf{T}}x) = a,$]
3) We can write (1) as:
$x^2 + 2xy - 3zx + 2xy + 2y^2 + 4yz - 3zx + 4yz - 7z^2$

$= (x+2y-3z)x+(2x+2y+4z)y+(-3x+4y-7z)z$

$= [x+2y-3z \quad 2x+2y+4z \quad -3x+4y-7z] \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$= [x \quad y \quad z] \begin{bmatrix} 1 & 2 & -3 \\ 2 & 2 & 4 \\ -3 & 2 & -7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

4)

i. $[x\ y] \begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

ii. $[x\ y] \begin{bmatrix} 5 & 10 \\ 10 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

iii. $[x\ y] \begin{bmatrix} 1 & 6 \\ 6 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

5)

i. $[x \quad y \quad z] \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

ii. $[x \quad y \quad z] \begin{bmatrix} 1 & -3 & 10 \\ 3 & 2 & 0 \\ 10 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

iii. $[x \quad y \quad z] \begin{bmatrix} 3 & 2 & 3 \\ 2 & 3 & 5 \\ 3 & 5 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

Answer to Check Your Progress -2

1) **[**Hint: $\mathbf{x^T Ax} = 3x_1^2 + 2x_3^2 + 2x_1x_2 - 4x_1x_3 + 6x_2x_3$ then find $\frac{\partial}{\partial x_1}(x^{\mathbf{T}}Ax) = 6x_1 + 2x_2 - 4x_3, \frac{\partial}{\partial x_2}(x^{\mathbf{T}}Ax) = 2x_1 + 6x_3$ and $\frac{\partial}{\partial x_3}(x^{\mathbf{T}}Ax) = -4x_1 + 6x_2 + 4x_3$]
2) Follow the same logic as given above.

9.10 REFERENCES AND FURTHER READINGS

1. Mathematics Class XII Part – 2 By Ravi Prakash , GRB Books, Second Edition 2025

ignou
THE PEOPLE'S
UNIVERSITY