Linear Algebra

Linear algebra is a branch of mathematics that is concerned with mathematical structures closed under the operations of addition and scalar multiplication and that includes the theory of systems of linear equations, matrices, determinants, vector spaces, and linear transformations

Why learn Linear Algebra for Machine Learning?

What do you see when you look at the image above? You most likely said flower, leaves -not too difficult. But, if I ask you to write that logic so that a computer can do the same for you – it will be a very difficult task (to say the least).
You were able to identify the flower because the human brain has gone through million years of evolution. We do not understand what goes in the background to be able to tell whether the colour in the picture is red or black. We have somehow trained our brains to automatically perform this task.

But making a computer do the same task is not an easy task, and is an active area of research in Machine Learning and Computer Science in general. But before we work on identifying attributes in an image, let us ponder over a particular question- How does a machine stores this image?

You probably know that computers of today are designed to process only 0 and 1. So how can an image such as above with multiple attributes like colour be stored in a computer? This is achieved by storing the pixel intensities in a construct called **Matrix.** Then, this matrix can be processed to identify colours etc.

So any operation which you want to perform on this image would likely use Linear Algebra and matrices at the back end.

Matrix

Matrix is a way of writing similar things together to handle and manipulate them as per our requirements easily. In Data Science, it is generally used to store information like weights in an Artificial Neural Network while training various algorithms. You will be able to understand my point by the end of this article.
Technically, a matrix is a 2-D array of numbers (as far as Data Science is concerned). For example look at the matrix A below.

1	2	3
4	5	6
7	8	9

Generally, rows are denoted by ‘i’ and column are denoted by ‘j’. The elements are indexed by ‘i’th row and ‘j’th column.We denote the matrix by some alphabet e.g. A and its elements by A(ij).
In above matrix
$A12 = 2$
To reach to the result, go along first row and reach to second column.

Basic operations on matrix

Let’s play with matrices and realise the capabilities of matrix operations.

- **Addition**:

Addition of matrices is almost similar to basic arithmetic addition. All you need is the order of all the matrices being added should be same. This point will become obvious once you will do matrix addition by yourself.
Suppose we have 2 matrices ‘A’ and ‘B’ and the resultant matrix after the addition is ‘C’.
Then
$Cij = Aij + Bij$
For example, let’s take two matrices and solve them.

A =

1	0
2	3

B =

4	-1
0	5

Then,

C =

5	-1
2	8

Observe that to get the elements of C matrix, I have added A and B element-wise i.e. 1 to 4, 3 to 5 and so on.

- **<u>Scalar Multiplication</u>**:

Multiplication of a matrix with a scalar constant is called scalar multiplication. All we have to do in a scalar multiplication is to multiply each element of the matrix with the given constant. Suppose we have a constant scalar ‘c’ and a matrix ‘A’. Then multiplying ‘c’ with ‘A’ gives-

c[Aij] = [c*Aij]

- **<u>Transposition</u>**:

Transposition simply means interchanging the row and column index. For example-

AijT= Aji

- **<u>Matrix multiplication:</u>**

Matrix multiplication is one of the most frequently used operations in linear algebra. We will learn to multiply two matrices as well as go through its important properties.

Before landing to algorithms, there are a few points to be kept in mind.

The multiplication of two matrices of orders i*j and j*k results into a matrix of order i*k. Just keep the outer indices in order to get the indices of the final matrix.
Two matrices will be compatible for multiplication only if the number of columns of the first matrix and the number of rows of the second one are same.
The third point is that order of multiplication matters.
Don’t worry if you can’t get these points. You will be able to understand by the end of this section.

Suppose, we are given two matrices A and B to multiply. I will write the final expression first and then will explain the steps.

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \text{x} \begin{bmatrix} 10 & 11 \\ 20 & 21 \\ 30 & 31 \end{bmatrix}$$

$$= \begin{bmatrix} 1\text{x}10 + 2\text{x}20 + 3\text{x}30 & 1\text{x}11 + 2\text{x}21 + 3\text{x}31 \\ 4\text{x}10 + 5\text{x}20 + 6\text{x}30 & 4\text{x}11 + 5\text{x}21 + 6\text{x}31 \end{bmatrix}$$

$$= \begin{bmatrix} 10+40+90 & 11+42+93 \\ 40+100+180 & 44+105+186 \end{bmatrix} = \begin{bmatrix} 140 & 146 \\ 320 & 335 \end{bmatrix}$$