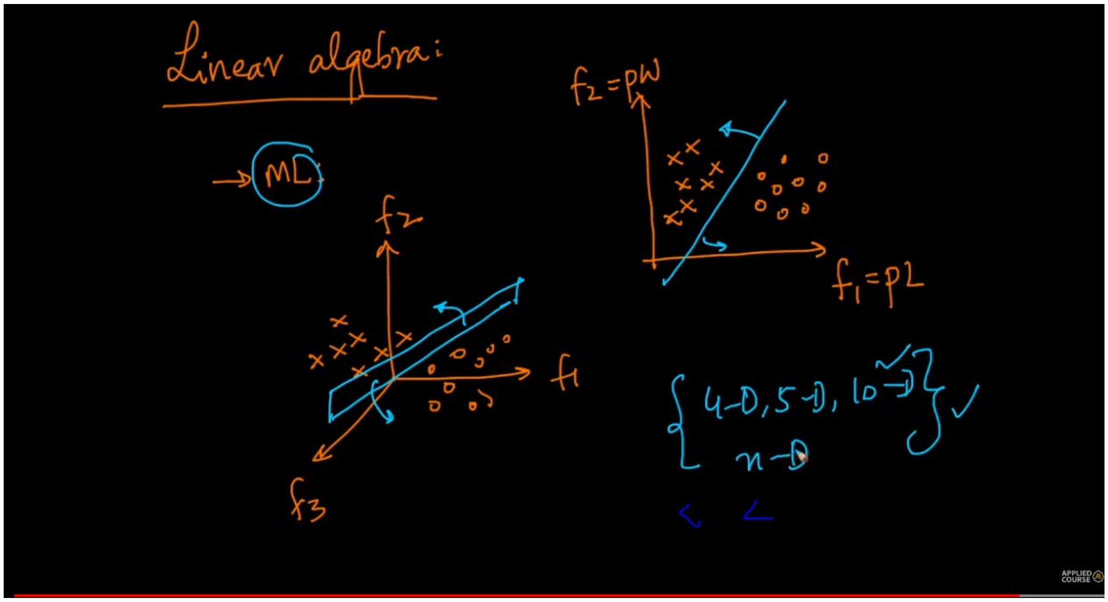
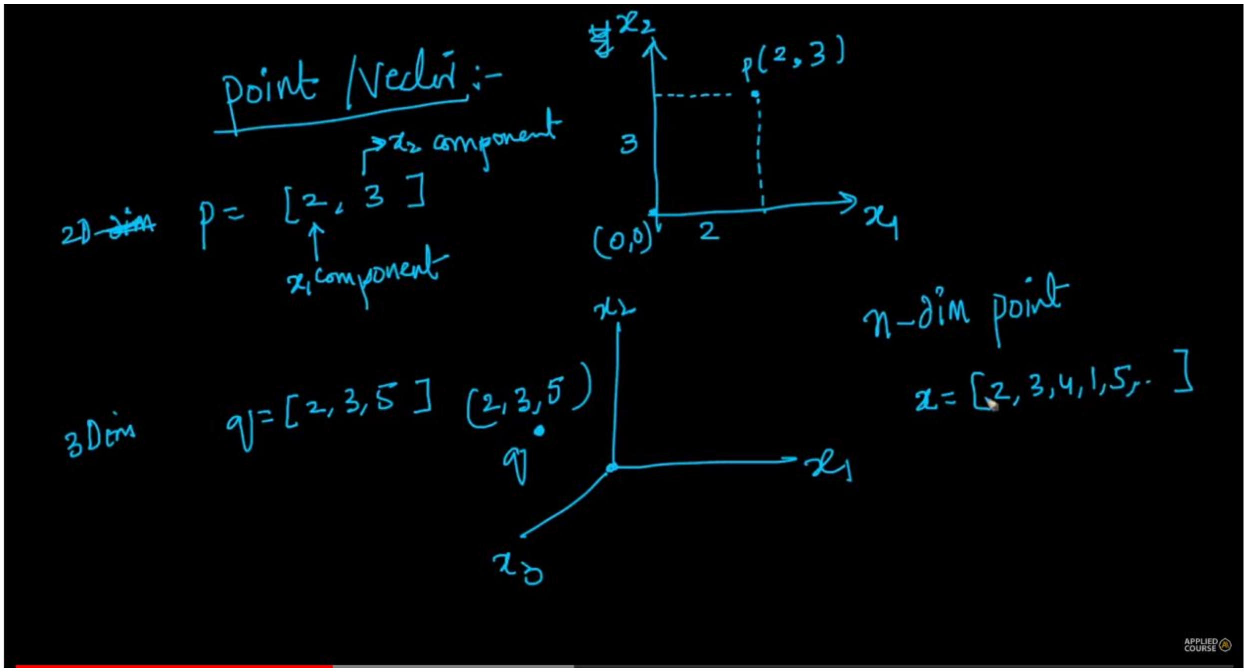
LINEAR ALGEBRA

Why Linear Algebra is used in Machine Learning?



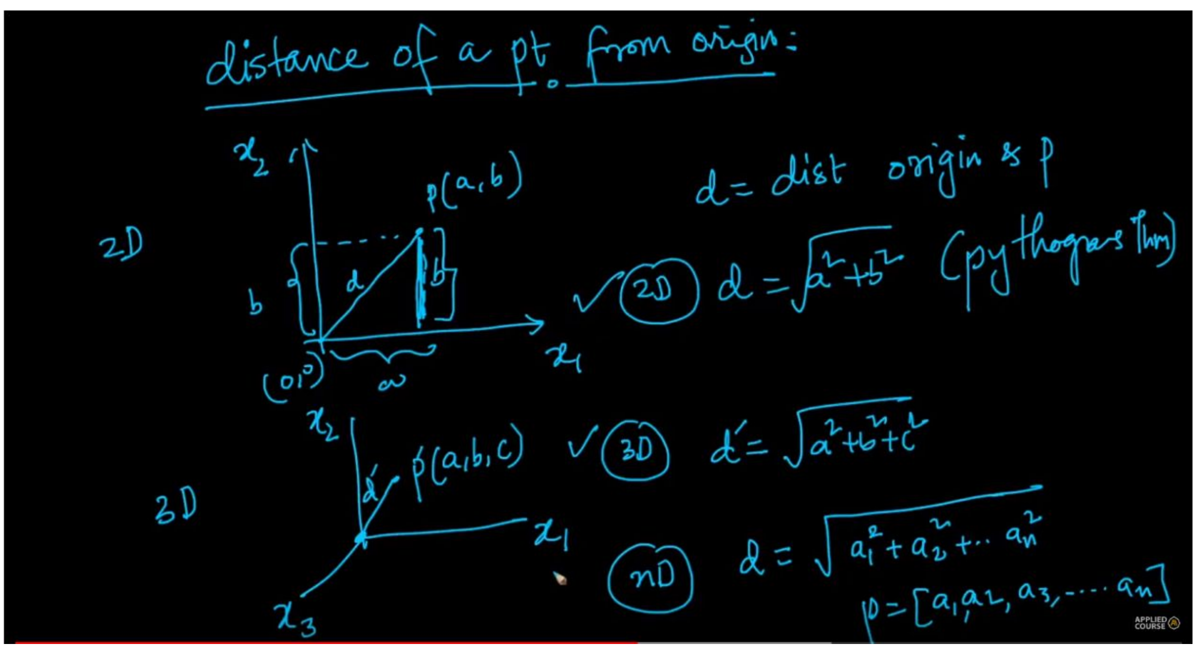
As we can see the line or the plane separates the data points perfectly w.r.t 2-D and 3-D but what if we have 10 dimensions or 1000 dimensions.
Linear Algebra provides us the necessary tools to operate in higher dimensions

Point/ Vector

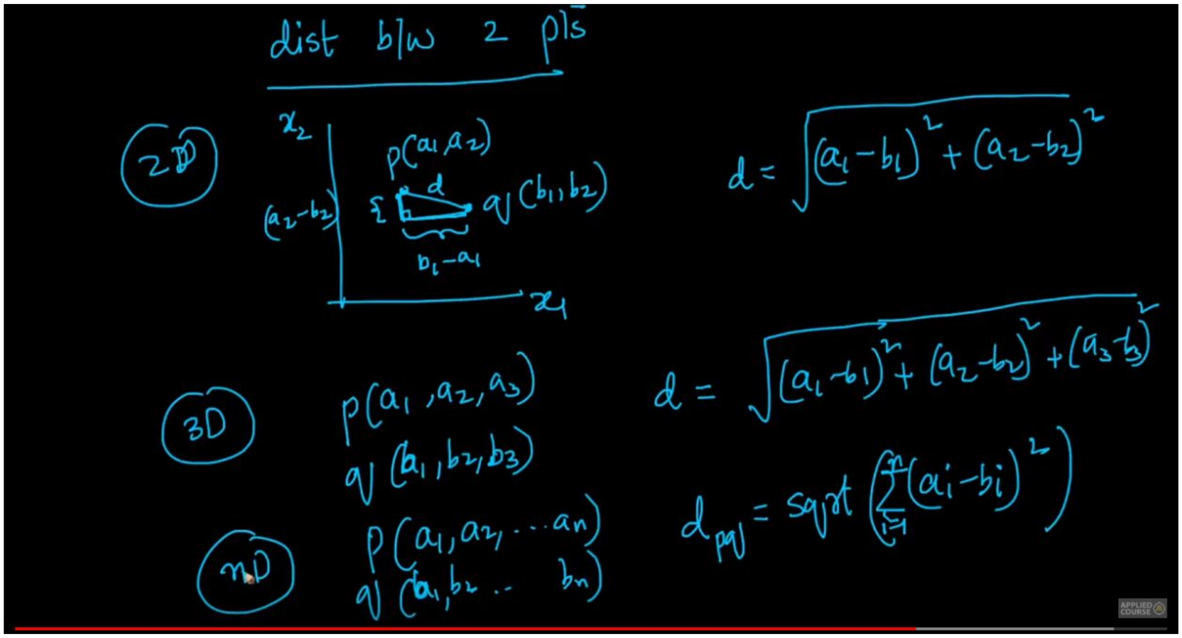


Self explanatory but suppose if we have n- points then we can represent it as x = [2,3,4,1,6…..n]

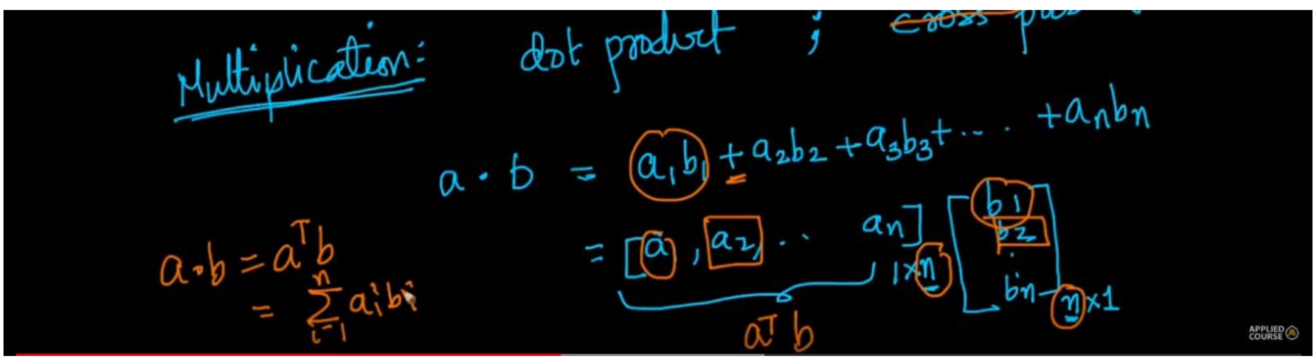
Distance of a point from origin



Distance between two points

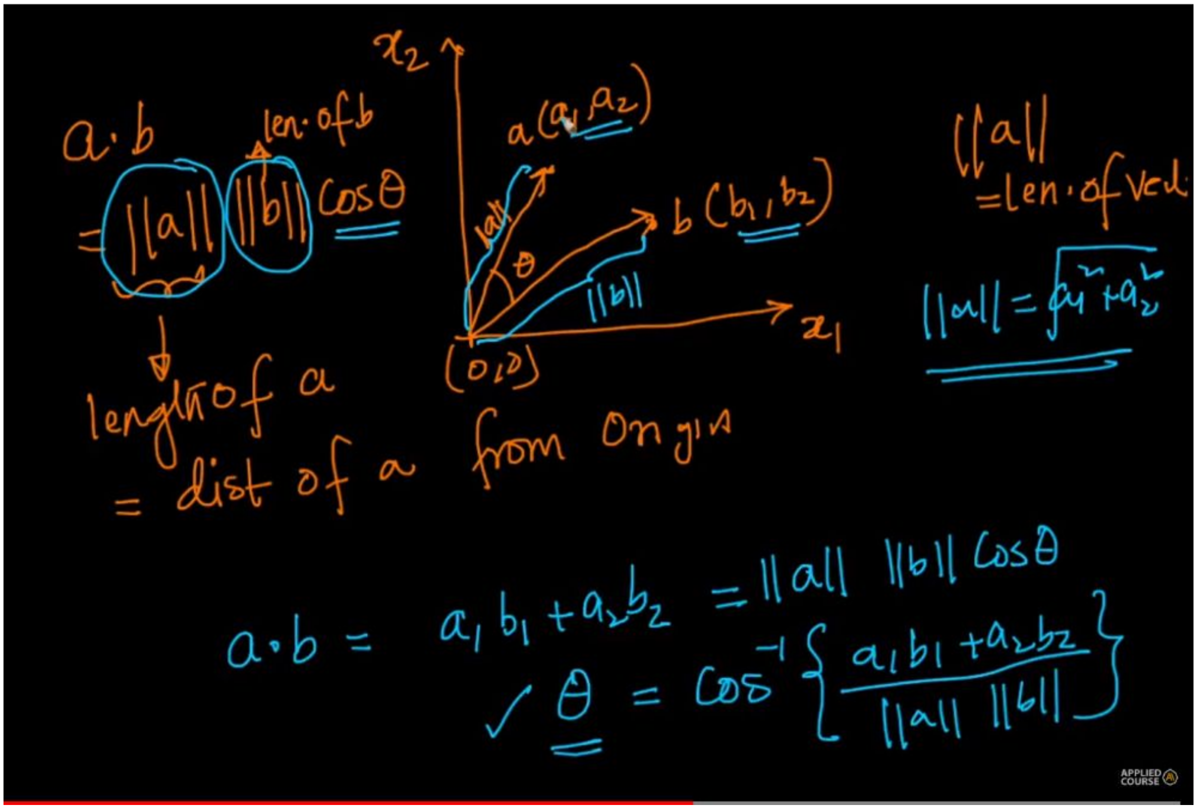


Dot product (Linear algebra point of view)



One question : Why are we summing up the products ? Will be answered in the geometrical point of view of dot product.

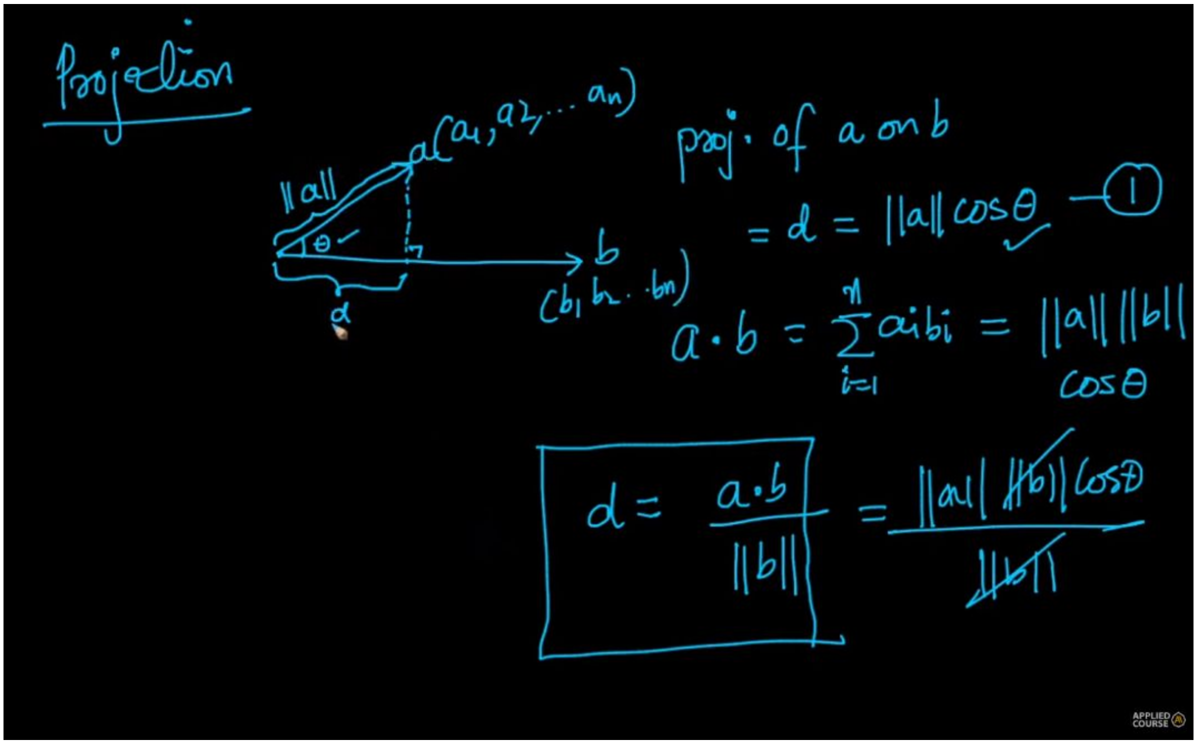
Dot Product (Geometric understanding)



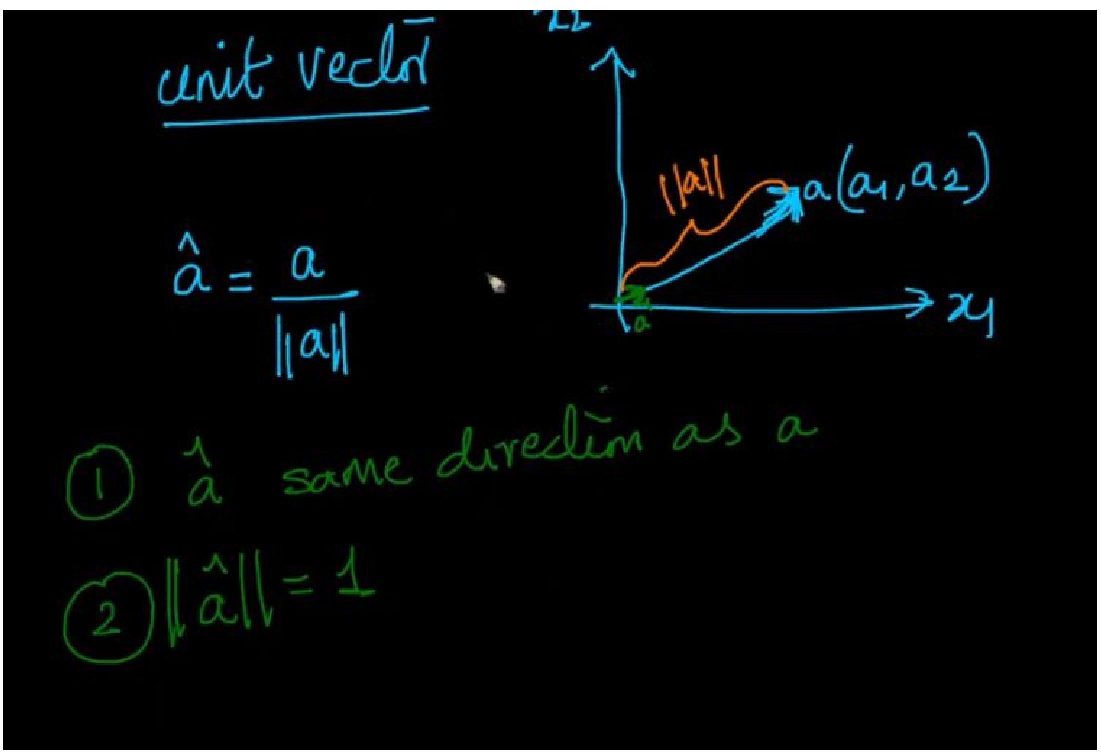
This theory can be applied to any dimensions of vectors.
Note : If the angle between two vectors is 90 then dot is 0 (cos 90 = 0)

For dot product intuition and the answer to above question
https://math.stackexchange.com/questions/348717/dot-product-intuition

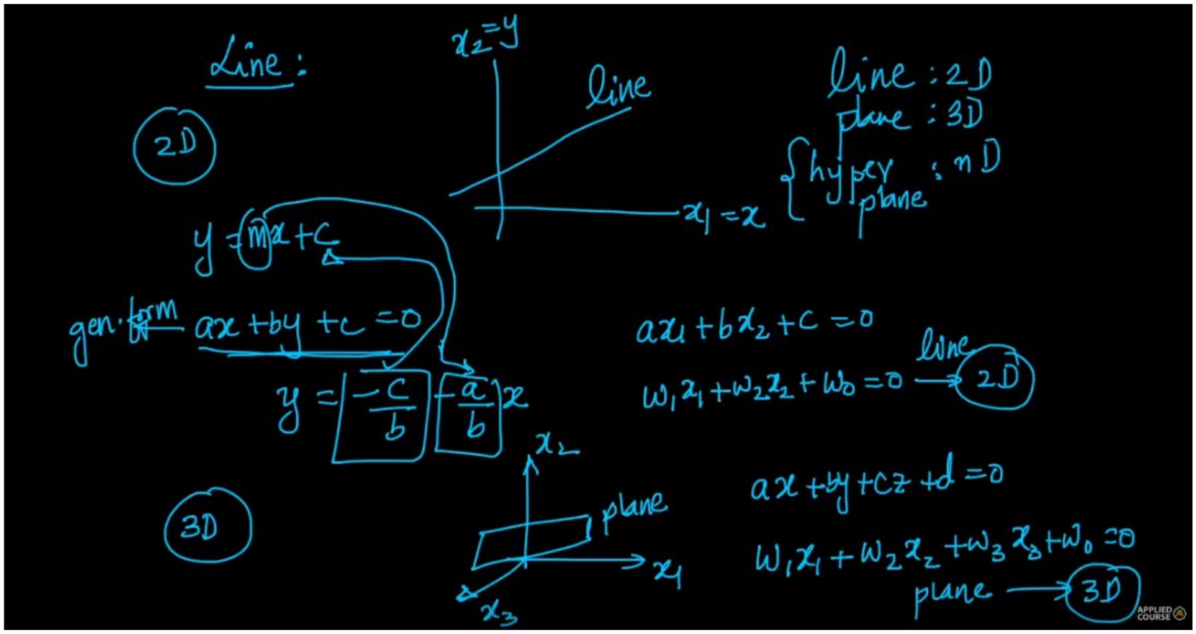
Projection



Unit Vector



Line and Plane (or for n-d array)



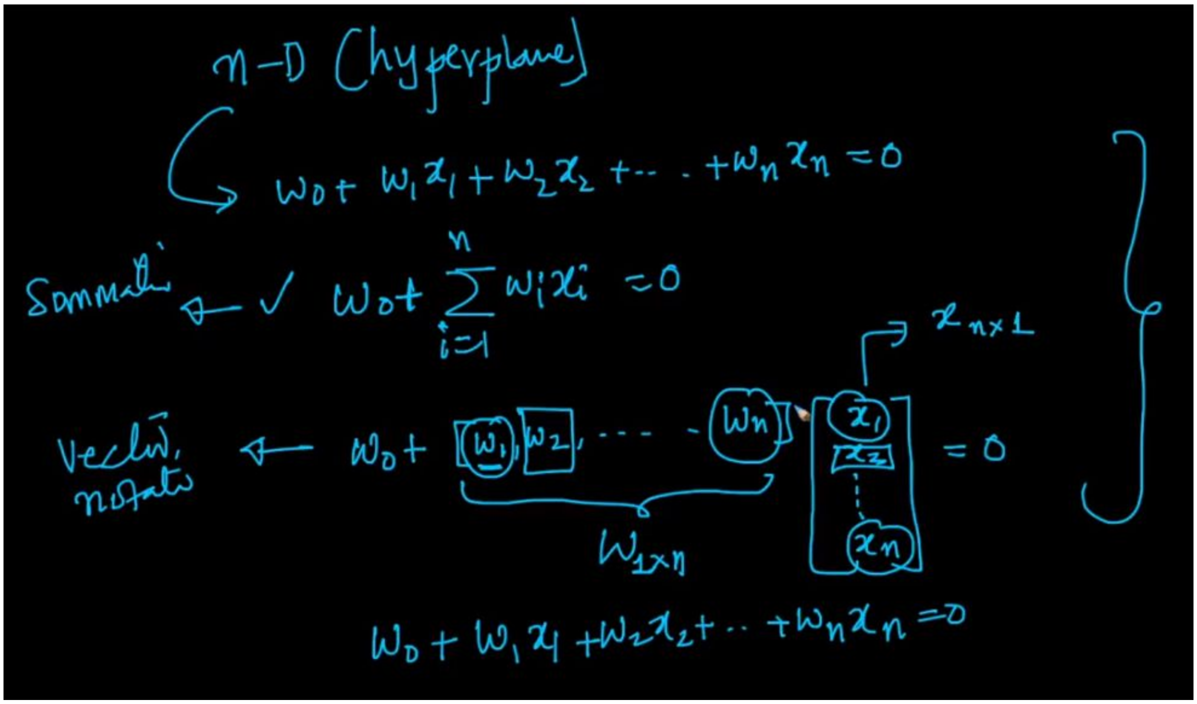
A line divides the space into two parts. Above it and below it. Same exists for plane in 3d. For n= dimensional vector it is hyperplane.

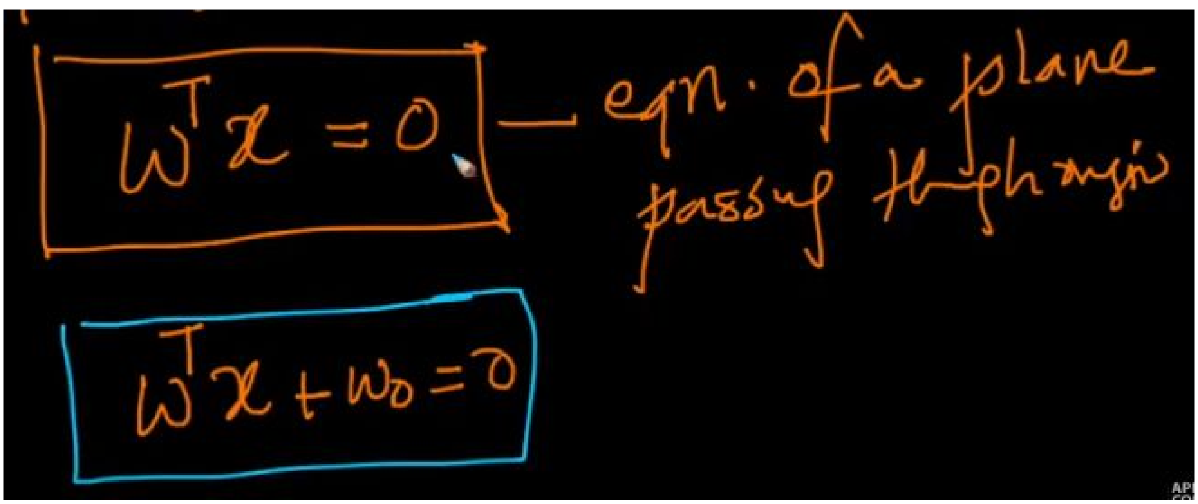
Gen eqn of line : ax + by + c = 0

w1x1 + w2x2 + w0 = 0 (Commonly used for representation)

General Equation of plane : ax + by + cz + d = 0

For N- d vectors

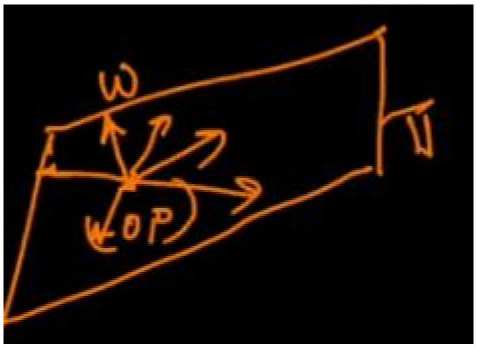




If both **W** and **x** are **both** (1*n) vectors then this equation . w0 is the intercept same as c in Equation of line : y = mx + c (where c is the y intercept)

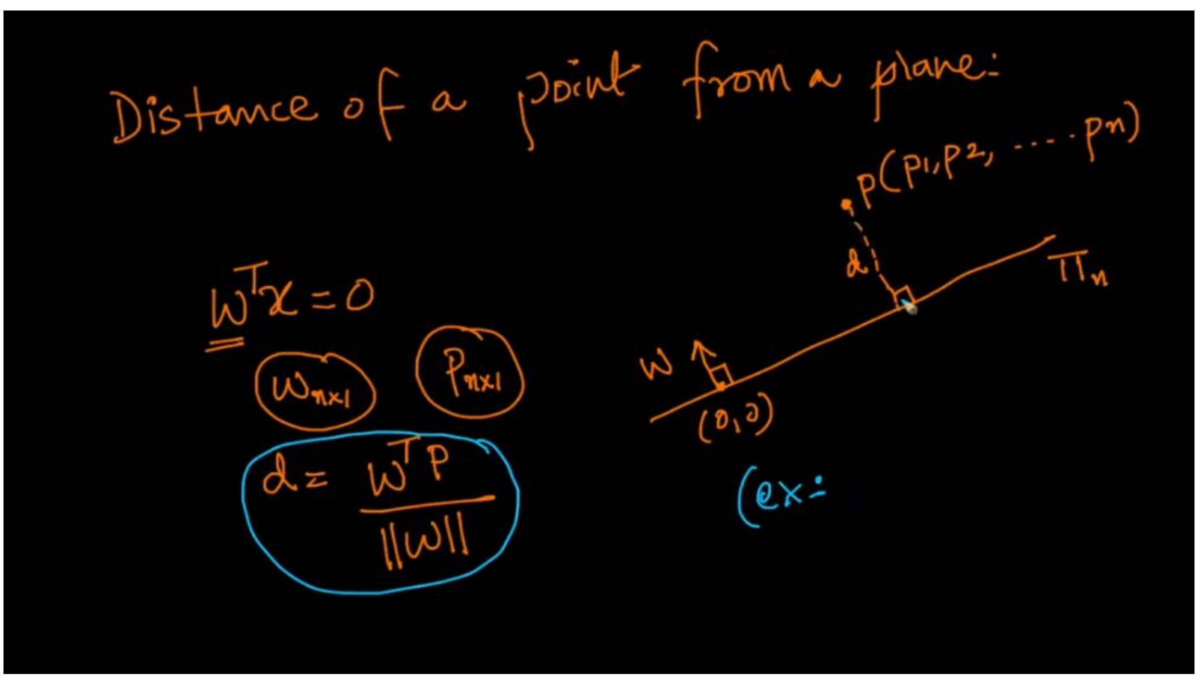
What is the significance of W in $W^T x = 0$?

We know that $W^T x = 0$ => ||W|| .||x|| cos Θ therefore theta is 90 which leads us to know that W is vector perpendicular to the plane passing through **origin**

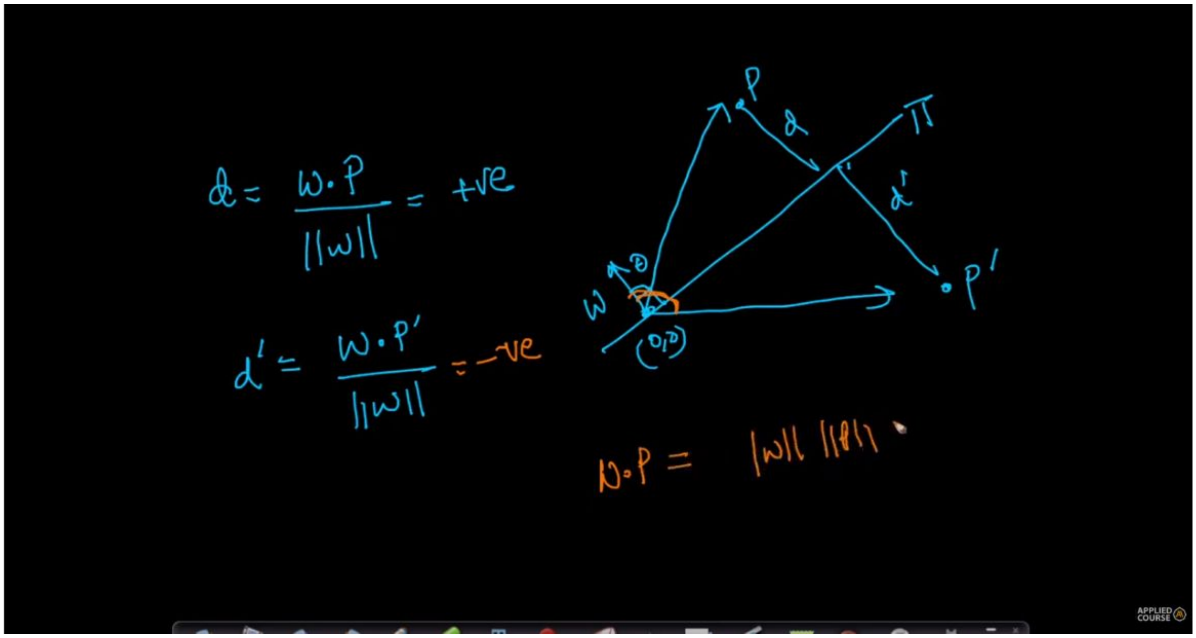


It also means that W is perpendicular to any vector in the plane which makes their dot product = 0 as well

Distance of a point from a plane



Let us cover what it can tell below



The angle between W and P vector is less than 90 ∘ it is +ve because cos is +ve between 0 and 90. The angle between W and P’ vector is more than 90 ∘ it is -ve because cos is -ve between 90 and 180. It can also be inferred that the distance points lying in half space of P will be positive and negative for P’. It can be used in ***classification.***

For circle