LECTURE NOTES – 28-01-2018

MATRICES

$$\begin{bmatrix} a_{11} & a_{12} & - & a_{1n} \\ a_{21} & - & - & - \\ - & - & - & - \\ a_{m1} & a_{m2} & - & a_{mn} \end{bmatrix}$$

Each row has **n** elements. Each column has **m** elements.
Every element is identified uniquely by indices, the first index indicates the row and the second index is the column.
An array with dimensions $m \times n$ is denoted as $A_{m \times n}$.
If either **m** or **n** is **1**, we call it as a **vector**.

$$1 \times n \ \rightarrow row\ vector$$
$$m \times 1 \ \rightarrow column\ vector$$

Matrices are denoted by upper case letters, ex – A, B. Vectors are denoted by lower case letters, ex – *v*, *w*.

If A and B are matrices, they are said to be equal iff every element of A is equal to every element of B.

$$A = B \ \ iff \ \ A_{ij} \ = B_{ij} \ \forall \begin{matrix} 1 \le i \le m \\ 1 \le j \le n \end{matrix}$$

Any higher dimensional matrix is known as **<u>tensor</u>**.

A matrix is a **square matrix** if $m = n$, i.e., no. of rows equals the no. of columns.
A general matrix, where $m \neq n$, is known as a **rectangular matrix**.

Let's take an example.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 8 & 9 & 3 \\ 0 & 1 & 4 \end{bmatrix}$$

A is a 3 X 3 square matrix.
The **transpose** of matrix A is given as A^T. The transpose of a matrix is when we interchange the rows and the columns. So, for the above example, A^T will be:

$$A^T = \begin{bmatrix} 1 & 8 & 0 \\ 2 & 9 & 1 \\ 3 & 3 & 4 \end{bmatrix}$$

We can define transpose as:

$$a_{ij}^T = a_{ji} \; \forall \, i, j$$

Transpose is applicable to all types of matrices, not just square matrices.

$$Ax = b$$

For the above system of equations, A is the coefficient matrix, and x and b are generally row vectors. As per convention, we should be using the transpose of the row vectors:

$$Ax^T = b^T$$

Algebra of Matrices

Addition of Matrices

$$C = A + B$$
$$C_{ij} = A_{ij} + B_{ij} \; \forall \; i, j$$

The assumption here is that A and B are of the same dimensions, and so is C. **m** and **n** are the same for all. If the dimensions are not the same, the operation is not defined.

Addition is *associative* for matrices.

$$(A + B) + C = A + (B + C)$$

Scalar Multiplication

If **s** is a scalar and **A** is a matrix, scalar multiplication of s X A is defined as every element of A being multiplied into **s**.

$$sA = sA_{ij} \;\; \forall \;\; i, j$$

If we multiply A by (-1), we get

$$(-1)A = -A$$

Subtraction of Matrices

Subtraction of matrices is just an extension of scalar multiplication and addition.

$$A - B = A + (-1)B$$

Null Matrix

If we subtract A with itself, we get null matrix.

$$A - A = 0$$

Similarly if we multiply any matrix with scalar 0, we get a null matrix.

$$0. A = 0$$

Null matrix, also known as **zero matrix** is a matrix with all elements as zero.
Note: Two null matrices with different dimensions are not the same.

A square matrix is said to be **symmetric** if the value of components are identical along its diagonal. If A is the same as its transpose, it is said to be a symmetric matrix.

$$A = A^T \quad \rightarrow symmetric$$
$$A = -A^T \quad \rightarrow skew\ symmetric$$

In a square matrix, if we have all elements except diagonal elements as 0, we call it as **diagonal matrix**.

$$a_{ij} = 0 \ \forall \ i \neq j$$

A diagonal matrix where all diagonal elements have the value 1, is known as **identity matrix**.

$$a_{ij} = \begin{cases} 0 \ \forall \ i \neq j \\ 1 \ \forall \ i = j \end{cases}$$

Essentially, a diagonal matrix is symmetric because it is a square matrix and the transpose is the same as itself. Similarly, since identity matrix is also a diagonal matrix, it is symmetric.

Trace of a matrix is defined as the sum of all elements of a diagonal.

$$Trace = \sum a_{ij} \ \ \forall \ i = j$$

Trace is defined for any square matrix.

There are two types of diagonals – **leading diagonal** (from top left corner to bottom right corner) and **trailing diagonal** (from top right corner to bottom left corner). For all purposes and matrix operations, we use the term diagonal to refer to the leading diagonal.

Triangular matrix

If all elements below leading diagonal are 0, it is known as the *upper* triangular matrix.
If all elements above leading diagonal are 0, it is known as the *lower* triangular matrix.

$$a_{ij} = \begin{cases} 0 \;\forall\; i < j \\ \quad or \\ 0 \;\forall\; i > j \end{cases}$$

We can have a **matrix function** where each element of the matrix is a function of another variable.
Example:

$$A(t) = \begin{bmatrix} t^2 & 0 & (1-t) \\ \sin(t) & e^{2t} & 5 \end{bmatrix}$$

$\mathbf{I}$ is always used to denote the identity matrix.

We can differentiate and integrate matrix functions.

$$\mathrm{f(t)} = \begin{bmatrix} \mathrm{t}^2 & \ln(\mathrm{t}) & (1-\mathrm{t}) \\ \sin(\mathrm{t}) & \mathrm{e}^{2\mathrm{t}} & 5 \end{bmatrix}$$

In the above matrix function, since *ln(t)* is not defined at *t = 0*, hence the entire function is not defined at *t = 0*.

Multiplication of matrices

Two matrices can only be multiplied if the number of columns of the first matrix is equal to the number of rows of the second matrix.

$$C = A_{m\times k} . B_{k\times n}$$
$$c_{ij} = \sum_k a_{ik} . b_{kj}$$

Resulting matrix has the dimensions m X n.
Example 1:

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

A is 2X2 and B is 2 X 3. No. of columns of A = No. of rows of B, hence we can multiply A and B.

$$A \times B = \begin{bmatrix} 1 \times 2 + 3 \times 1 & 1 \times (-2) + 3 \times 5 & 1 \times 0 + 3 \times 3 \\ 2 \times 2 + 4 \times 1 & 2 \times (-2) + 4 \times 5 & 2 \times 0 + 4 \times 3 \end{bmatrix}$$
$$A \times B = \begin{bmatrix} 5 & 13 & 9 \\ 8 & 16 & 12 \end{bmatrix}$$

The result is a 2 X 3 matrix.

Example 2:

$$A = [1\ 2 - 1]$$
$$B = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix}$$
$$A \times B = [-2\ 1]$$

B X A is not feasible/not defined as number of columns of B (3) is not equal to number of rows of A (1).

Example 3:

$$A = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$$
$$B = \begin{bmatrix} 1 & 4 & -6 \end{bmatrix}$$
$$AB = \begin{bmatrix} 2 & 8 & -12 \\ -1 & -4 & 6 \\ 3 & 12 & -18 \end{bmatrix}$$
$$BA = [-20]$$

Provided all the below equations involve feasible multiplication, we have the rules:

$$A(BC) = (AB)C \ \rightarrow Associative\ Property$$
$$A(B + C) = AB + AC \ \rightarrow Distributive\ Property$$
$$(A + B)C = AC + BC \ \rightarrow Distributive\ Property$$

Let us define the **inverse** of a matrix. Inverse of a matrix A = A^{-1}. It is defined as:

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

We have a system of equations:

$$Ax = b$$

Multiplying both sides by A^{-1},

$$A^{-1}Ax = A^{-1}b$$
$$Ix = A^{-1}b$$
$$x = A^{-1}b$$

We wanted to solve the vector of unknowns x. If we're able to find the inverse of A, and able to pre-multiply it to b, we get the solution of x. Motivation to find inverse is to find the solution for x.
There are mainly two techniques to solve these equations:

- Gauss Seidel
- Gauss Jordan

(Also read Gauss elimination).

If we multiply vector by a scalar, the property of the vector doesn't change, it just scales up/down.
If we have a matrix A = $\begin{bmatrix} 1 & 0 & 3 \\ 4 & 2 & 5 \\ 8 & 4 & 10 \end{bmatrix}$ such that:

$$\begin{bmatrix} 1 & 0 & 3 \\ 4 & 2 & 5 \\ 8 & 4 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Then, this matrix A is said to be singular, if by any linear transformation, we can convert at least one entire row or column to zeroes.

<u>Linear Transformation</u> -> add/subtract two rows or two columns, the nature of the matrix remains the same. Multiply a row or column by a scalar, nature of the matrix remains same. *Note*: whatever operations you do on A, you have to do the same on the matrix on the RHS (in the above case: the *b* vector).

<u>Singular matrices do not have a unique solution.</u>

In the above example, we can multiply second row of A with scalar 2 and subtract from the third row, we get one complete row of zeroes. Hence it has no unique solution.

"No unique solution" means either 0 solution (none) or infinitely many solutions. In case of infinitely many solutions, the solutions are generally linear transformations of each other (ex. One solution will be (1,2,3) and other solutions could be (2,4,6) or (3,6,9)).

Also Read Theory of Equations.

Norm

$$\|x_i\|_p = (\sum |x_i|^p)^{1/p}$$

The LHS is known as the **norm** of **order** <u>*p*</u>.
When order is 2, or p=2, it is known as the <u>Euclidian norm</u> – which is basically the Euclidian distance – distance of (x,y) from origin.

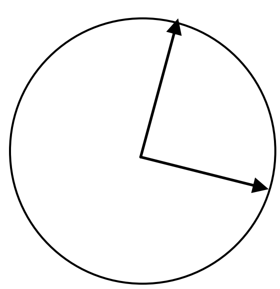
If $X^T Y = 0$, then X and Y are <u>**orthogonal**</u> only if they have non-zero norms.

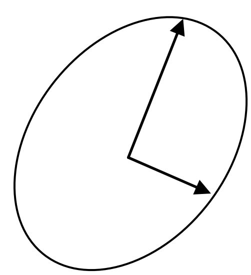
Eigen Decomposition

If **A** is a square matrix and ***v*** is a vector,

$$Av = \lambda v$$

If multiplying a vector by matrix results only in scaling of the vector, then ***v*** is known as <u>**eigenvector**</u> of **A** and **λ** is known as <u>**eigenvalue**</u>.
<u>Eigenvectors are always orthogonal.</u> There can be multiple eigenvectors for given matrix A, each eigenvector is associated with that eigenvalue **λ** and it stretches proportional to **λ**.





Geometrically, an eigenvector corresponding to a real non-zero eigenvalue points in a direction that is stretched by the transformation and the eigenvalue is the factor by which it is stretched.