Basic Linear Algebra in Machine Learning Techniques

Machine learning is nothing but value added applied Mathematics and Statistics.

A n- dimensional column vector x and its transpose x^T (an n-dimensional row vector) can be written as

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} ; \quad x^T = [x_1, x_2 \cdots x_n]$$

x_j ; $j = 1, \cdots, n$, are the elements of the vector.

We denote the $m \times n$ (rectangular) matrix A and its $n \times m$ transpose A^T as,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} ; \quad A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

A has m rows and n columns; a_{ij} denotes (i,j)th element, i.e, the element located in ith row and jth column.

Vectors may also be viewed as rectangular matrices. x is thus an $n \times 1$ matrix, and x^T is $1 \times n$ matrix.

Note that, we have used lower case italic letters for scalars, lower case bold non-italic letters for vectors, and upper case bold non-italic letters for matrices.

When $m = n$ i.e when the number of columns is equal to the number of rows, the matrix is said to be a square matrix of order n. A square matrix is called symmetric when its entries obey $a_{ij} = a_{ji}$.

Matrix can be expressed in terms of its column/rows.

for Example, a square matrix

$$A = [a_1 \; a_2 \; \cdots \; a_n];$$

$$a_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$$ of j^{th} column in A.

$$A = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

$$\alpha_i = [\alpha_{i1} \; \alpha_{i2} \; \cdots \; \alpha_{in}]$$ of i^{th} rows in A

A diagonal matrix is a square matrix whose elements off the principal diagonal are all zeros.

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{mn} \end{bmatrix}$$

A particular important matrix is the identity matrix I - an $n \times n$ (square) diagonal matrix whose principal diagonal entries are all 1's, and all other entries zeros

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

A null matrix 0 is a matrix whose elements are all equal to zero.

$$O = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Some Properties of Transpose

(i) $(A^T)^T = A$

(ii) $(KA)^T = KA^T$; K is a scalar

(iii) $(A+B)^T = A^T + B^T$

(iv) $(AB)^T = B^T A^T$

(v) For any matrix A, A^TA and AA^T ~~are~~ both Symmetric

(vi) When a square matrix A is symmetric, $A = A^T$

Addition of vectors ~~and~~ of matrices is component by component

The product AB of an $m \times n$ matrix A by an ~~[illegible]~~ $n \times p$ matrix B (number of columns of A must be equal to number of rows of B) is an $m \times p$ matrix C.

$$C = AB \text{ or } c_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj} \; ; \; i = 1 \cdots m \; ; \; j = 1 \cdots p$$

$AB \neq BA$; $(AB)C = A(BC)$; $(A+B)C = AC + BC$

matrix follows Associative Property

Matrix does not follow Commutative Property

Determinant of a Matrix :→

Determinants are defined for square matrices only. The determinant of an $n \times n$ matrix A, written as $|A|$, is a scalar-valued function of A. if we have 1×1 matrix A Then $|A| = A$ itself.

If A is 2×2 matrix, Then $|A| = a_{11}a_{22} - a_{21}a_{12}$.
The determinant of a ~~square~~ general square matrix can be computed by expansion by Minors (M_{ij}).
Cofactor C_{ij} of The element a_{ij} is defined by the eqn :-

$$C_{ij} = (-1)^{i+j} m_{ij}$$

where m_{ij} is a Minor, for an $n \times n$ Matrix A, we define Minor m_{ij} to be the $(n-1) \times (n-1)$ matrix obtained by deleting i^{th} Row and j^{th} column of A.

Some Properties of determinants are

(i) $|AB| = |A||B|$

(ii) $|A^T| = |A|$

(iii) $|KA| = K^n |A|$; A is an $n \times n$ matrix and K is a scalar

A square matrix is called singular if The associated determinant is zero ; it is called nonsingular if associated determinant is non zero.

The rank $\rho(A)$ of a Matrix A is The ~~dim~~ dimension of The largest array in A with a non zero determinant.

Adjoint of a Matrix

Let $A = [a_{ij}]$ be a square matrix of order n. The adjoint of matrix A is the transpose of the cofactor of A. It is denoted by adj A. An adjoint matrix is also called adjugate matrix.

Example:-

Find the adjoint of the matrix

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

To find the adjoint of a matrix, first find the cofactor matrix of the given matrix. Then find the transpose of the co-factor matrix.

$$\text{Cofactor of } 3 = A_{ij} = \begin{vmatrix} -2 & 0 \\ 2 & -1 \end{vmatrix} = 2$$

$$\text{Cofactor of } 1 = A_{12} = - \begin{vmatrix} 2 & 0 \\ 1 & -1 \end{vmatrix} = 2$$

$$\text{Cofactor of } -1 = A_{13} = \begin{vmatrix} 2 & -2 \\ 1 & 2 \end{vmatrix} = 6$$

$$\text{Cofactor of } 2 = A_{21} = - \begin{vmatrix} 1 & -1 \\ 2 & -1 \end{vmatrix} = -1$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\text{Cofactor of } -1 \quad A_{33} = \begin{vmatrix} 3 & 1 \\ 2 & -2 \end{vmatrix} = -8$$

$$\text{The cofactor matrix of A is } |A_{ij}| = \begin{bmatrix} 2 & 2 & 6 \\ -1 & -2 & -5 \\ -2 & -2 & -8 \end{bmatrix}$$

Now find the transpose of A_{ij}

$$\operatorname{adj} A = (A_{ij})^T = \begin{bmatrix} 2 & -1 & -2 \\ 2 & -2 & -2 \\ 6 & -5 & -8 \end{bmatrix}$$

Inverse of Matrix :-

The inverse of an $n \times n$ matrix A, denoted by A^{-1}, is the $n \times n$ matrix such that

$$AA^{-1} = A^{-1}A = I$$

We can write the inverse of Matrix $A = \frac{adj(A)}{|A|}$

Some Properties of Matrix Inverse are

1. $(A^{-1}) = A$
2. $(A^T)^{-1} = (A^{-1})^T$
3. $(AB)^{-1} = B^{-1}A^{-1}$
4. $|A^{-1}| = \frac{1}{|A|}$
5. $|P^{-1}AP| = |A|$

* We cannot find A^{-1} for the non-square Matrix
* We cannot find A^{-1} for the square Matrix whose determinant is 0 (it means singular matrix)

Orthogonal Matrix :-

Suppose A is a square matrix with real elements and $n \times n$ order and A^T is the Transpose of A. Then according to the definition, if $A^T = A^{-1}$ is satisfied, Then

$$AA^T = I$$

where I is the identity matrix.

Characteristic Roots and vectors (or eigenvalues and eigenvectors)

Let A be a square matrix of order n, λ is a scalar and

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$ a column vector,

consider the equation

$$AX = \lambda X \qquad (1)$$

Clearly, $X = 0$ is the solution of eq (1) for any value λ

Now, if I_n is unit matrix of order n then eq (1) may be written in the form

$$(A - \lambda I)X = 0 \qquad (2)$$

eq^n (2) is the matrix form of a system of n homogeneous linear eq^n in n unknowns. This system will have non-trivial solution if and only if the determinant of the coefficient matrix $A - \lambda I_n$ vanishes, i.e.

if $|A - \lambda I_n| = 0$

if $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{mn} \end{bmatrix}$

Then $\begin{bmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & & & \end{bmatrix} = 0$

The expansion of this determinant yields a polynomial of degree n in λ, ~~characteristic~~ called the characteristic polynomial of the matrix A.

The eqn $|A-\lambda I_n|=0$ is called the characteristic eqn.

The n^{th} roots of the characteristic eqn of matrix A of order n are called the characteristic roots, characteristic values, eigen values, ~~of~~ of the matrix A.

If λ is a eigen value of an $n\times n$ matrix A then a non-zero vector X such that

$$AX=\lambda X$$

is called characteristic vector, eigen vector of the matrix A corresponding to the characteristic root λ.

Ex:- find the eigen value and eigen vectors of the matrix

$$A=\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$$

Soln:- The characteristic eqn of the matrix A is

$$|A-\lambda I|=0 \quad ; \quad \begin{vmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{vmatrix}=0$$

or $$(8-\lambda)\{(7-\lambda)(3-\lambda)-16\}+6\{-6(3-\lambda)+8\}$$
$$+2\{24-2(7-\lambda)\}=0$$

$\Rightarrow$ $\lambda^3-18\lambda^2+45\lambda=0$

$\Rightarrow$ $\lambda(\lambda-3)(\lambda-15)=0$

$$\boxed{\lambda=0,3,15}$$

The eigen vector of A corresponding to the eigen value 0 is given by

$(A - 0I)X = 0$

$$\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $\begin{bmatrix} -2 & -4 & 3 \\ -6 & 7 & -4 \\ 8 & -6 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, by $R_1 \leftrightarrow R_3$

or $\begin{bmatrix} 2 & -4 & 3 \\ 0 & -5 & 5 \\ 0 & 10 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ by $R_2 \to R_2 + 3R_1$
$R_3 \to R_3 - 4R_1$

or $\begin{bmatrix} 2 & -4 & 3 \\ 0 & -5 & 5 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ by $R_3 \to R_3 + 2R_2$

The coefficient matrix is of rank 2 (number of rows e)

Therefore, these eqn's have $n - r = 3 - 2 = 1$ linearly independent solutions.

The above eqn are

$$2x_1 - 4x_2 + 3x_3 = 0$$
$$-5x_2 + 5x_3 = 0$$

from the last eqn, we get $\boxed{x_2 = x_3}$

choose $x_2 = K$, $x_3 = K$

then the first eqn gives $x_1 = \frac{K}{2}$, where K is any scalar

$\therefore X_1 = K\left[\frac{1}{2}, 1, 1\right]$

$X = [1, 2, 2]'$ is an eigen vector of A corresponding to eigen value 0.

The eigenvector of A corresponding to the eigenvalue 3 are given by $(A-3I)X=0$

or $\begin{bmatrix} 5 & -6 & 2 \\ -6 & 4 & -4 \\ 2 & -4 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

or $\begin{bmatrix} -1 & -2 & -2 \\ -6 & 4 & -4 \\ 2 & 4 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, by $R_1 \to R_1 + R_2$

or $\begin{bmatrix} -1 & -2 & -2 \\ 0 & 16 & 8 \\ 0 & -8 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, by $R_2 \to R_2 - 6R_1$, $R_3 \to R_3 + 2R_1$

or $\begin{bmatrix} -1 & -2 & -2 \\ 0 & 16 & 8 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$,

by $R_3 \to R_3 + \frac{1}{2}R_2$

The coefficient matrix of these eqn is of rank 2. Therefore, This eqn is of rank 2.

Therefore, These eqns have $n-r = 3-2 = 1$ linearly Independent solutions.

The above eqn are

$$-x_1 - 2x_2 - 2x_3 = 0$$
$$16x_2 + 8x_3 = 0$$

from The last eqn, we set $x_2 = -\frac{1}{2}x_3$

Choose $x_3 = 4K$, $x_2 = -2K$; Then from The first eqn $X_1 = -4K$, where K is any scalar

$\therefore$ $X_2 = K[-4, -2, 4]'$ is an eigenvector of A corresponding to the eigen value 3

Now, The eigenvector of A corresponding to The eigen value 15 is given by $(A-15I)X=0$

or $\begin{bmatrix} -7 & -6 & 2 \\ -6 & -8 & -4 \\ 2 & -4 & -12 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

or $\begin{bmatrix} -1 & 2 & 6 \\ -6 & -8 & -4 \\ 2 & -4 & -12 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, by $R_1 \to R_1 - R_2$

or $\begin{bmatrix} -1 & 2 & 6 \\ 0 & -20 & -40 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, by $R_2 \to R_2 - 6R_1$
$R_3 \to R_3 + 2R_1$

The coefficient matrix of These eqn is of rank 2.

$\therefore$ These eqn have $n-r = 3-2 = 1$ linearly independent solution.

The above eqn are

$-x_1 + 2x_2 + 6x_3 = 0$

$-20x_2 - 40x_3 = 0$

from The last eqn we get $x_2 = -2x_3$

Choose $x_3 = k$, $x_2 = -2k$

Then from The first eqn, we get $x_1 = 2k$, where k is any scalar

$\therefore$ $x_3 = k[2, -2, 1]'$ is an eigen vector of A corresponding to the eigen value $= 15$.