

2.1 Introduction



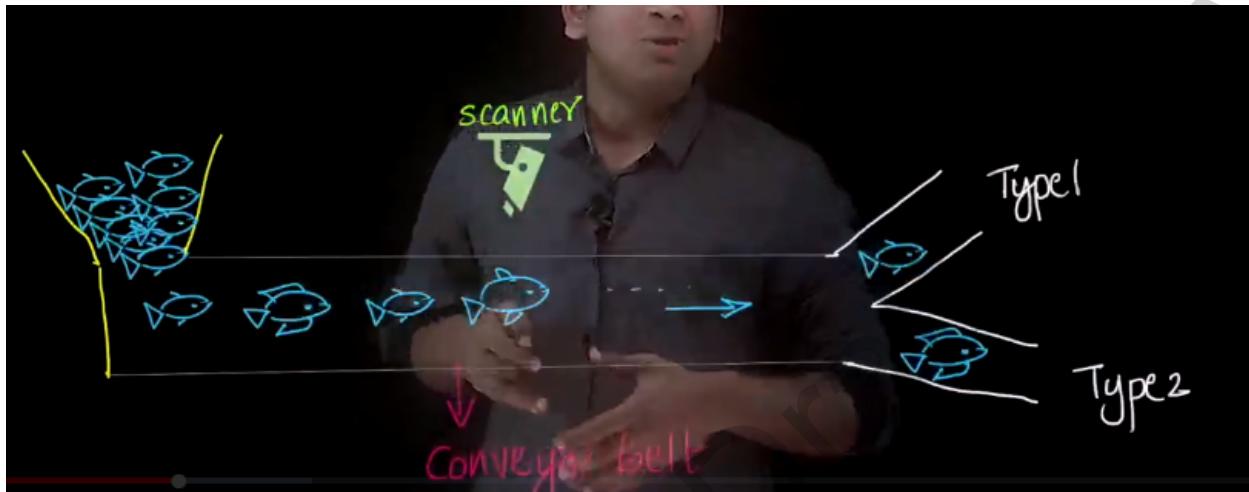
- In this chapter we will be learning about basics in Linear algebra (lines, planes, vectors, matrices etc.)



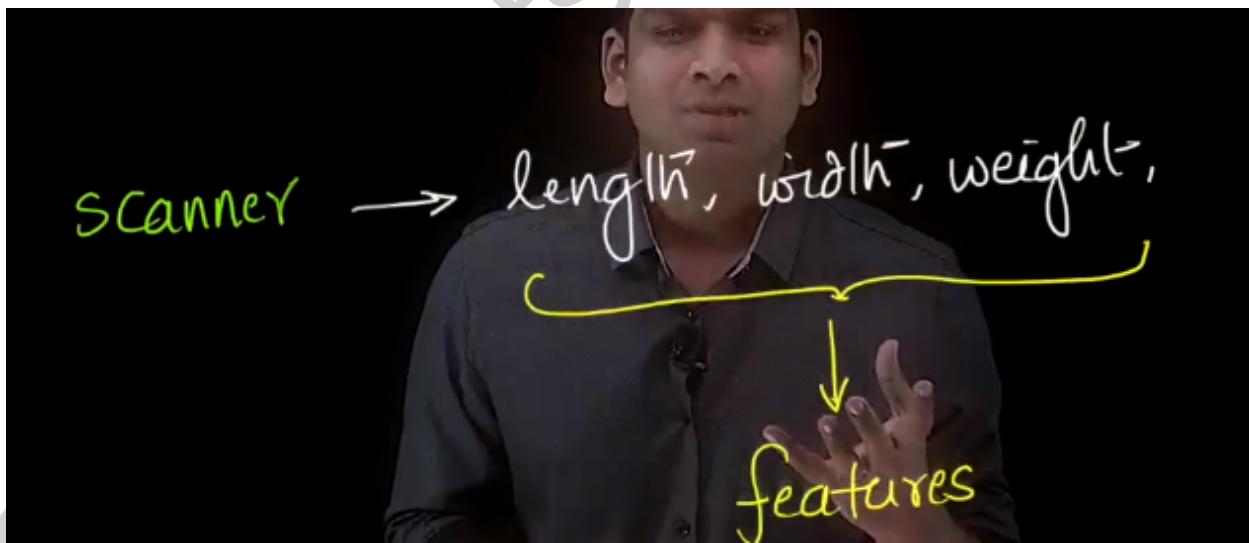
- We will discuss concepts which are relevant to ML/AI and solve them using real world problem context.

2.2 Real world ML classification problem: Sorting Fish in a Factory

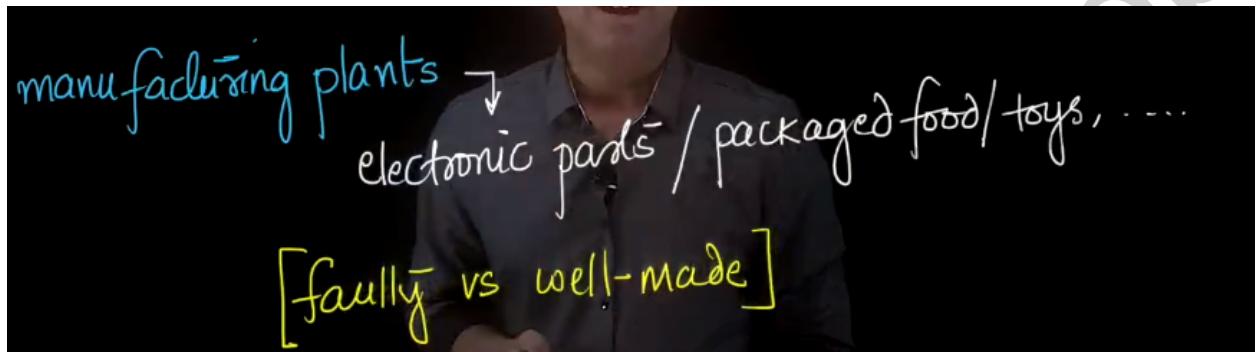
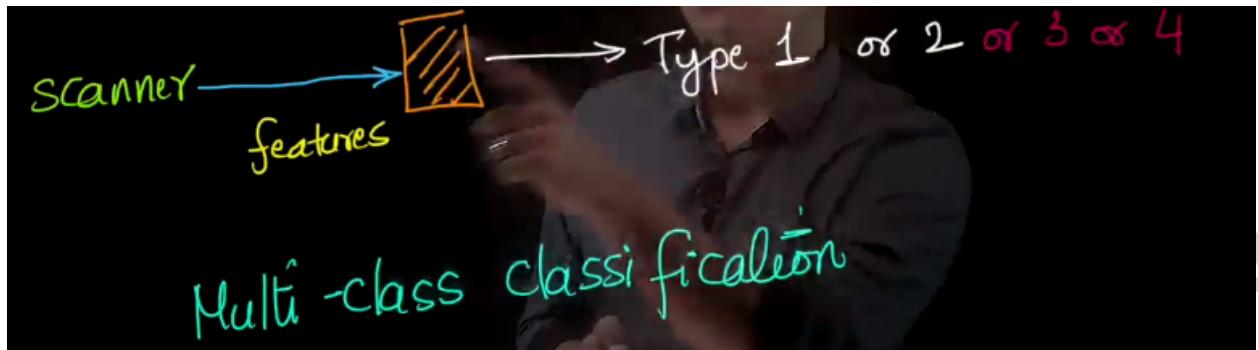
Below is an example shown at timestamp 2.28 in the video



Labour intensive & need for human experts



- Sorting fish problem can be treated as classification problem where we have to classify whether a given fish is of type 1 or type2. It is clearly an labour intensive task.
- We try to automate the task by identifying the features of the fish such as length, width, weight and using these features we classify each fish whether it is type 1 or type 2



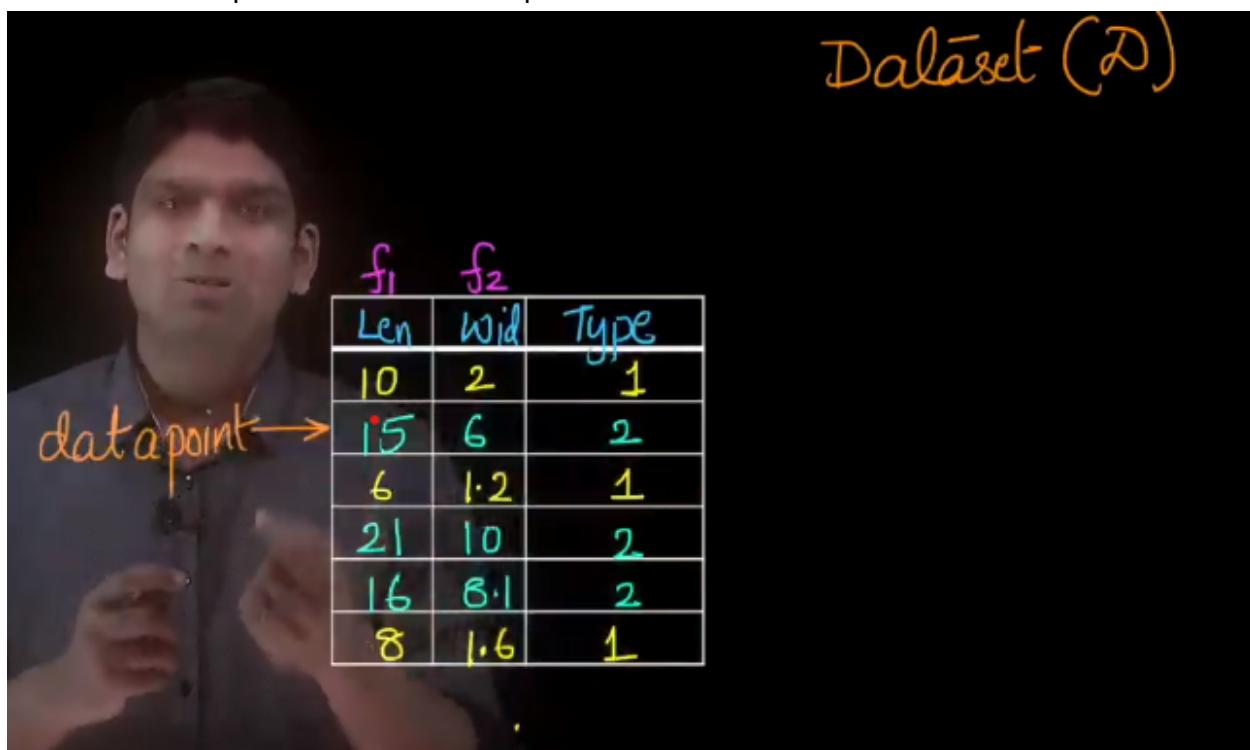
- Our task here is a binary classification task since we have two types of fish but this can be extended to multiclass classification as well.
- Similar classification tasks are performed across the manufacturing plants for identifying faulty vs good products.

2.3.Dataset & Plotting

At timestamp 0.30 in the video



Below is an example shown at timestamp 2.02 in the video



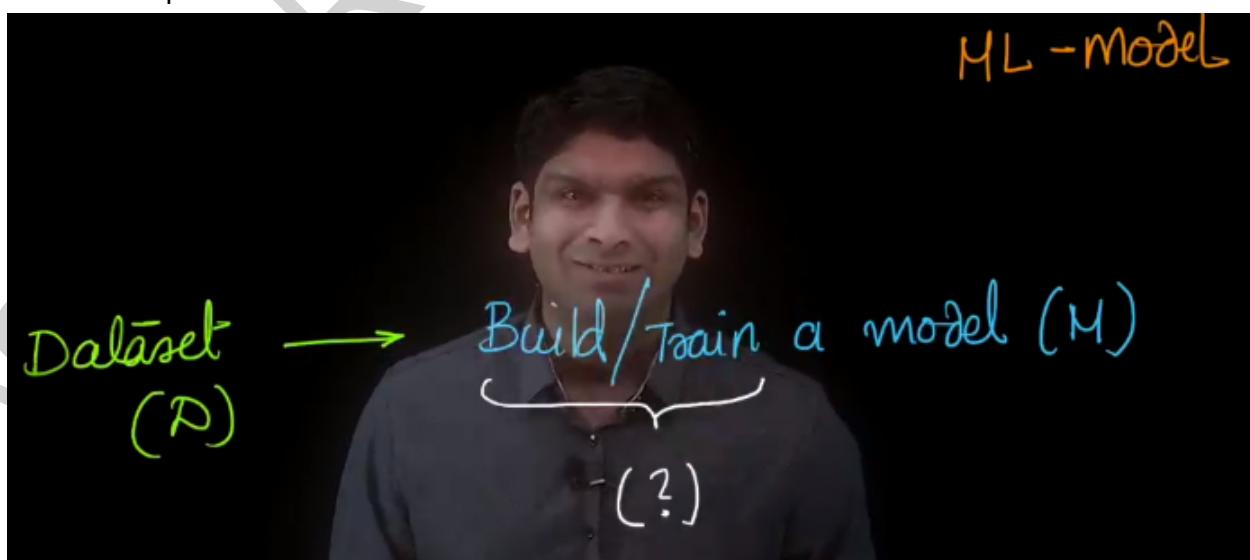
Dataset (\mathcal{D})

datapoint →

f_1	f_2	Type
Len	wid	
10	2	1
15	6	2
6	1.2	1
21	10	2
16	8.1	2
8	1.6	1

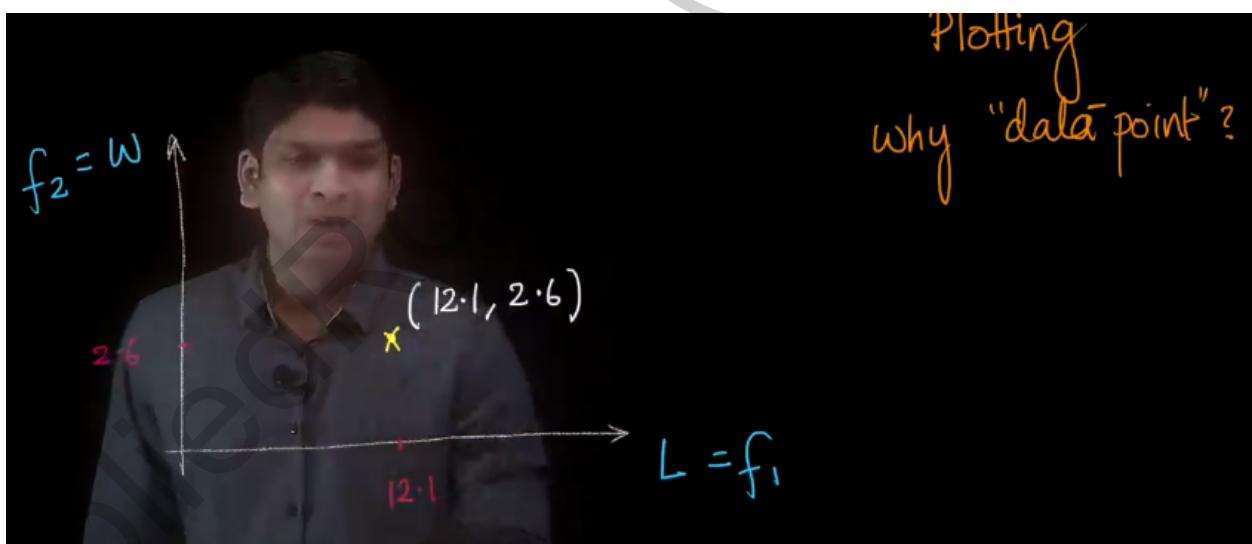
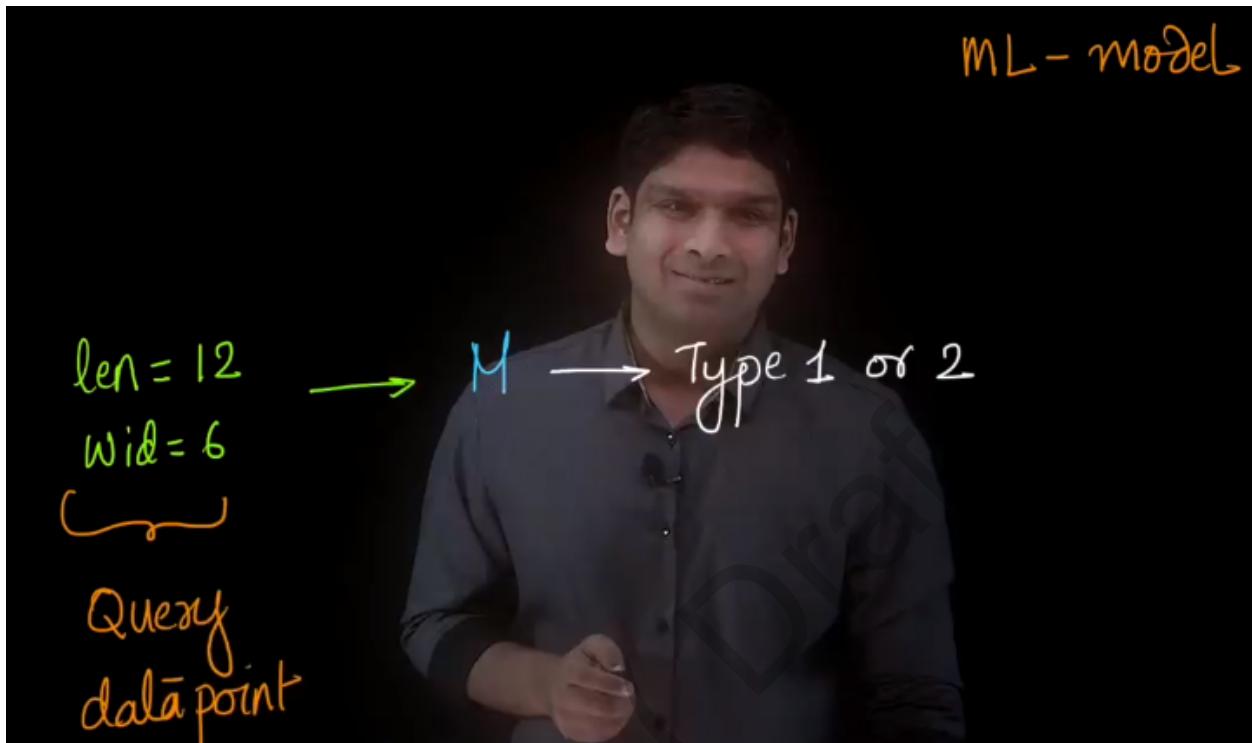
- The problem we are about to solve is a binary classification problem and we use the features of the fish for classification.
- We have a dataset(\mathcal{D}) in which each row describes the length,width,type of a fish.We call this information of each fish as a datapoint which is collected manually.
- The yellow data points represent fish of type1 and green data points represent fish of type 2.

At timestamp 3.31 in the video

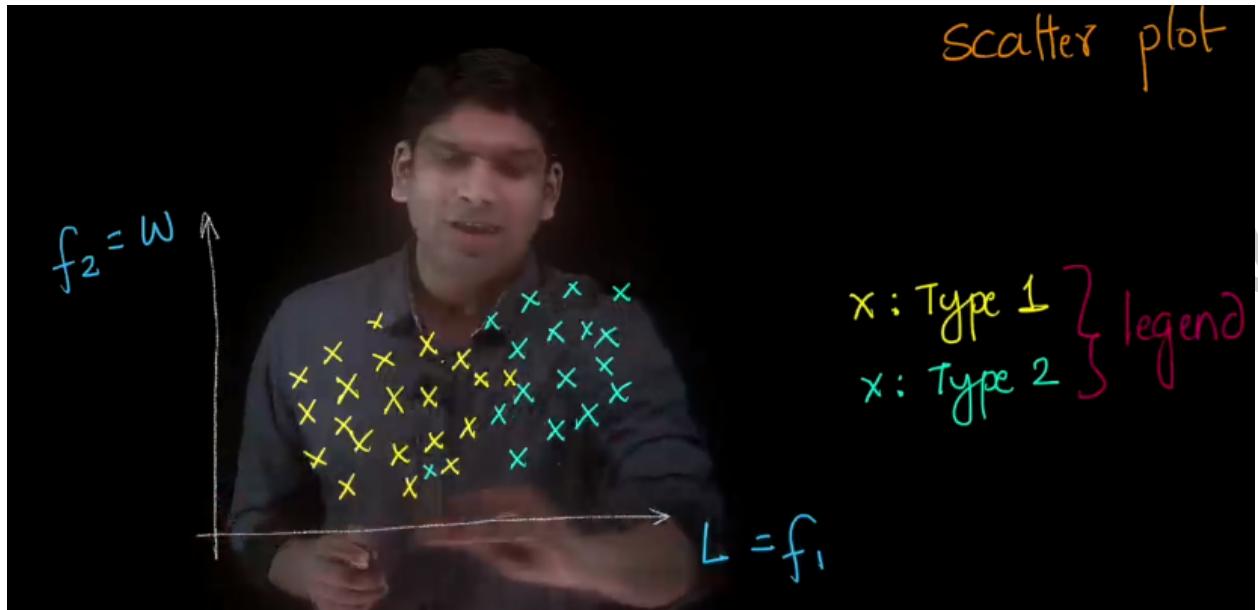


- Given the dataset D we are going to build or train a machine learning model (M).

At timestamp 6.03 in the video



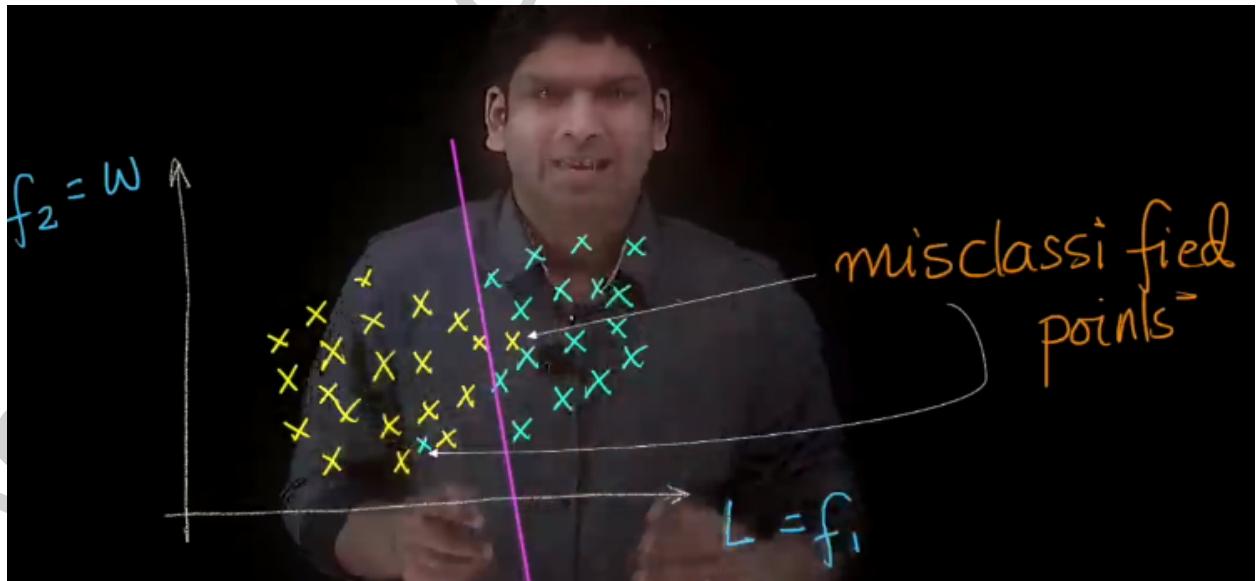
- Assume we have built model M ,consider we have a fish with features length=12,width=6 and we want to find whether it is of type 1 or type2. We call this a query datapoint.we can visualize the data point as shown.
- feature 1(length),feature 2(width) are represented along the X,Y axis respectively.(f_1, f_2) gives data about 1 fish.



- We can visualize the entire dataset in 2D using a scatter plot since we have 2 features ,and we can add legend to the plot for separating data points.

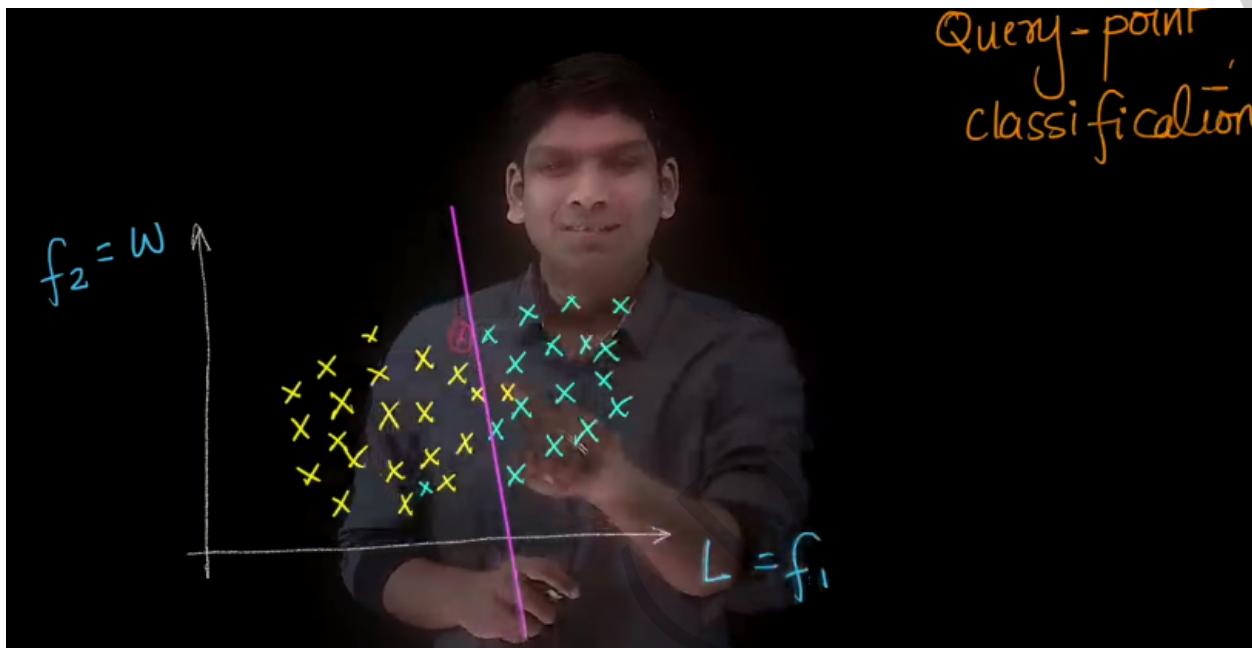
2.4 Lines, Circles and Ellipses for ML Classification

At timestamp 3.18 in video



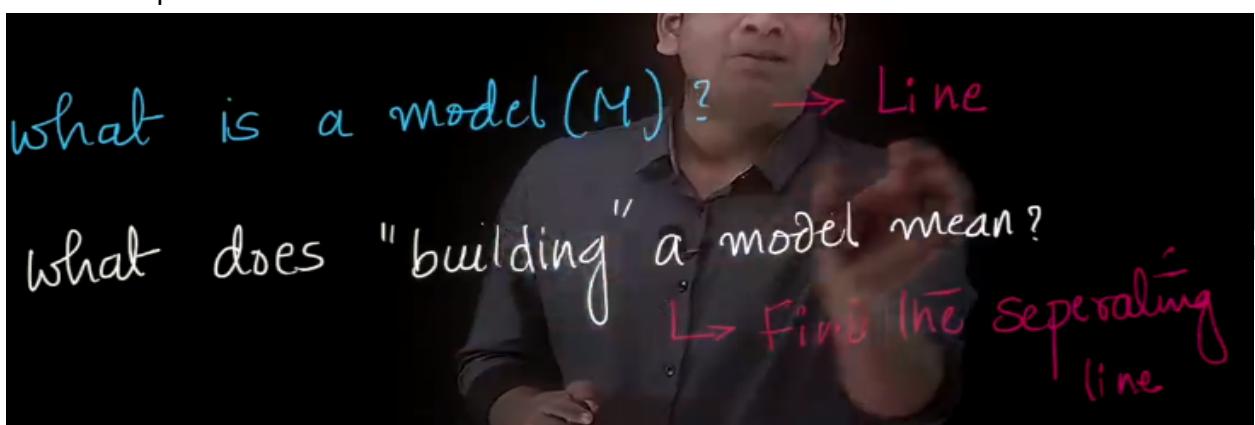
- You can clearly observe that the pink line is more or less separating the two types of fish. Let's say that the line is our model. It is separating our yellow points and green points fairly well except for the two misclassified points. Majority of the points are classified correctly.

At timestamp 6.13

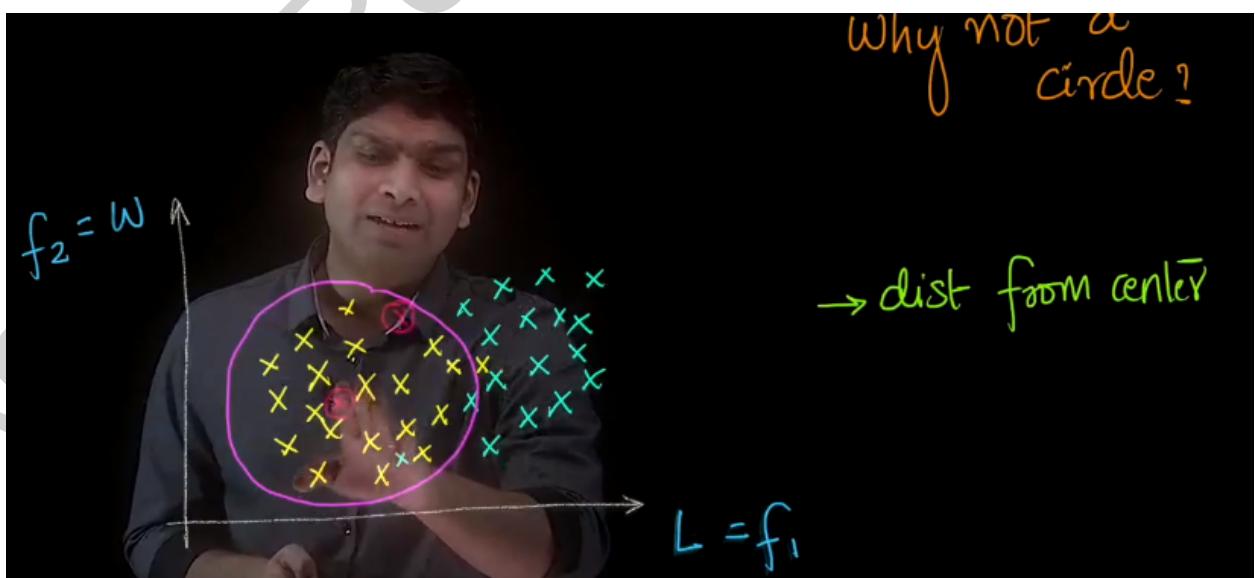
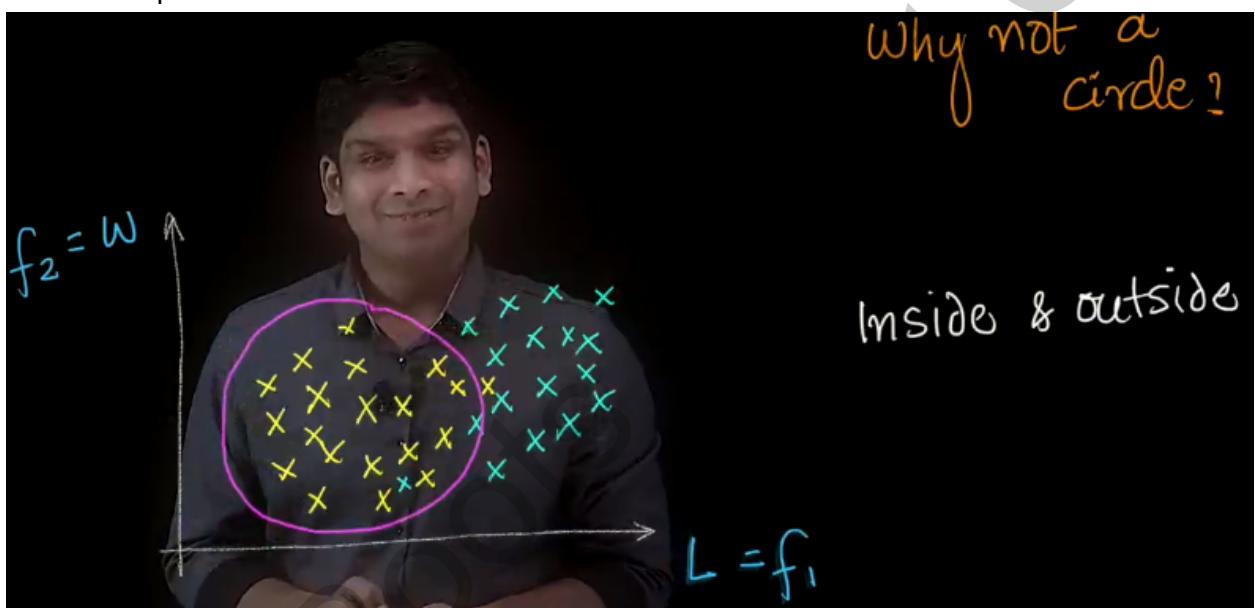


- Now that we have our line as a model ,we can have a query point and determine whether it is type 1 or type 2 fish.The closer the query point towards the line our model is less certain about its type.If the point is farther away from the line then model will be confident about the type of fish.
- Finding the classifier or model is nothing but finding a good line that separates the points.
- We classify a query point based on two things ,
 1. Which side of line point lies
 2. Distance of point from the line

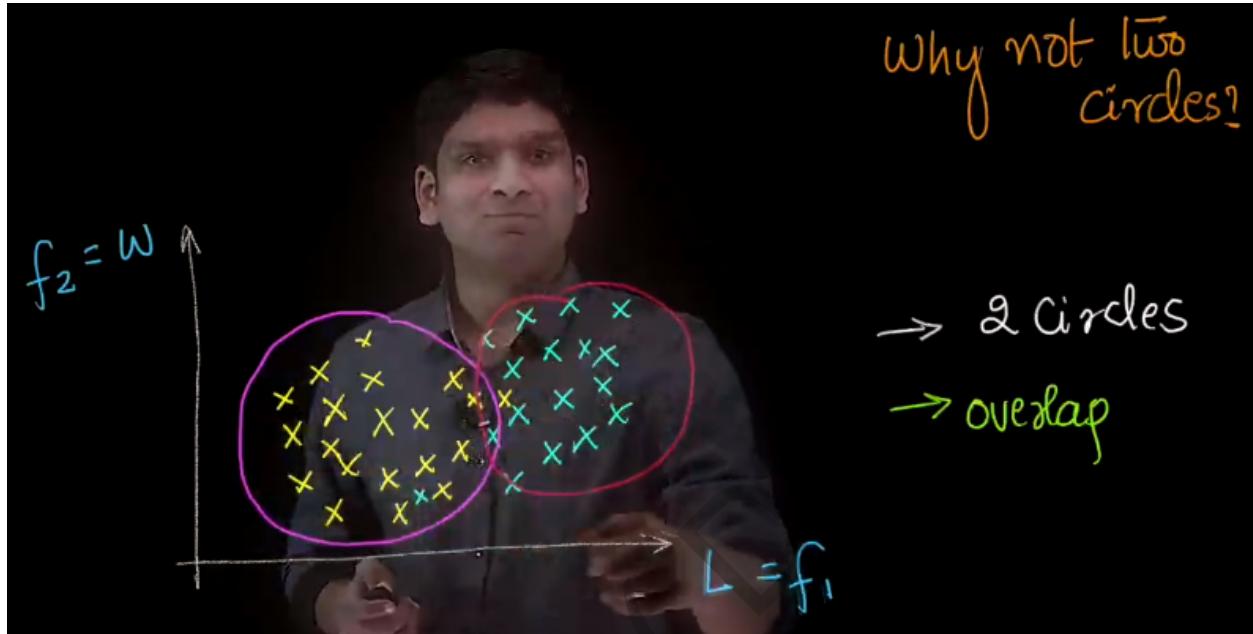
At timestamp 6.13



At timestamp 10.32



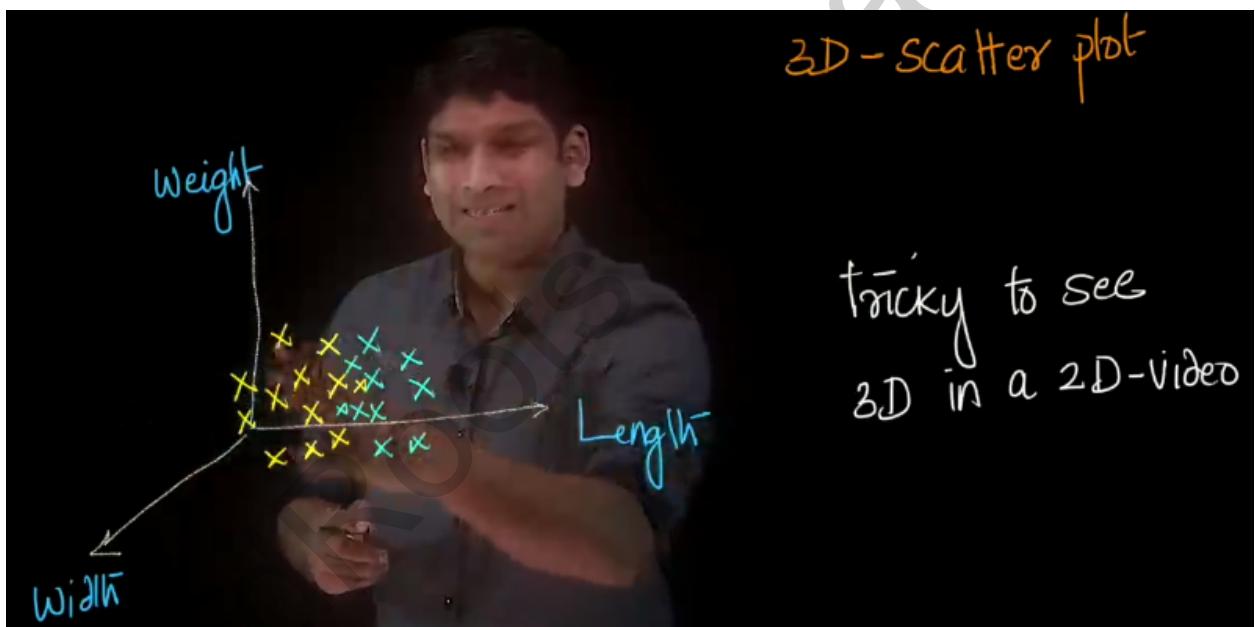
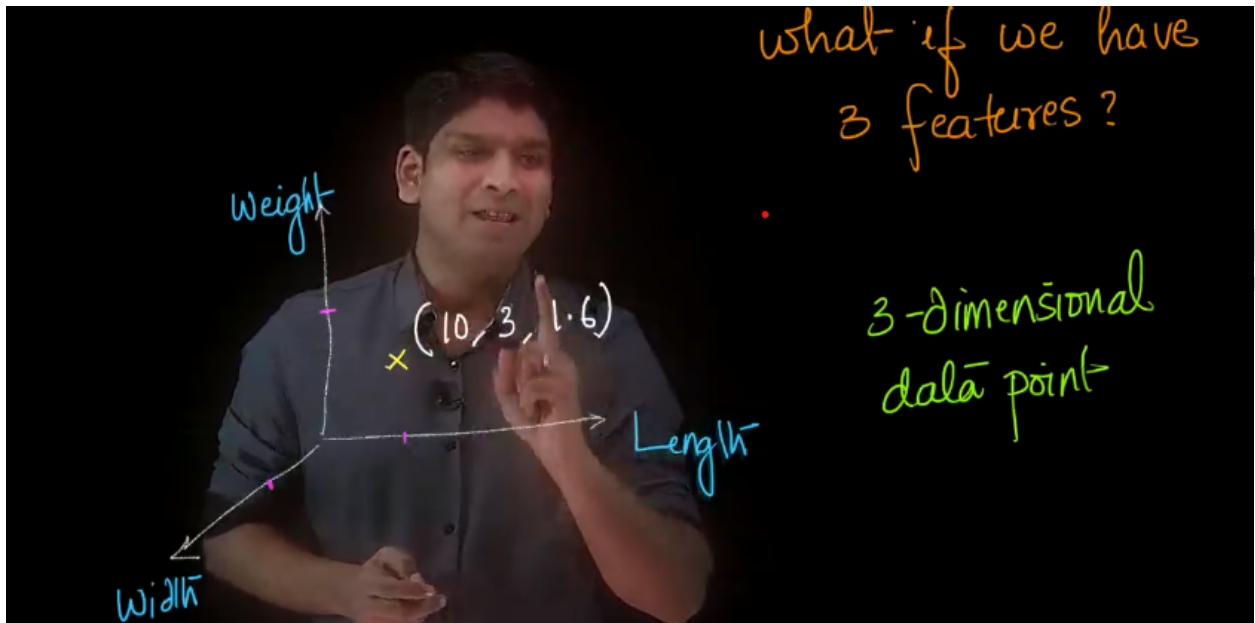
- The model can also be a circle and we classify the data points based on whether the data point lies inside or outside of the circle. We do this by calculating the distance of the point from the centre of the circle.
- The closer we are to the centre the more certain we are about the type of fish vice versa.



- We can also have a model which is represented using two circles
- For every model there are pros and cons
for example
 1. What if a point lies exactly on the line, circle
 2. What if the point lies in the region where circles overlap, and if point lies outside the circles
- Similarly we can use ellipse also for our binary classification problem.

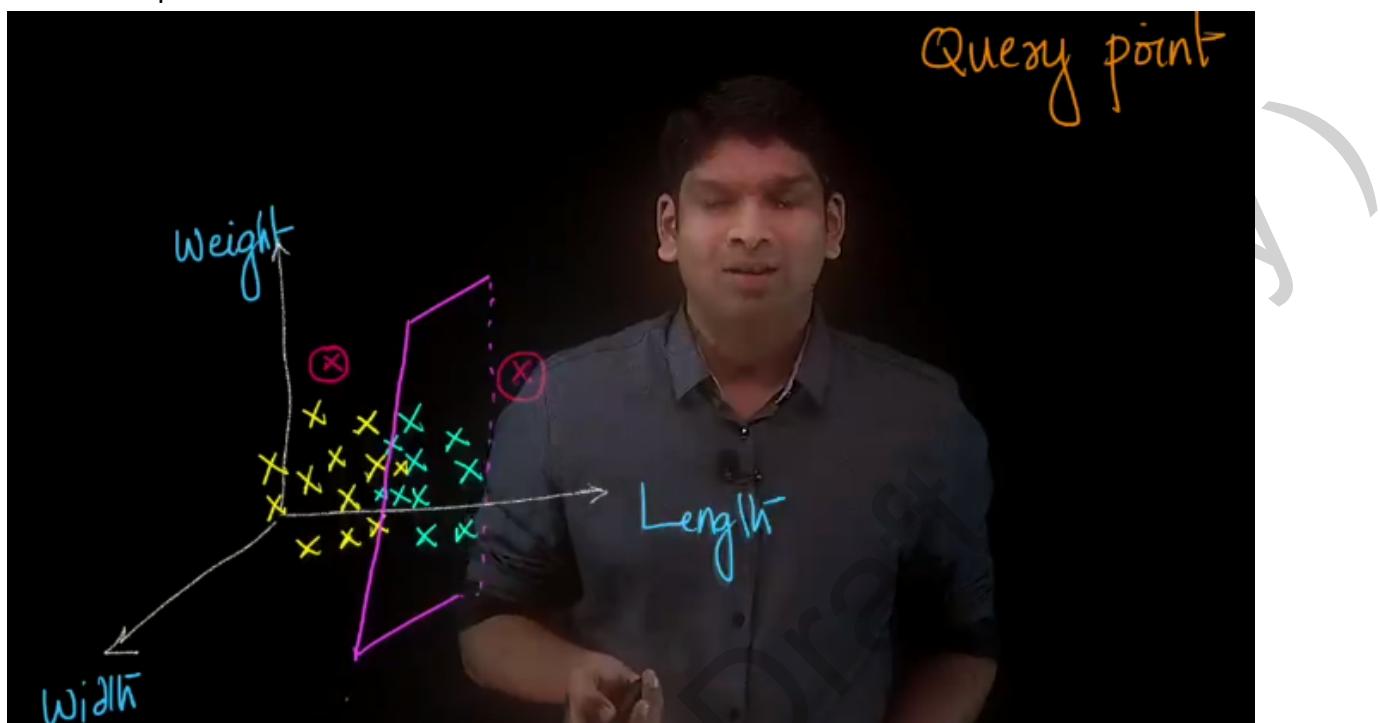
2.5 Planes, Spheres and Ellipsoids for ML classification

At timestamp 1.53 in the video



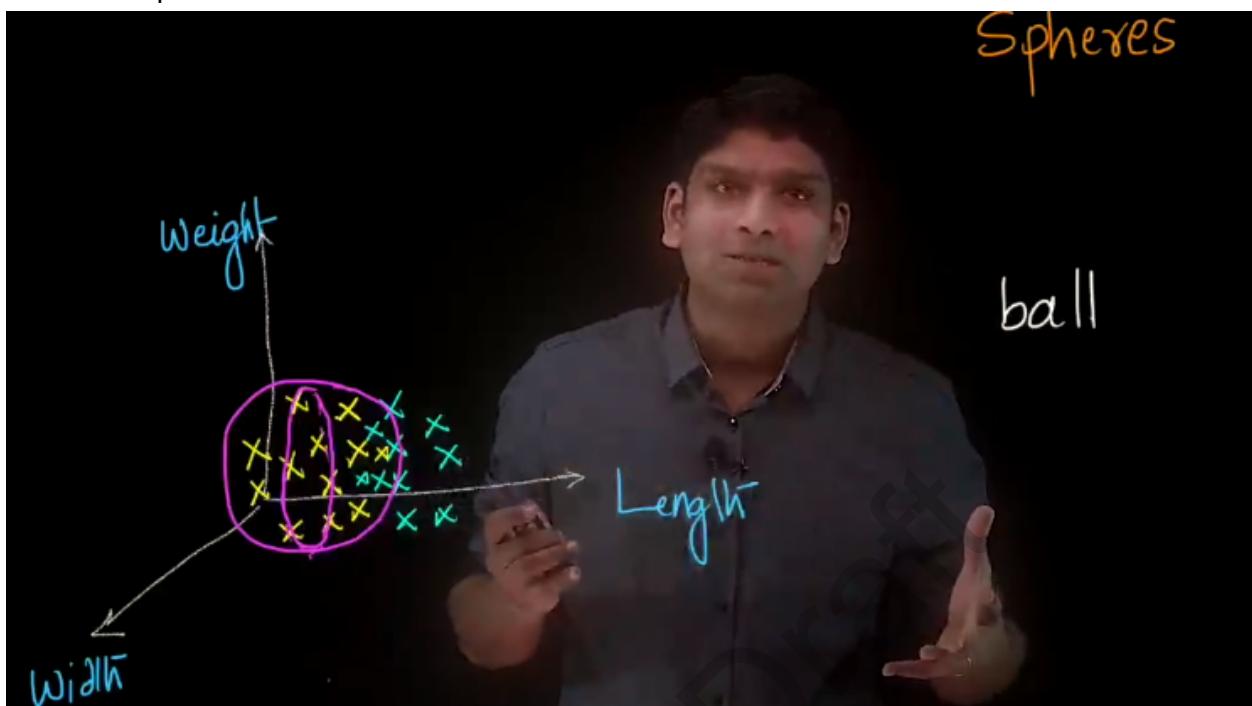
- The same way we visualize a datapoint in 2D we can visualize a data point in 3 D space if we have 3 features to represent a data point along the x,y,z axis respectively.
- We can also visualize the entire dataset in 3D space.

At timestamp 8.20 in video

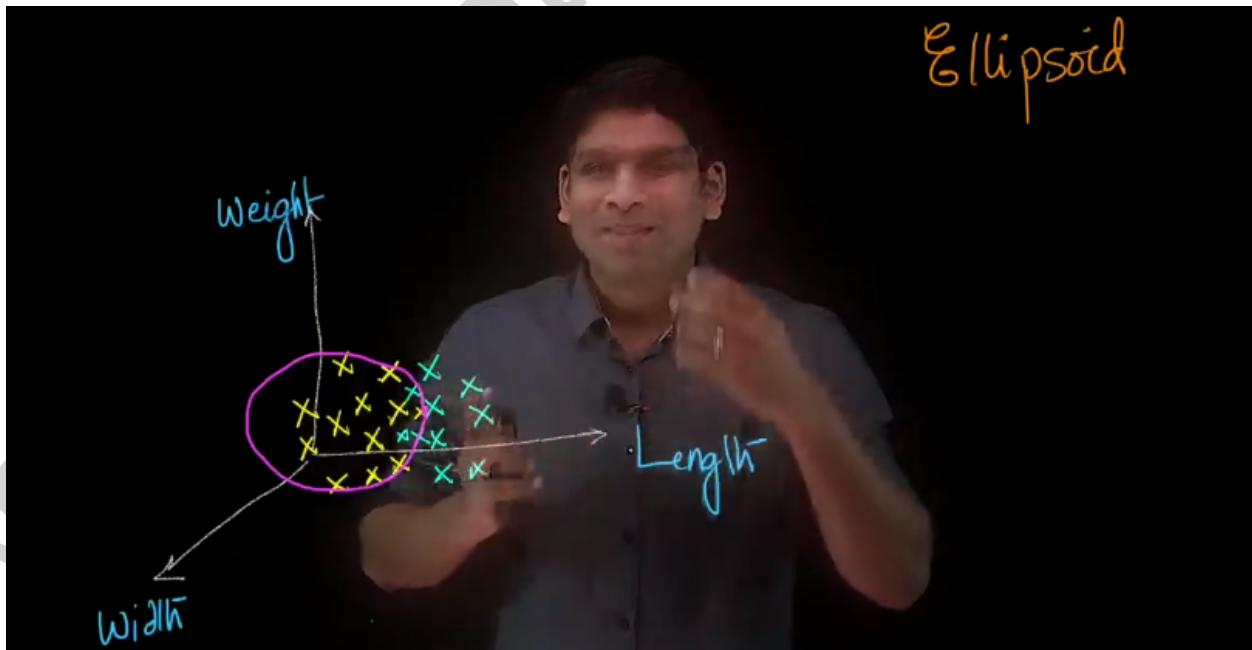


- We cannot use a line to classify the data points in 3D space; we need a plane for separating them. Plane is the simplest model in 3D
- Plane separates the 3D space into two regions , given a query point we find the distance of the point from the plane and which side of the plane the point belongs .
- This is similar to what we have discussed with line in 2D

At timestamp 9.43 in video



- Just like we have extended the concept line in 2D to plane in 3D ;we can extend circles in 2D to spheres in 3D .Everything inside the sphere is one class of data points and outside will be another class.How close the point is to the centre of the sphere determines how certain we are that the point belongs to a particular class.



- Similarly we can extend ellipse in 2D to ellipsoid in 3D for the classification.

3-D data \Rightarrow use planes/spheres/ellipsoids

At timestamp 13.2 min

what if 4-D,
5-D, 10-D, 100-D?

Let's say we have > 3 features

At timestamp 14.34 min

Hyper-X :

hyper-plane, hyper-sphere, hyper- ellipsoid

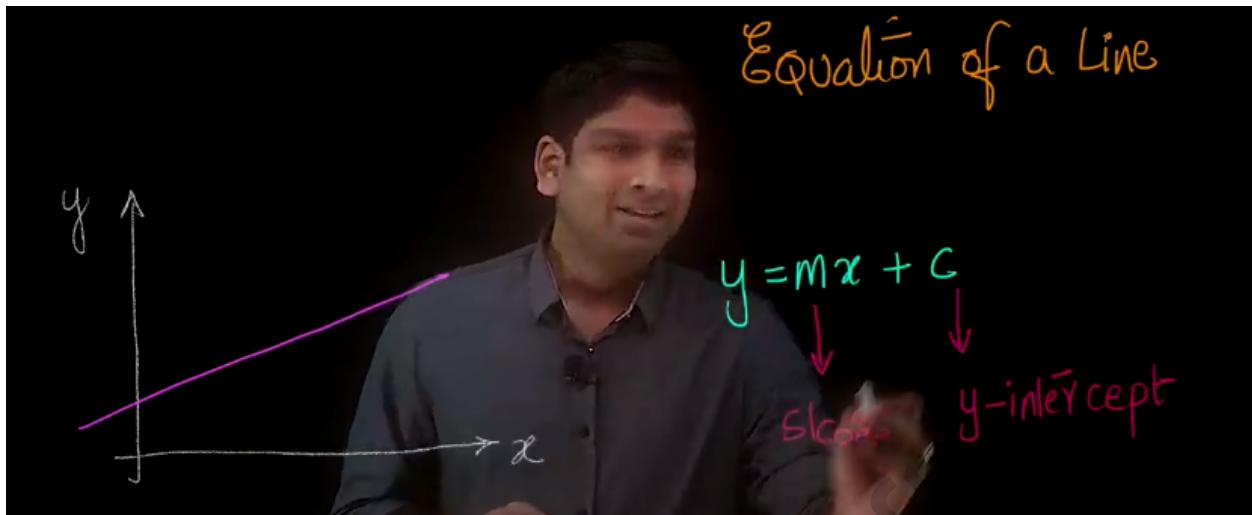


new concepts [use Linear Algebra]

- To summarize we use planes ,spheres,ellipsoids in 3D space and when we have more than 3 features we extend the same concept to higher dimensions as well.But we cannot visualize data in higher dimensions like we did in case of 2D and 3D;so we use linear algebra to understand it mathematically.We call them hyper planes ,hyper-spheres and hyper ellipsoids in higher dimensions.

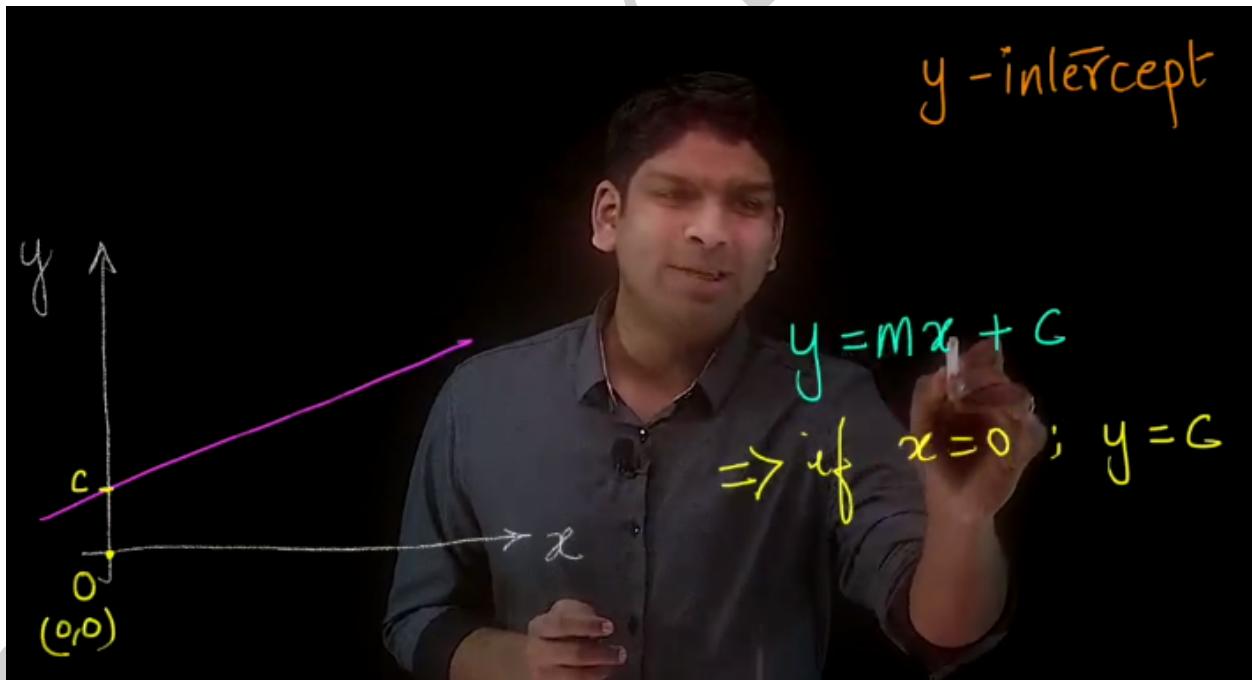
2.6 Equation of a line

At timestamp 0.22



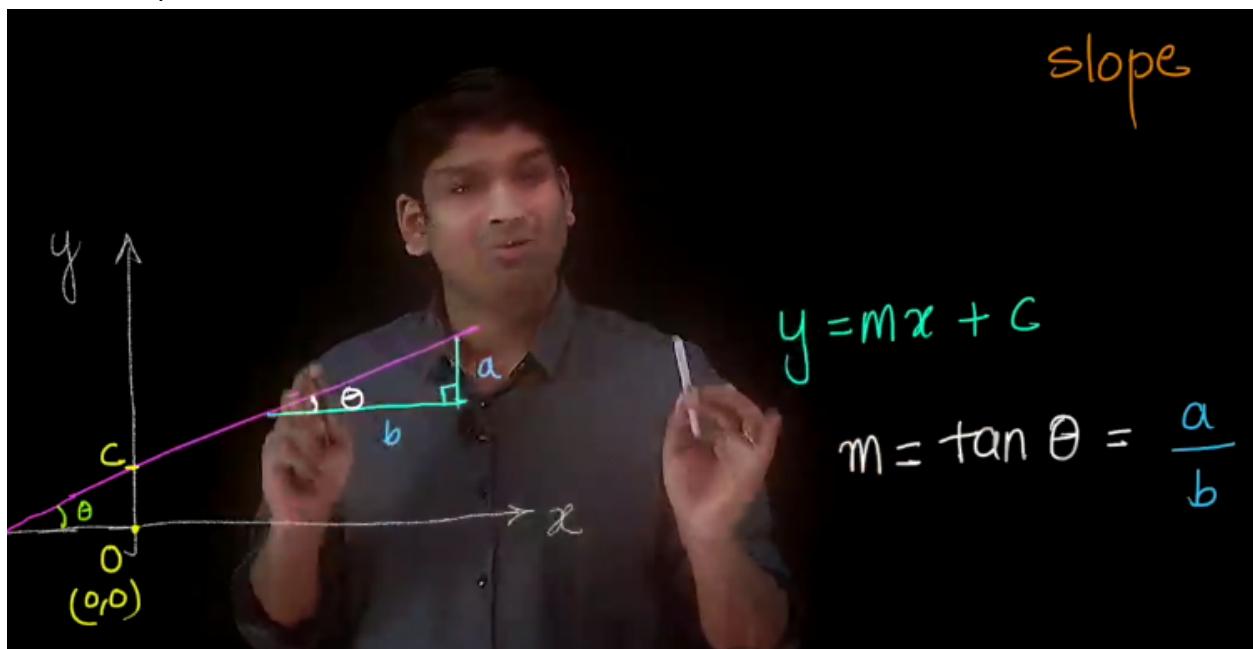
- Slope intercept formulation of a line is $y = m \cdot x + c$, whereas m is the slope and c is y-intercept.
- y intercept is the point at which the line intercepts the y axis.

At timestamp 2.32



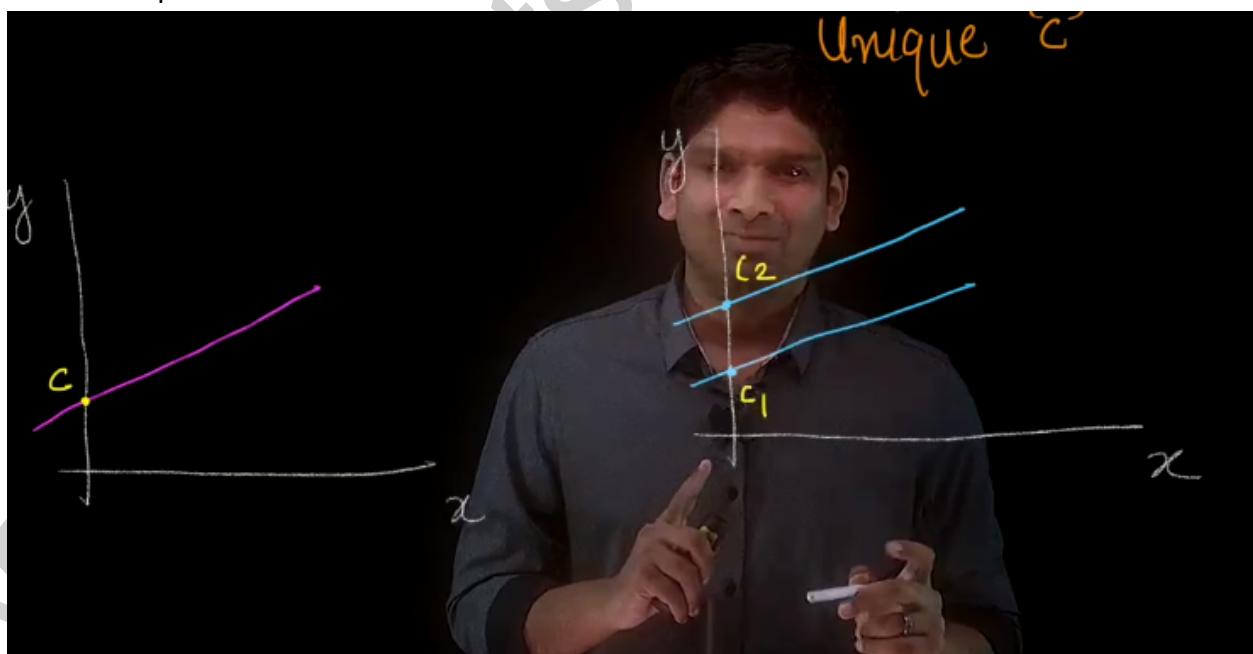
- For the line shown above the line is c units away from the origin ,hence for this line $y\text{-intercept}=c$.Algebraically at $x=0$ we have $y=c$

At timestamp 4.5 min



- Slope m is the angle made by the line with x -axis, you can calculate slope $\tan\theta = a/b$ as shown above.

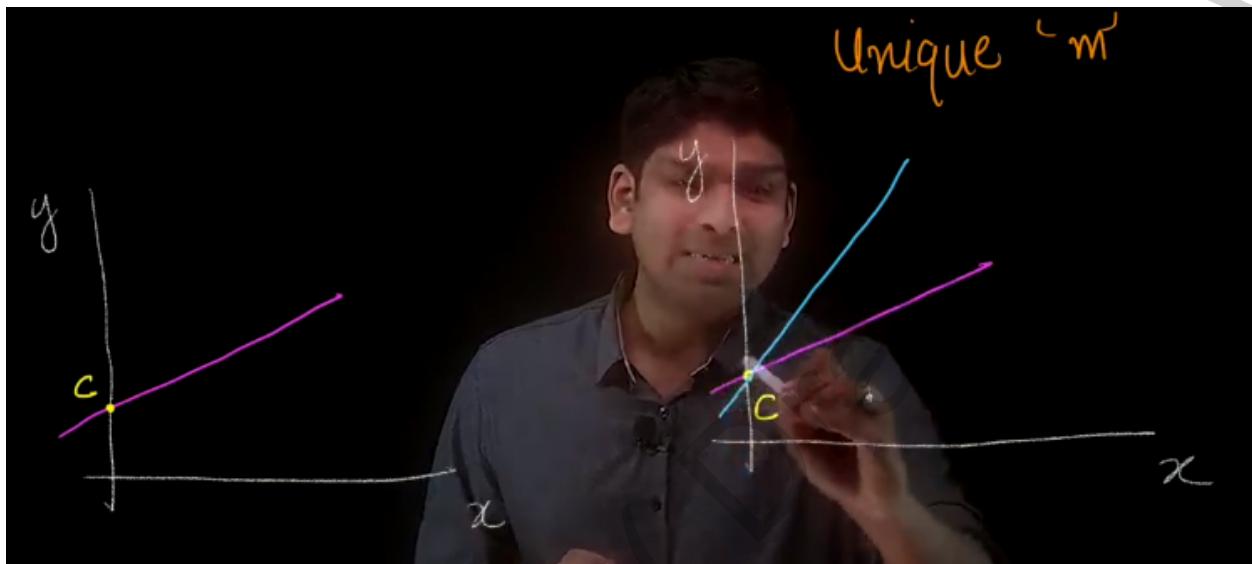
At timestamp 7.02



- Intercept c will always be unique for a given line.

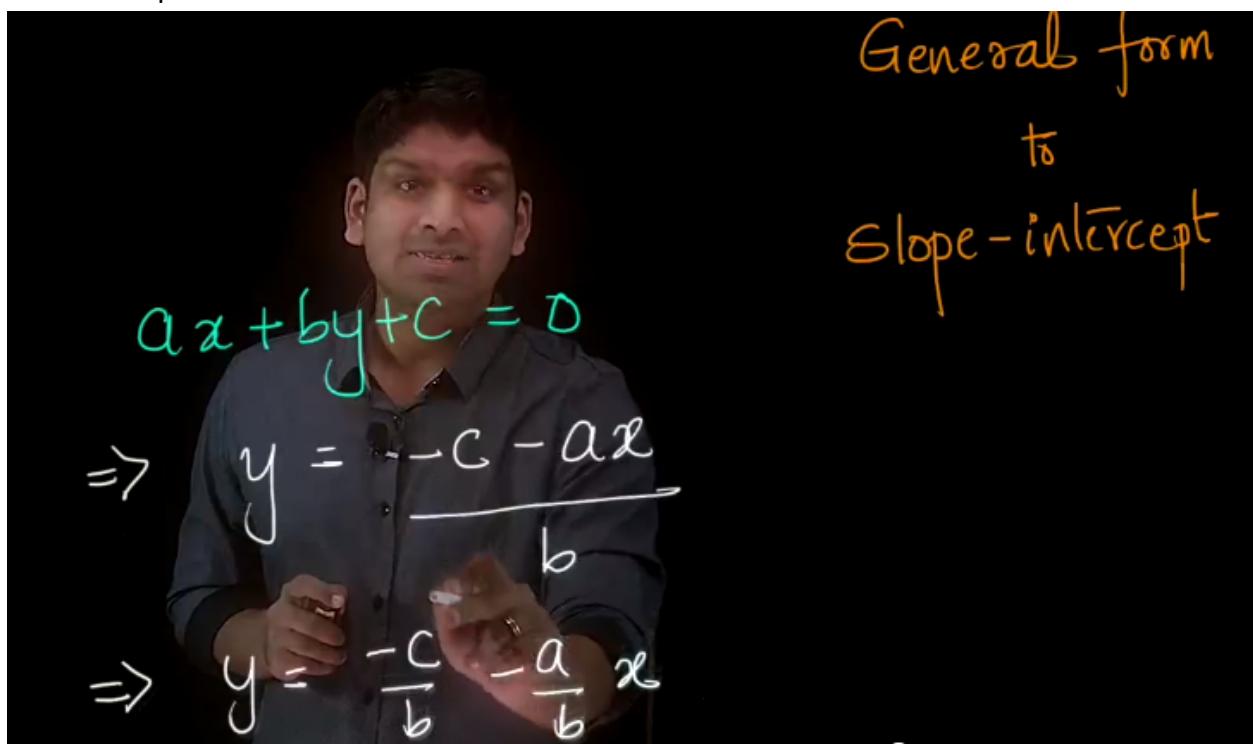
- If we have two lines parallel to each other their slopes will be same but intercept will not be same
- Given unique m and c there will only be one such line .

At timestamp 8.29



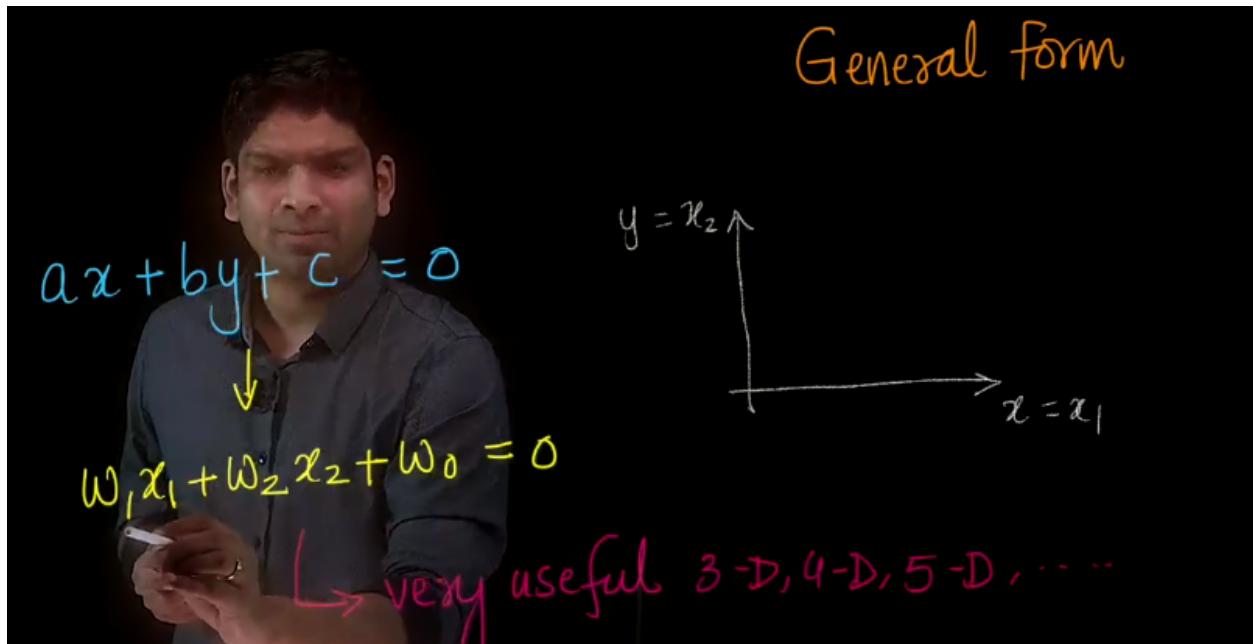
- The intercept c can be the same for two lines but their slopes will be different;because the angle made by each line with x axis will be different.

At timestamp 12.31



- General form of a line is given by the equation $a.x+b.y+c=0$. (x,y correspond to x and y coordinates). This equation is important because we use this to represent lines, planes, hyperplanes in ML.
- Given a general form of line we can easily convert it into slope-intercept form as shown above. Slope $m=-a/b$ and y -intercept is $-c/b$.

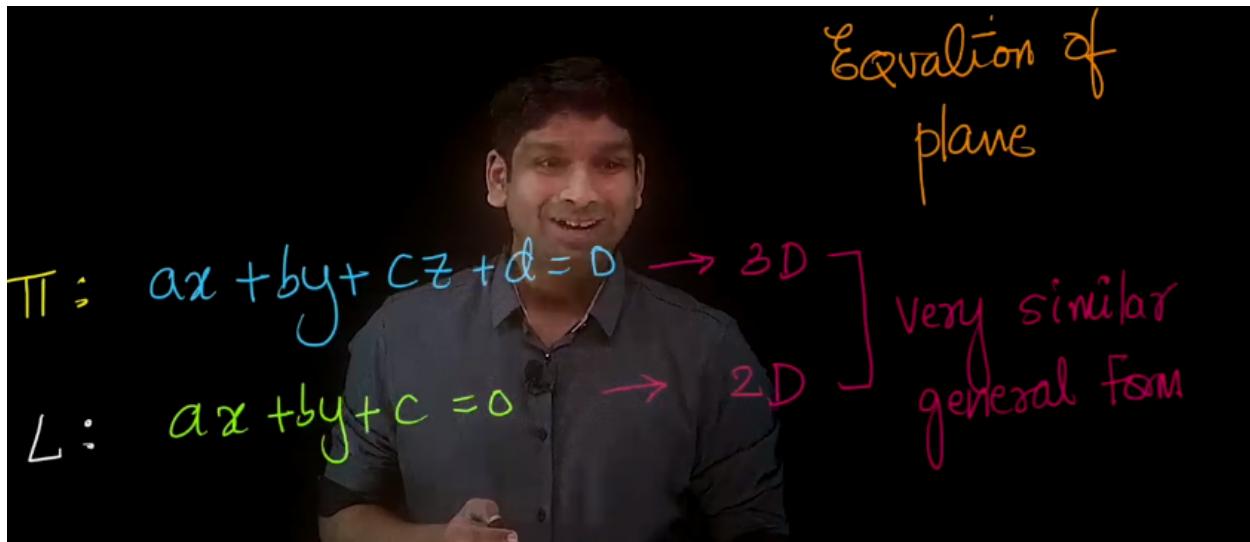
At timestamp 16.13



- Instead of using x,y as x coordinate and y coordinate from now on we use x1,x2 as the coordinates ;so that we can extend the same concept even if we have n dimensions (x1,x2,x3,x4....).
- Also for representing coefficients we are going to use w1,w2,w0 instead of a,b,c
- For understanding the geometry of line,plane and hyperplane in higher dimensions we are going to represent the equation as $w_1x_1+w_2x_2+w_0=0$

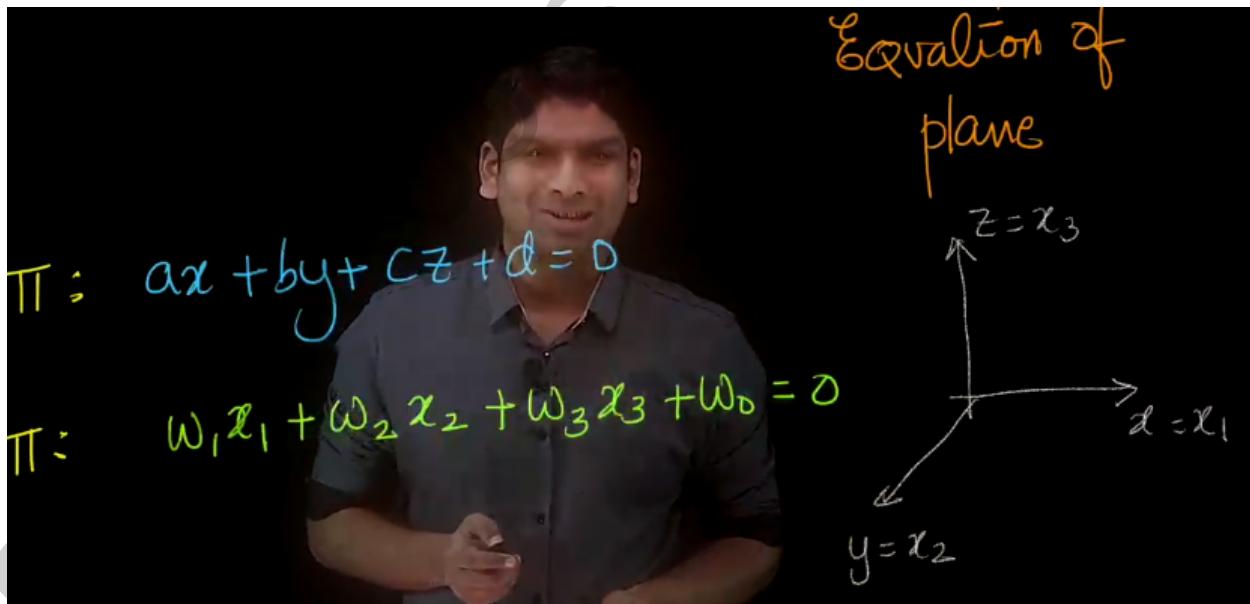
2.7 Equation of a plane & hyper planes

At timestamp 2.10



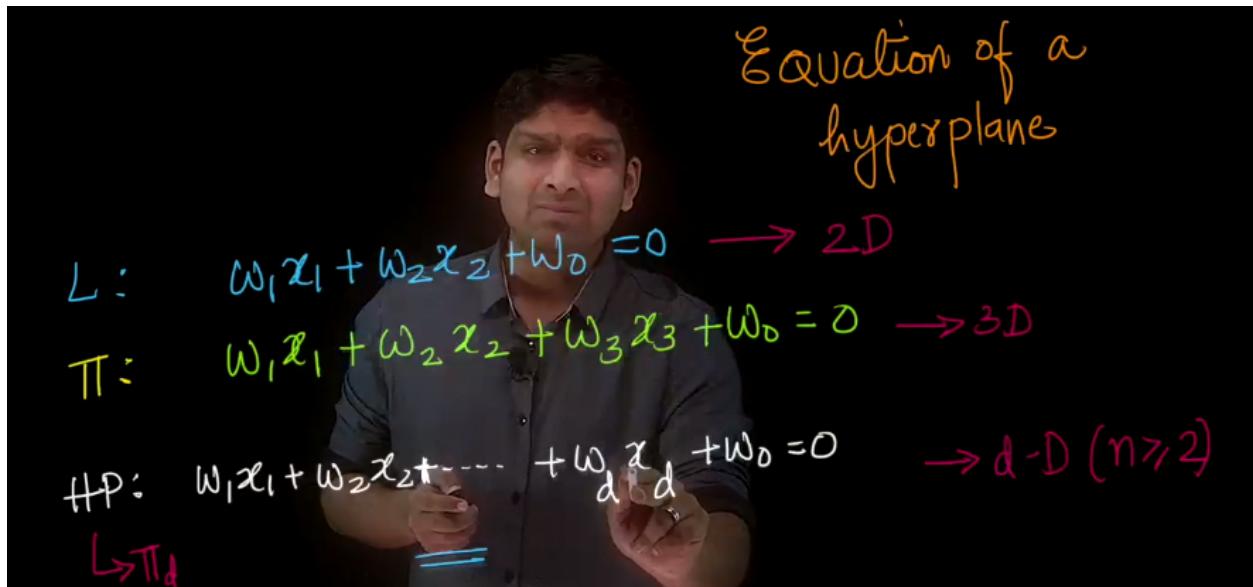
- Equation of the plane is $ax+by+cz+d=0$.
- We can clearly notice that the general form of line in 2D and plane in 3D are similar, just like line separates two regions in 2D, a plane separates two regions in 3D. In 2D we have only 2 axes and in 3D we have 3 axes.

At timestamp 3.1



- From now on we use the representation as shown above for representing a plane.

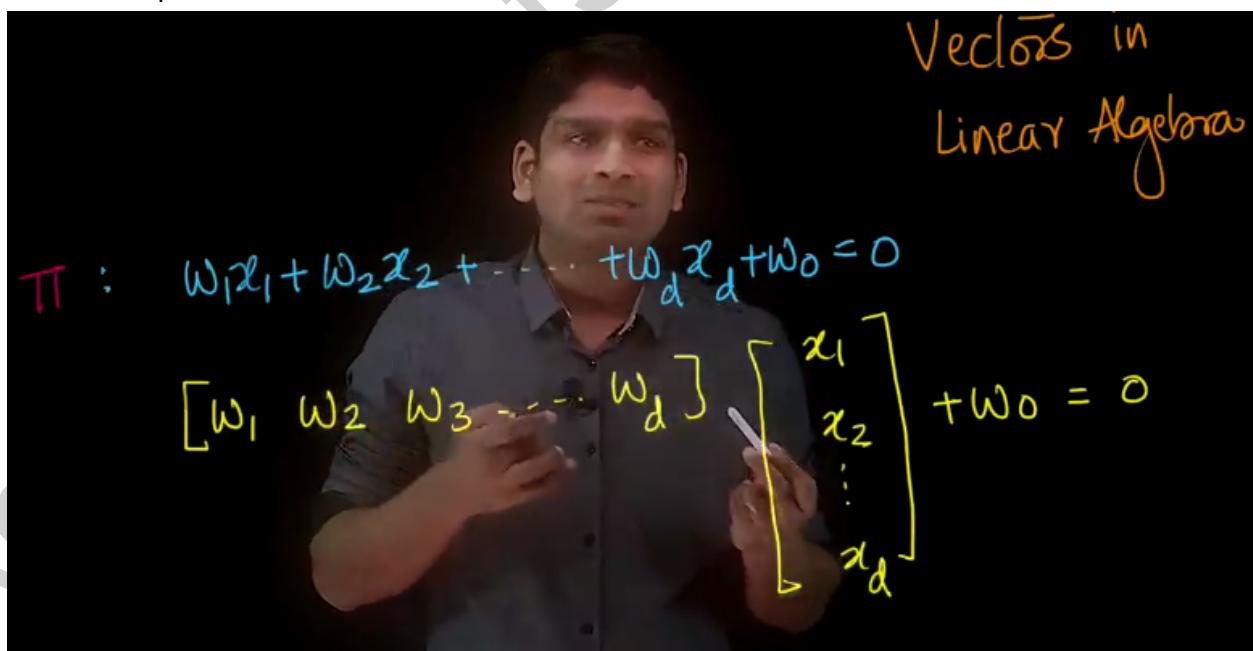
At timestamp 4.26



- Like we have lines in 2D, planes in 3D, we can have d dimensional hyperplanes in dD.
- A hyperplane in d-dimensions separates the d dimensional space into two regions.
- The equation for representing dD hyperplane is shown above.

2.8 Equations using Linear Algebra

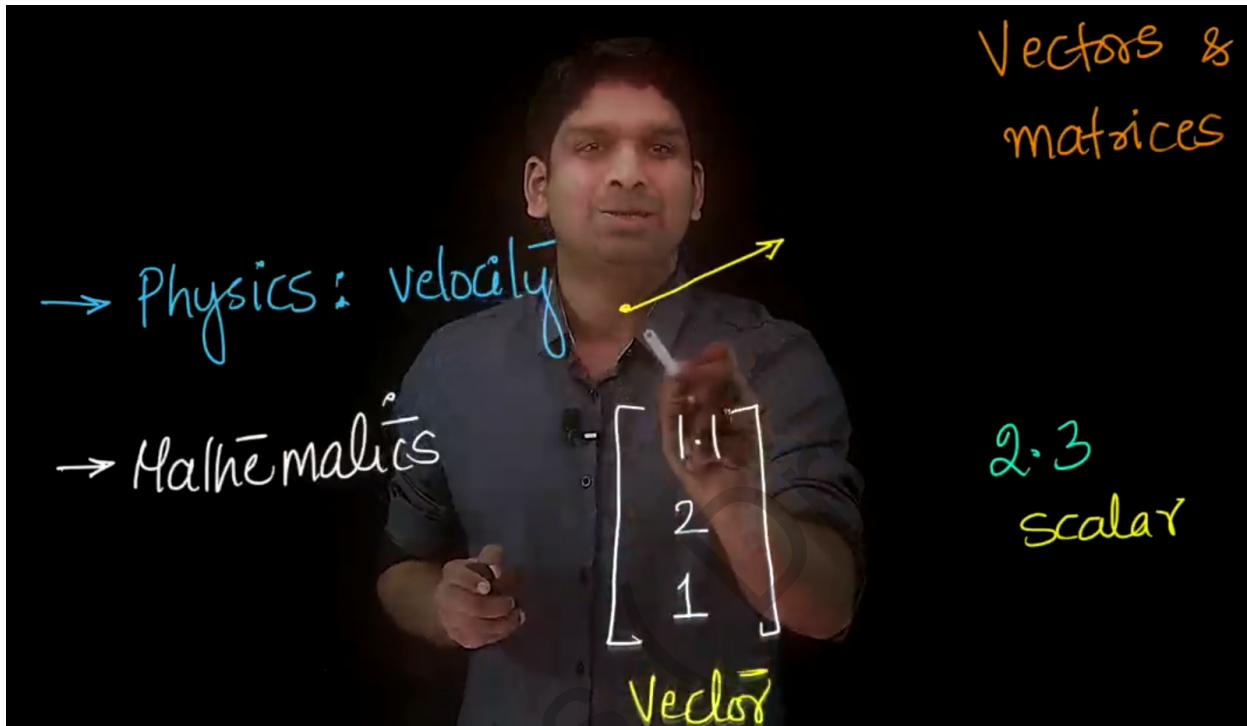
At timestamp 3.10



- We are familiar with the concept of vectors in linear algebra, we can write the equation of a plane in a crisp way using vectors as shown above.

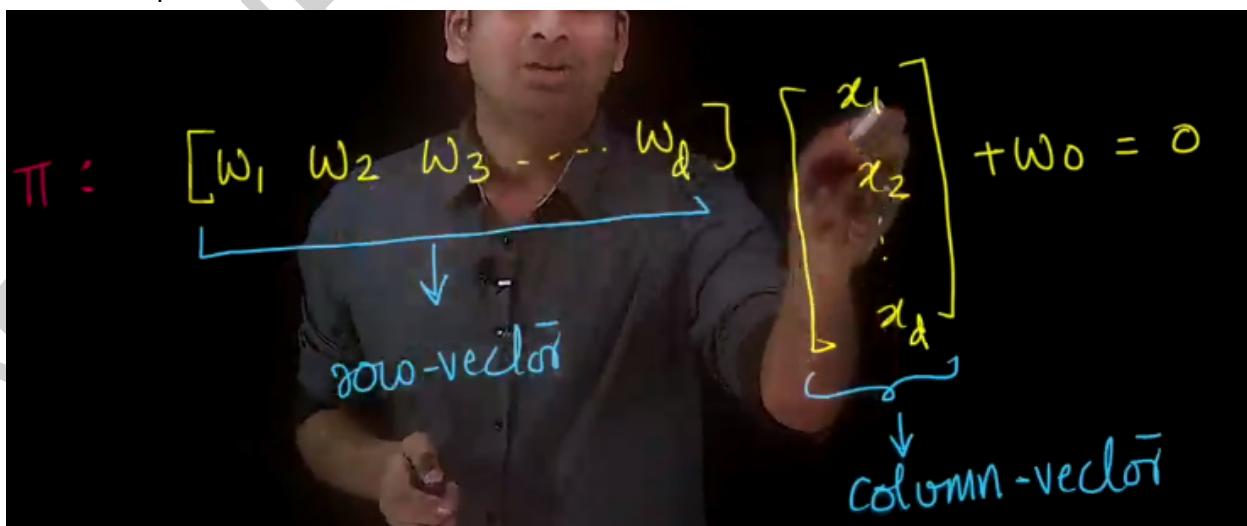
- W is called a row vector and X is called a column vector ,both vectors are of d dimensions. When we perform matrix multiplication for these two vectors we obtain the equation of the plane .

At timestamp 5.48



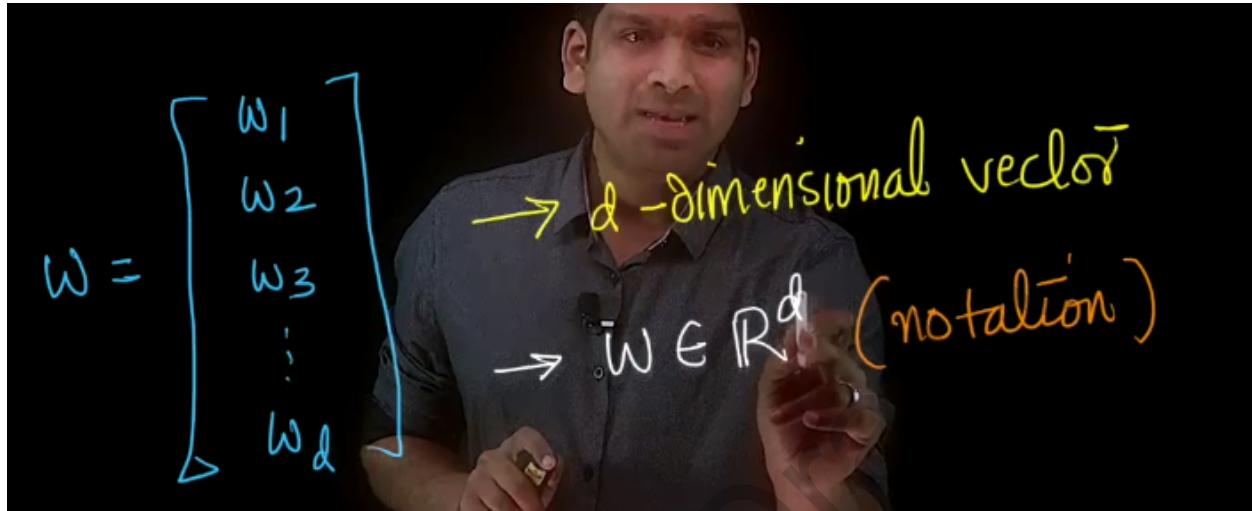
- In physics a vector has magnitude and direction, the concept of vectors arose both in physics and maths. Instead of using the physics approach of vectors we use mathematics interpretation vectors.
- In mathematics a vector and a scalar is represented as shown above.

At timestamp 6.58



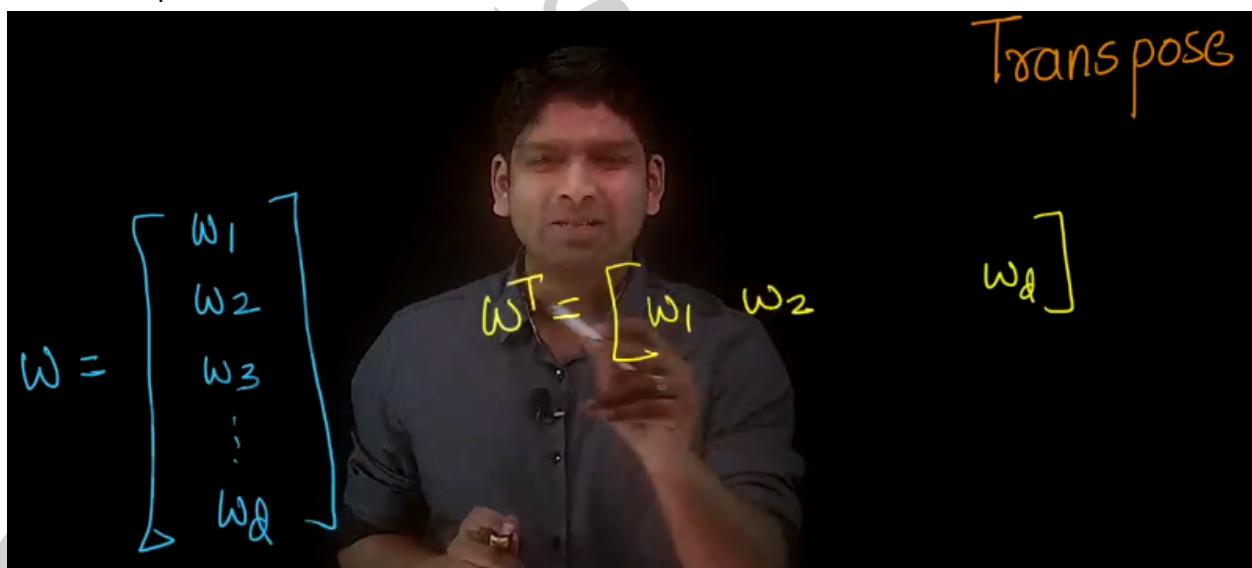
- A row vector and a column vector are represented as shown above, when you multiply a row vector with a column vector the resultant will be a scalar.
- When we say a vector in maths by default it is a column vector. (it is a convention)

At timestamp 9.28



- W is a d dimensional vector and we represent it as shown above. (R is set of real numbers)

At timestamp 10.15



- By default every vector is a column vector, we convert a column vector into a row vector by taking the transpose as shown above. Transpose is an operation we perform on vectors to convert column vectors to row vectors and vice versa.

At timestamp 11.15

Transpose of a matrix

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$
$$A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$$

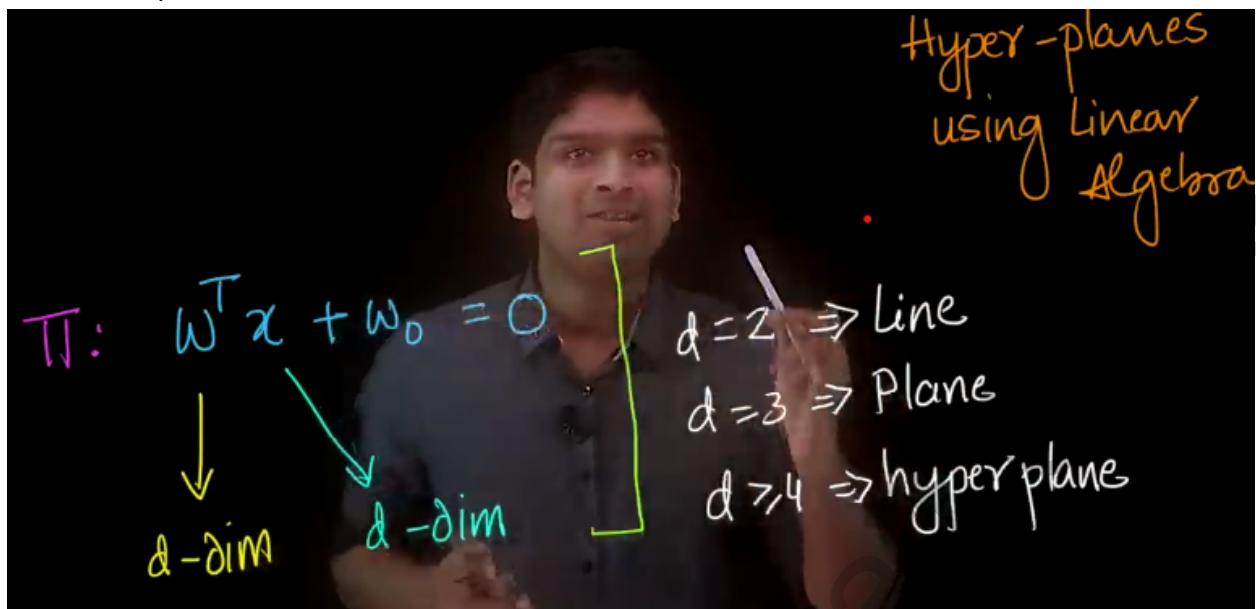
- Transpose can be used in matrices as well as shown above .Rows will be interchanged to columns and vice versa.

$$\Pi: w^T x + w_0 = 0$$
$$\Pi: w^T x + b = 0$$

NOTE: w & x are same dimensional vectors

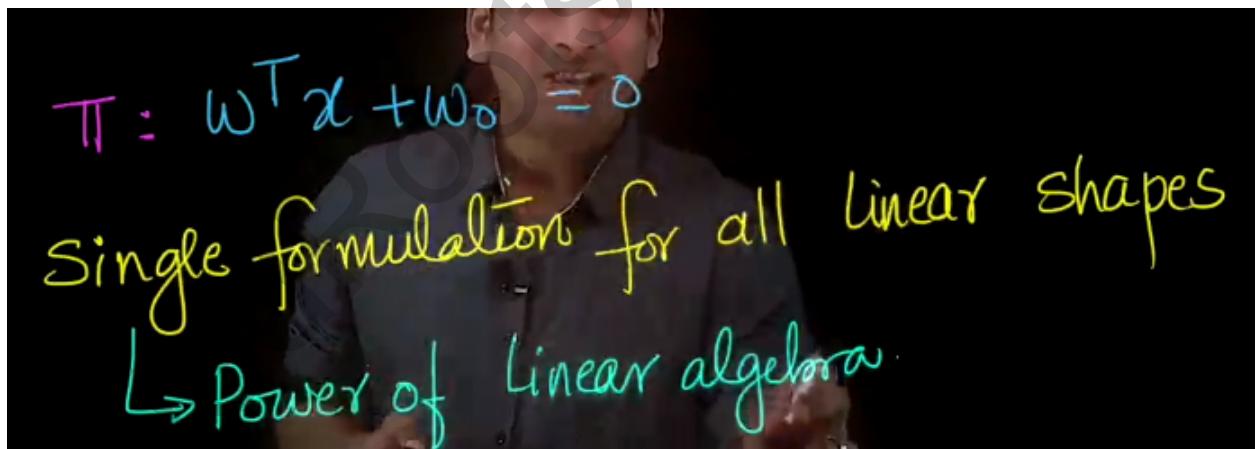
- So now we can represent the equation of the plane as shown above.
- Note that w , x should be of the same dimensions.

At timestamp 15.49



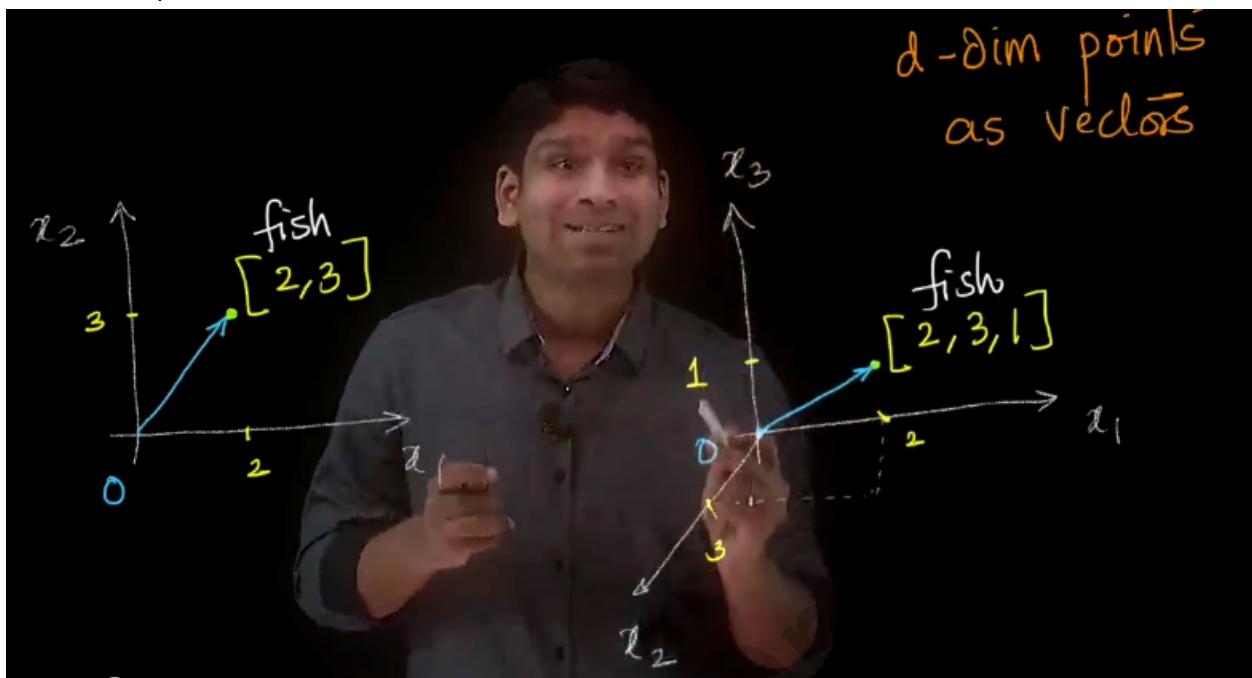
- The above equation when we have 2 dimensions we get a line, in 3 dimensions we get a plane and if $d \geq 4$ we get a hyperplane because w, x are d dimensional vectors.
- It is the most elegant and widely used equation in Machine learning.

At timestamp 16.7



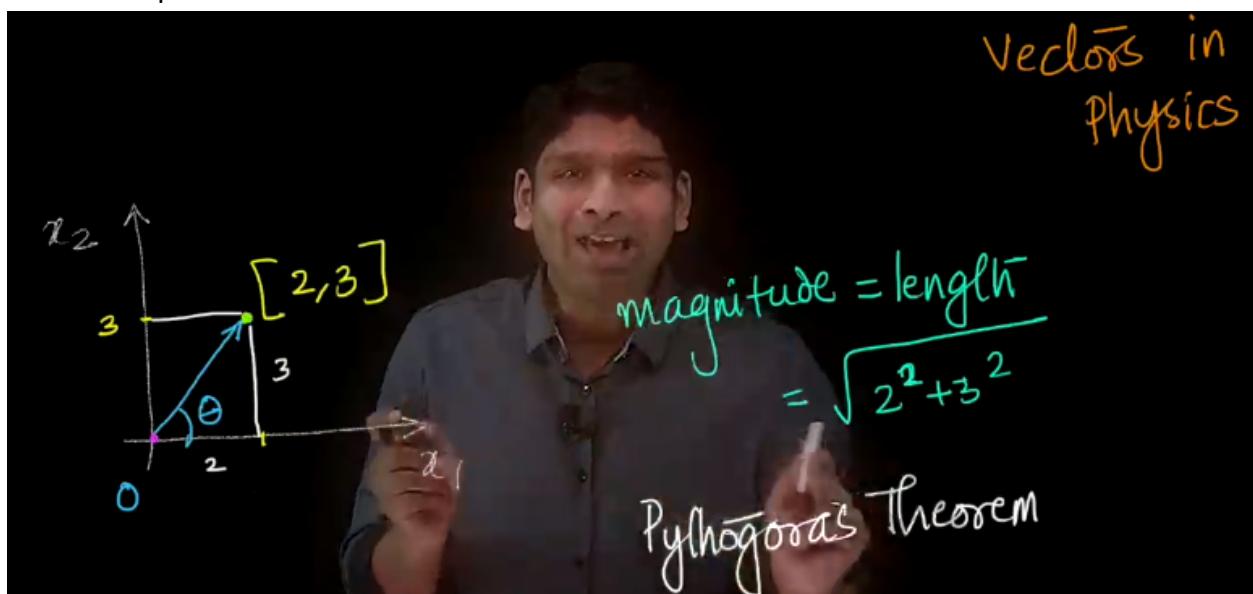
2.9 Dataset using vectors

At timestamp 2.11

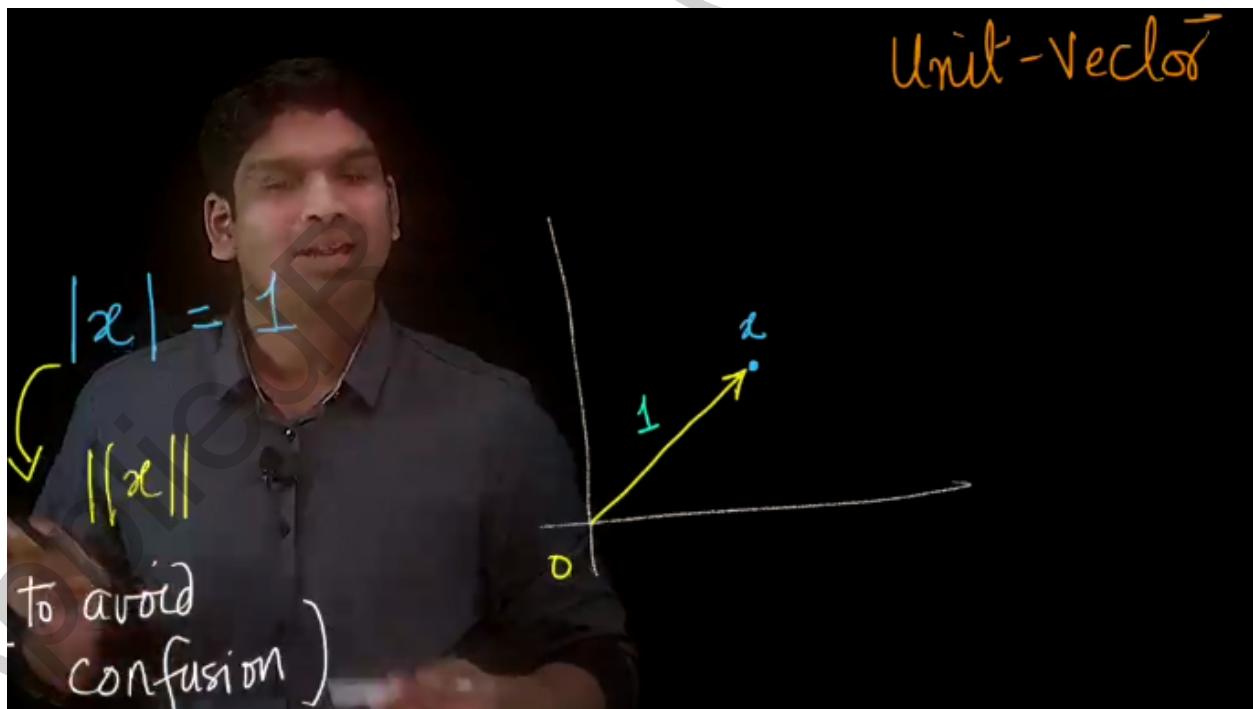


- We can use vectors to represent our dataset. Suppose we have two features to represent a fish(length,width) . Then we can represent a fish using the two values as a vector(fish is represented a point).We can think of d dimensional points as vectors.similarly we can represent a fish with 3 features.
- The first feature is 2 units away from x_1 axis,3 units away from x_2 axis and 1 unit away from x_3 axis.so we can represent a point in 3 D space and that point itself can be represented using a vector from origin to that point.

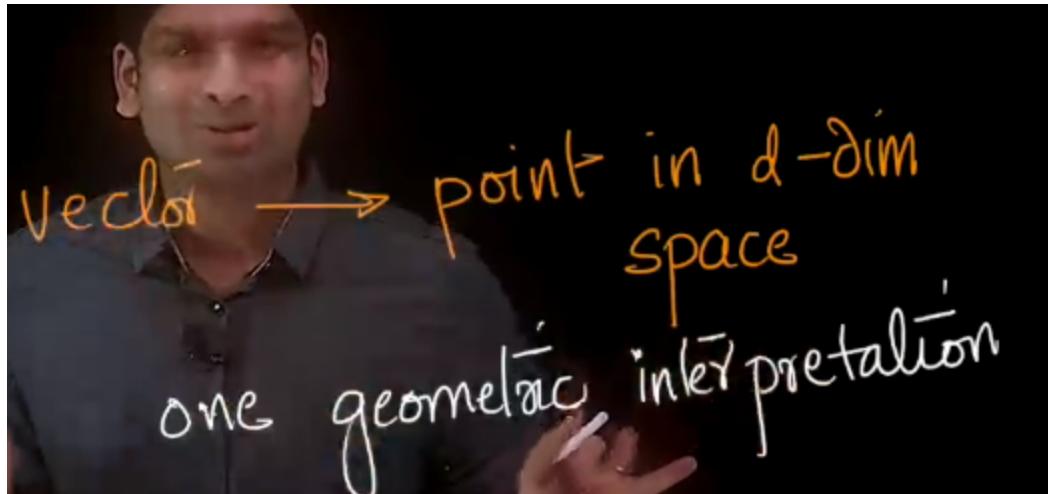
At timestamp 4.51



- Vectors in physics can be represented using magnitude and direction, we get the direction by obtaining the angle made by the vector with x-axis .
- Magnitude is nothing but length of the vector and can be calculated as shown above.\



- The magnitude of a vector is represented as shown above.
- A vector of magnitude or length 1 is a unit vector



- We can represent a vector as a point in d dimensional space.

At timestamp 11.45

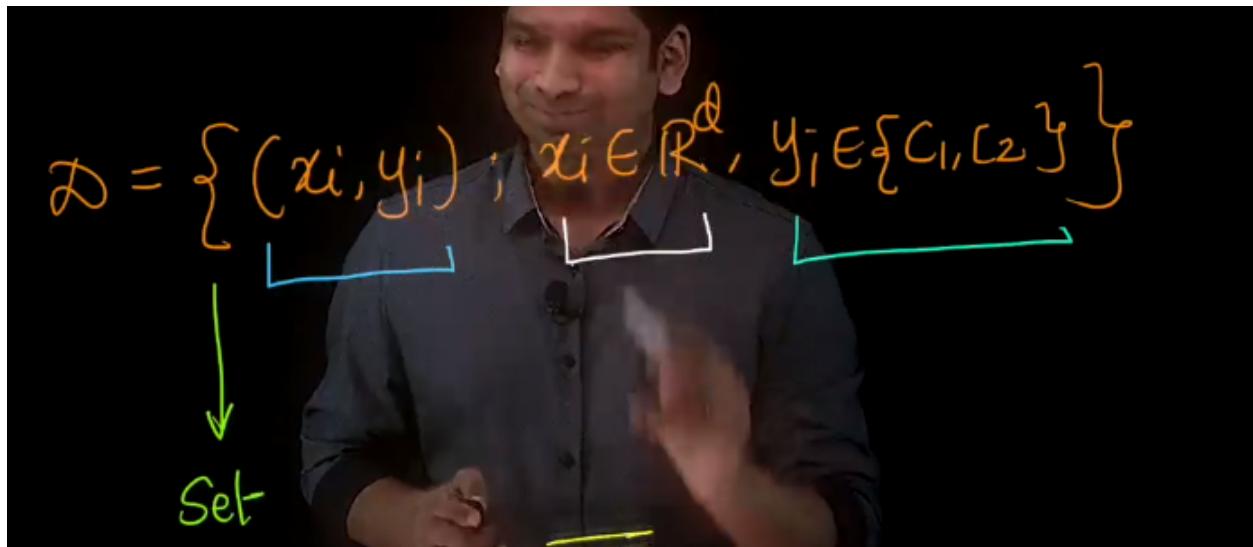
f_1	f_2	f_3	f_4	f_5	f_6	Type
		x_1				y_1
		x_2				y_2
						y_n

$x_i \in \mathbb{R}^6 \quad \forall i: 1 \rightarrow n$

$y_i \in \{C_1, C_2\}$

- Let our fish dataset has 6 features to represent each fish. So each row of our data set is a d dimensional row vector. The class to which the fish belongs to is given as type y_1, y_2 . Here we have two classes.
- You can see that x, y is data related to each fish
- Each x can be represented as a 6 dimensional vector.

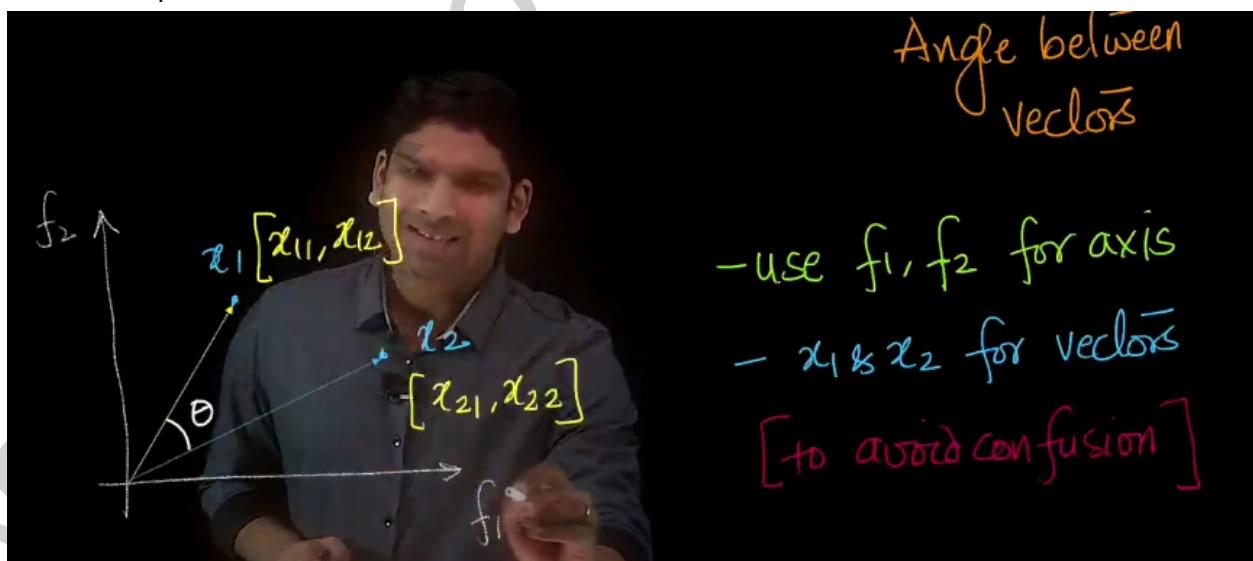
At timestamp 11.53



- We can represent our dataset mathematically as shown above .

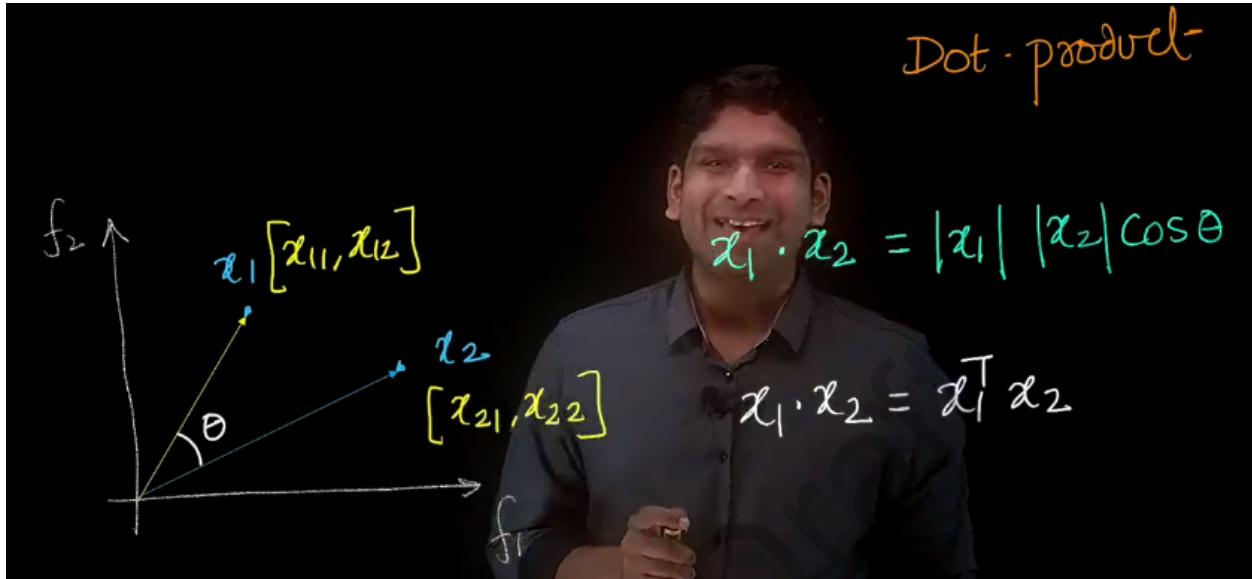
2.10 Dot product and angle between vectors

At timestamp 0.15



- In above snapshot f_1, f_2 are axis and x_1, x_2 are vectors θ is the angle between the vectors.

At timestamp 4.26



- Dot product is an operation between two vectors and is denoted as $x_1 \cdot x_2$ as shown above.
- For computing dot product between two vectors we need length of the two vectors and $\cos\theta$ (θ is angle between the vectors). Another way of interpreting dot product is if we have the vectors x_1 and x_2 (here we have two coordinates since 2D we can extend the same concept to nD as well) we can compute dot product by multiplying the vectors as shown.
- Given two vectors and their corresponding coordinates we can calculate $\cos\theta$ as shown below.

At timestamp 5.15 in video

$$\Rightarrow x_1^T x_2 = |x_1| |x_2| \cos \theta$$
$$\Rightarrow \begin{bmatrix} x_{11} & x_{12} \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} = \sqrt{x_{11}^2 + x_{12}^2} \cdot \sqrt{x_{21}^2 + x_{22}^2} \cdot \cos \theta$$

$$\Rightarrow \begin{bmatrix} x_{11} & x_{12} \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} = \sqrt{x_{11}^2 + x_{12}^2} \cdot \sqrt{x_{21}^2 + x_{22}^2} \cdot \cos\theta$$

$$\Rightarrow \frac{x_{11}x_{21} + x_{12}x_{22}}{\sqrt{x_{11}^2 + x_{12}^2} \cdot \sqrt{x_{21}^2 + x_{22}^2}} = \cos\theta$$

At timestamp 7.57

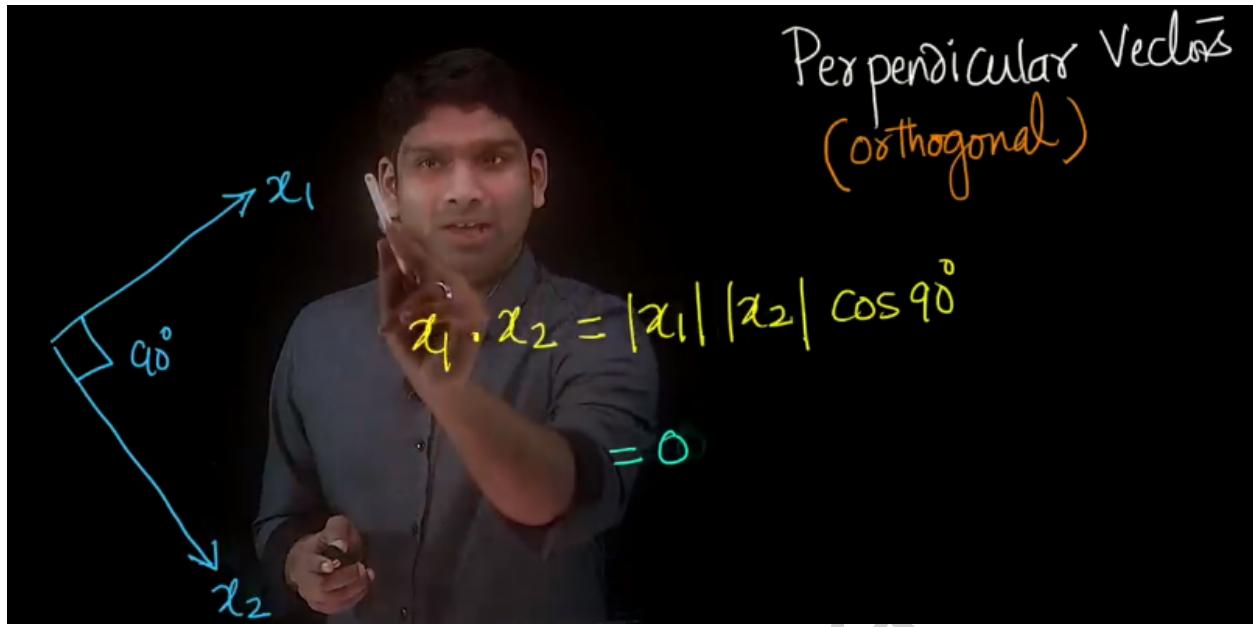
$$x_1 \cdot x_2 = |x_1| |x_2| \cos\theta$$

$$\cos\theta = \frac{x_1 \cdot x_2}{|x_1| |x_2|}$$

$x_1 \in \mathbb{R}^d$
 $x_2 \in \mathbb{R}^d$

- By equating and rearranging terms as shown above we calculated $\cos\theta$. (note that x_1 and x_2 should be of same dimensions because they both have to be in same d-dimensional space)
- What we learnt in 2D can transform it to higher dimensional space easily.

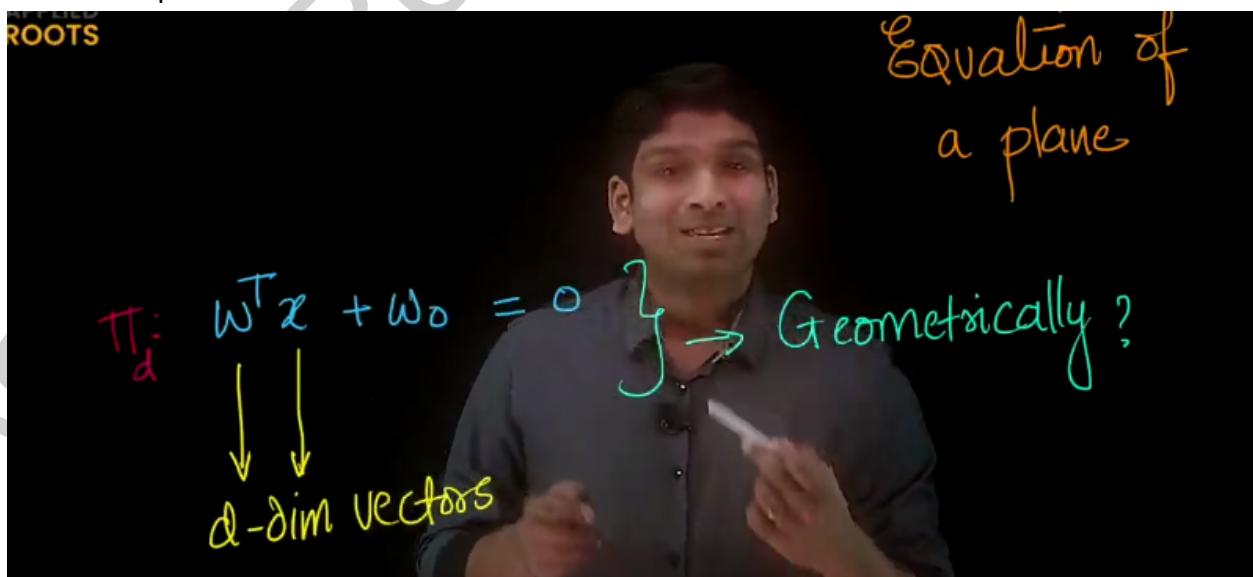
At timestamp 9.26



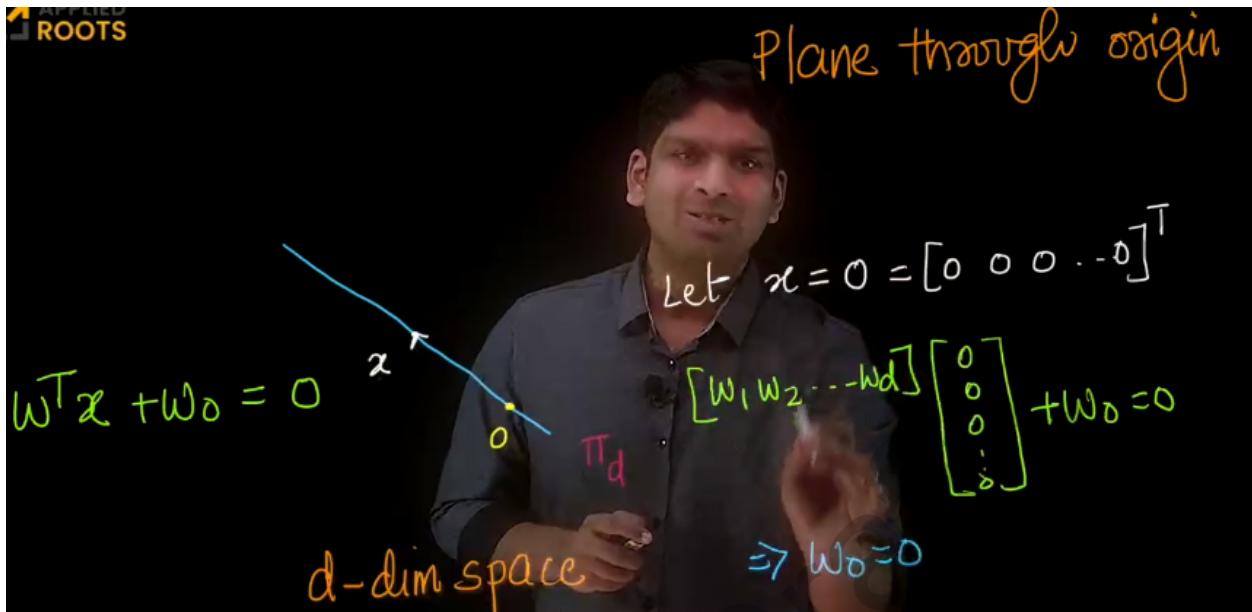
- When two vectors are perpendicular to each other then they are called Orthogonal vectors. (angle between the vectors is 90 degrees)
- The dot product of two perpendicular vectors is 0.
- If the dot product between two vectors is 0 we can conclude that they are Orthogonal vectors.

2.11 Geometry of the Equation of a plane

At timestamp 0.41 in video

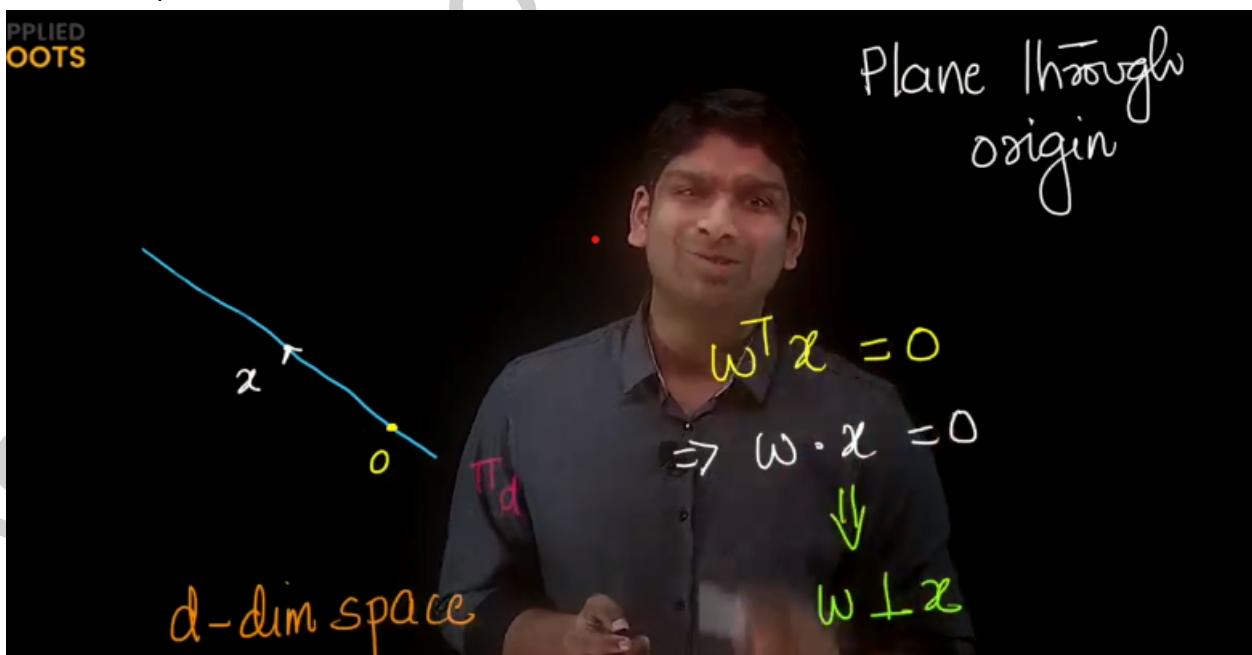


- Just like we understood the geometry of a line we can understand a plane geometrically.
- At timestamp 3.32 in video



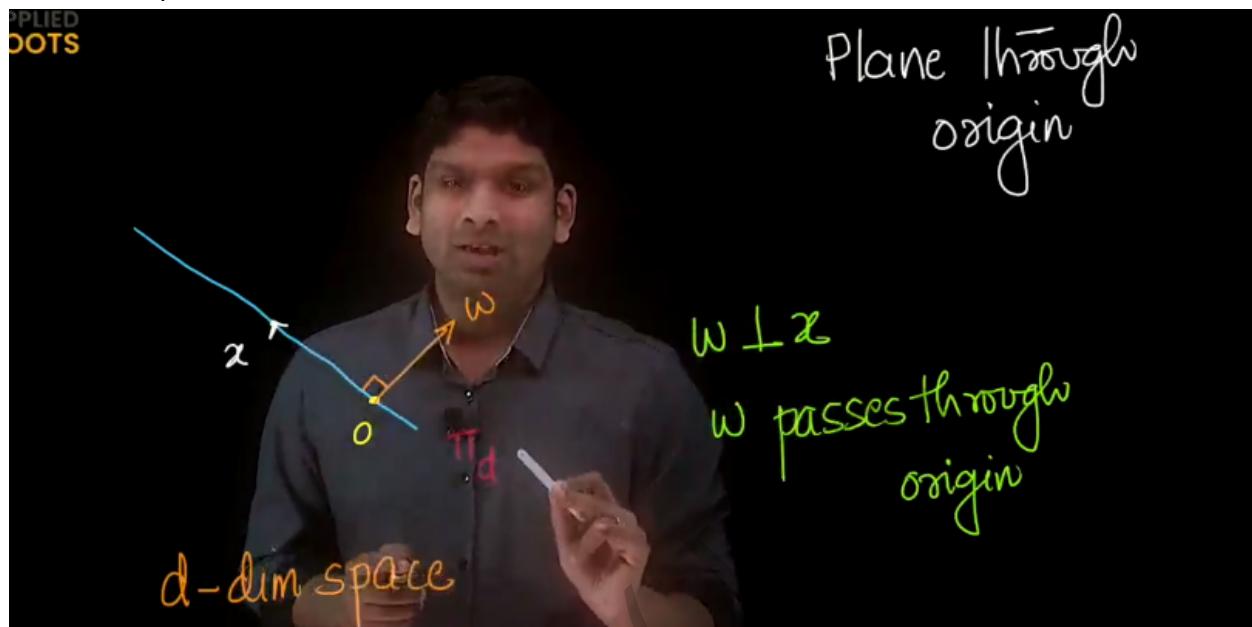
- Let's consider a plane passing through origin and the equation of plane is as shown above. This plane can be in any dimensional space if $d=2$ it's a line, if $d=3$ it's a plane, if $d>4$ they are all hyperplanes.
- Since our plane passes through origin we substitute $x=0$ in the equation, which gives us $w_0=0$. This implies when a plane passes through origin our $w_0=0$.

At timestamp 4.40 in video

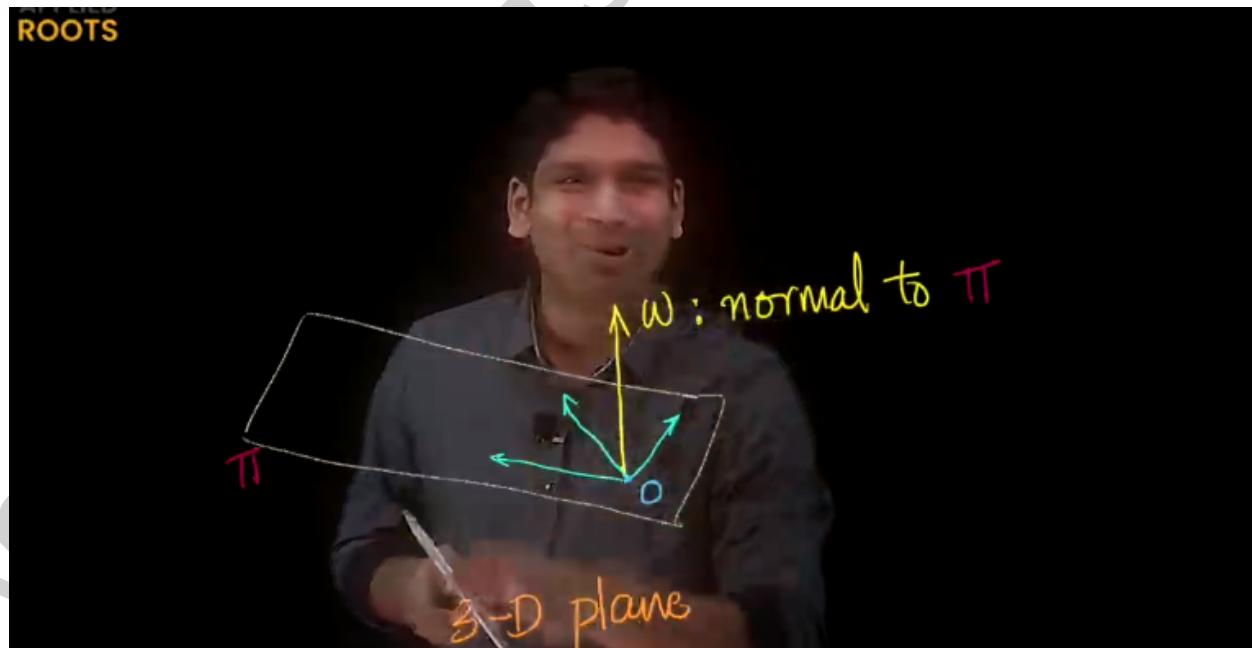


- For a plane passing through origin since $w_0=0$. The dot product of vectors $w \cdot x$ is 0, so they are perpendicular vectors.(angle between x and w is 90 degrees)
- w will be perpendicular to every point x on the plane.

At timestamp 5.50 in video

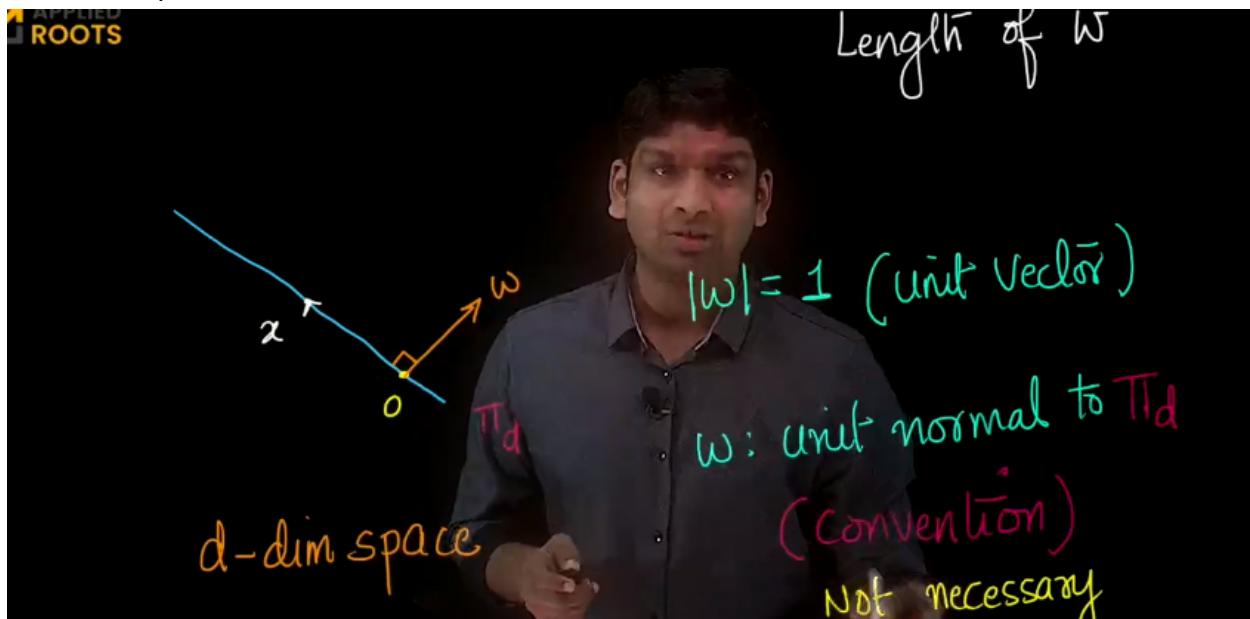


- w will be perpendicular to plane as shown above.



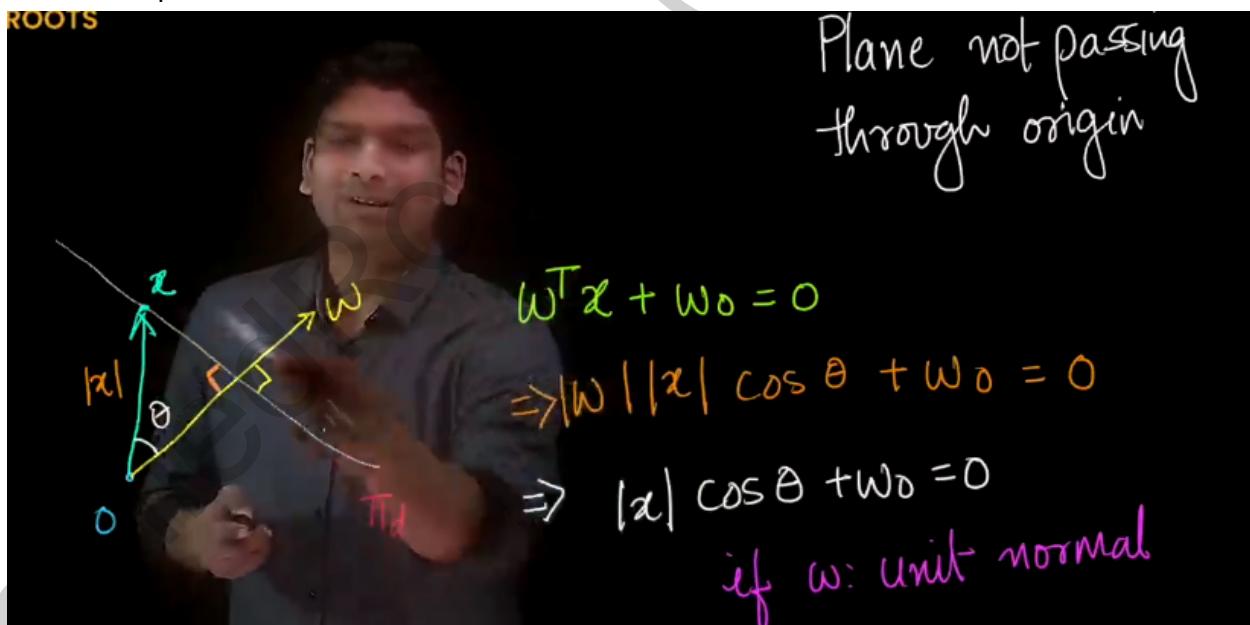
- You can visualize w in 3D as shown above.

At timestamp 8.33 in video



- By convention whenever we use a normal vector to a plane we typically use a unit vector but it is not necessary.
- When the length of vector $w=1$ we refer to it as the unit normal vector to plane.

At timestamp 13.8 in video



- For a plane not passing through origin the equation of plane is as shown above. w is normal vector to the plane. Lets consider a point x on the plane .(ox is a vector and w is our normal vector)

- Now we assume w to be unit vector and find the dot product between vectors x and w as shown above and $\cos\theta$ is not equal to 0. w_0 will not be equal to 0 since plane is not passing through origin

At timestamp 15.8 in video

Plane not passing through origin

$$|a| \cos \theta + w_0 = 0$$

$$\cos \theta = \frac{a}{|a|}$$

$$\Rightarrow a = -w_0$$

- We can calculate $\cos\theta$ as shown above and we can clearly see that $a=-w_0$.

At timestamp 16.26 in video

Plane not passing through origin

$$a = -w_0$$

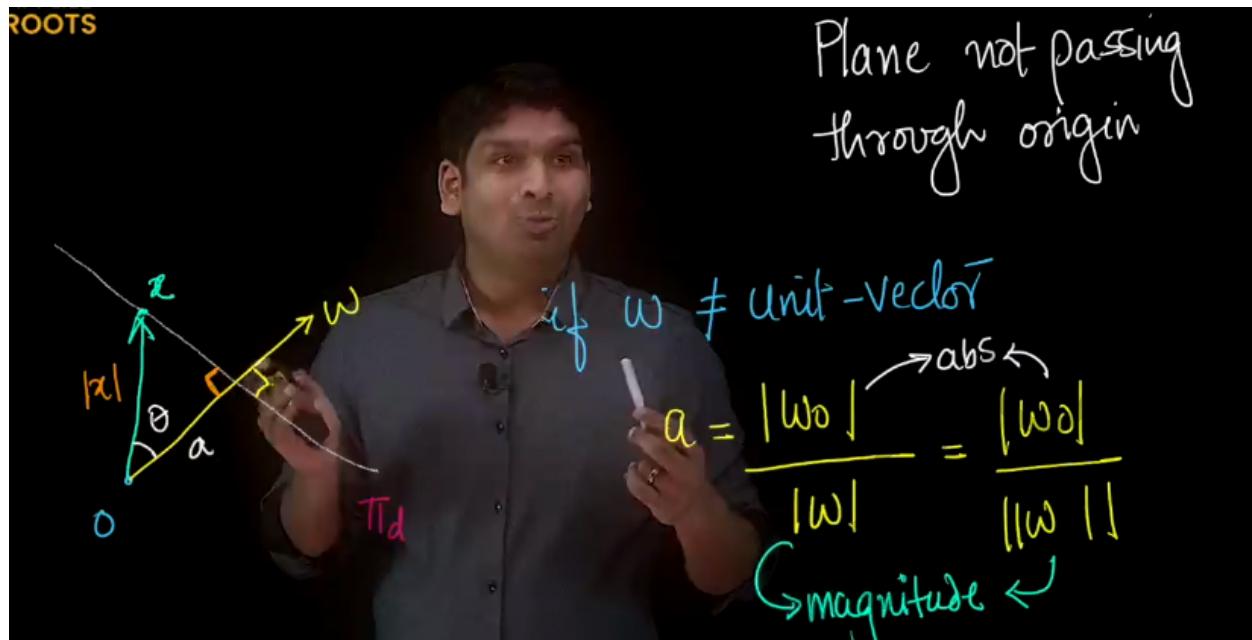
if w : unit-normal vector

$$w_0 = -2 \text{ (rr)} \quad w_0 = 2$$

$$\text{dist}(O, \Pi) = |w_0|$$

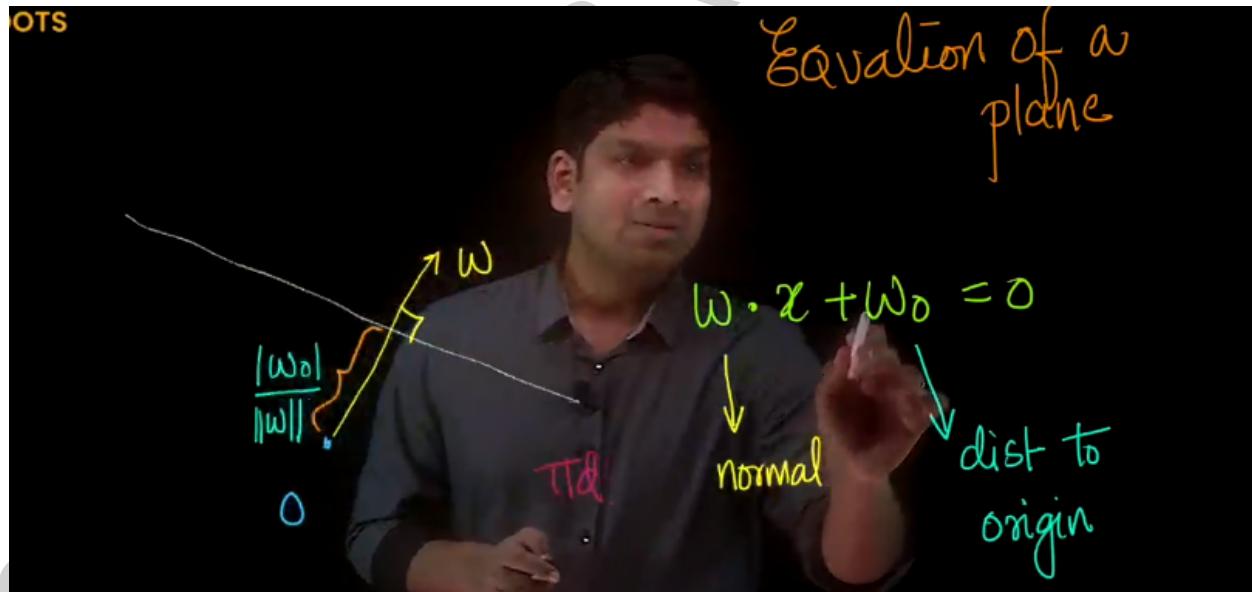
- a is nothing but the shortest distance from origin to the plane and we know that distance cannot have negative values (-ve represents the direction). So, the distance from origin to the plane is $\text{abs}(w_0)$

At timestamp 17.11 in video



- For plane not passing through origin and w is not a unit vector then a can be written as shown above.

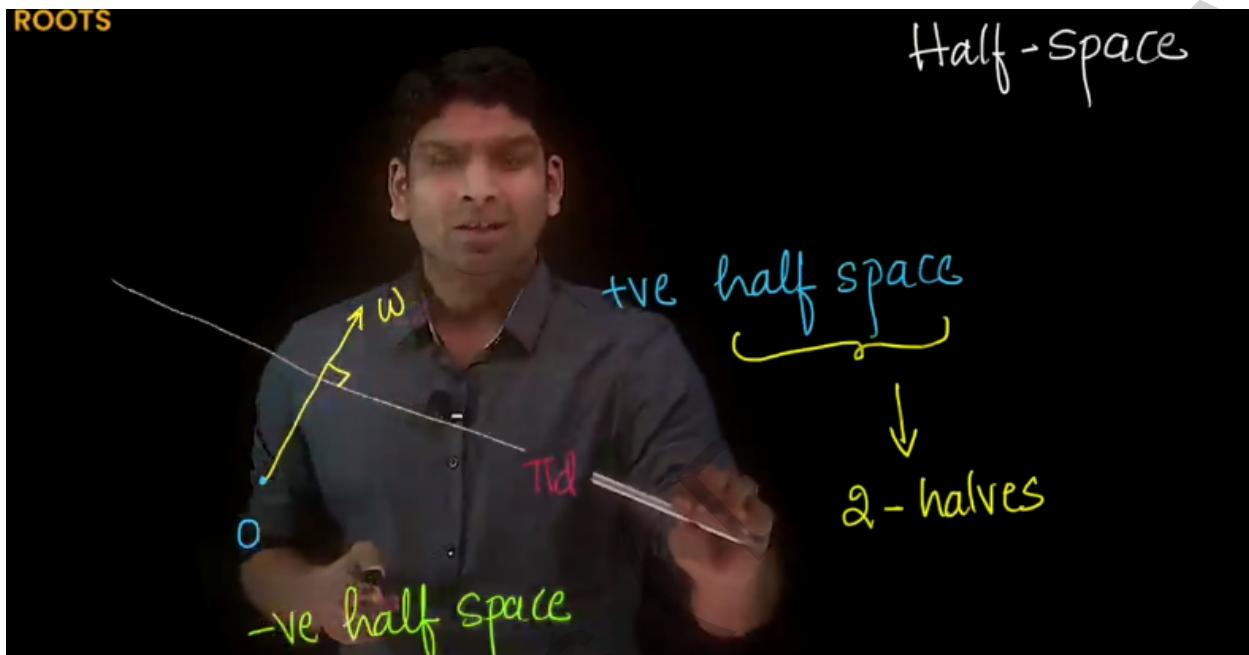
At timestamp 21.53 in video



- Given the equation of plane w is vector normal to the plane and passing through origin,w0 is the distance from origin to the plane.

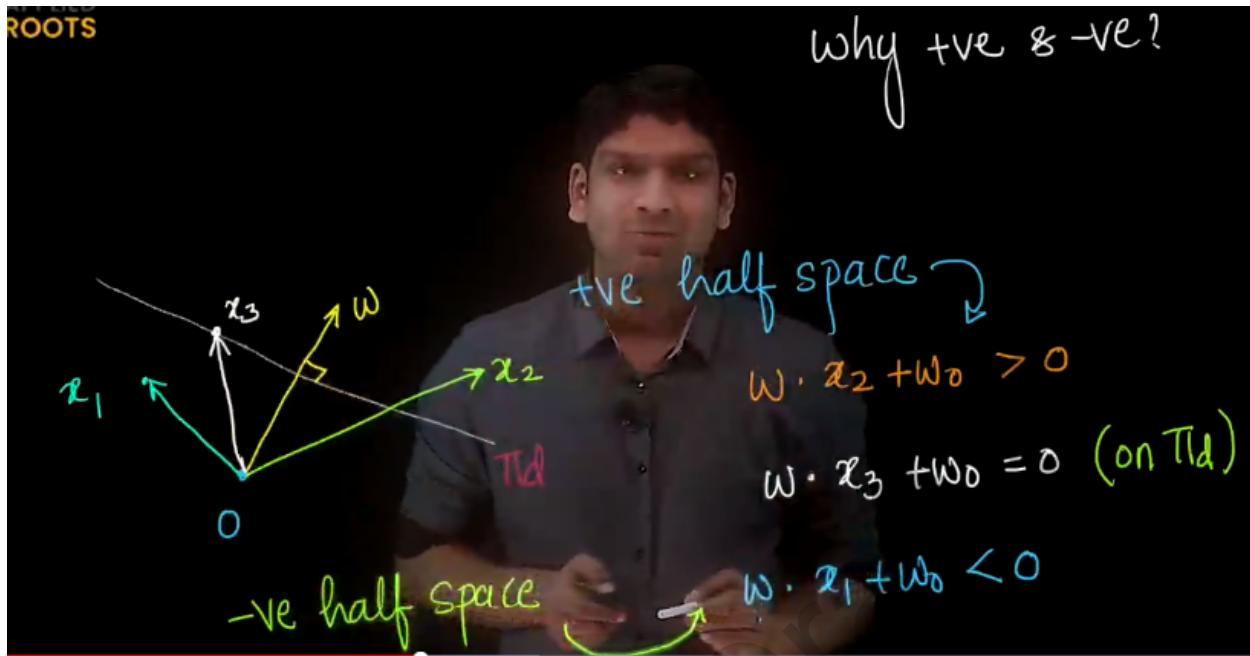
2.12 Half spaces & classification using a plane

At timestamp 0.42 in video



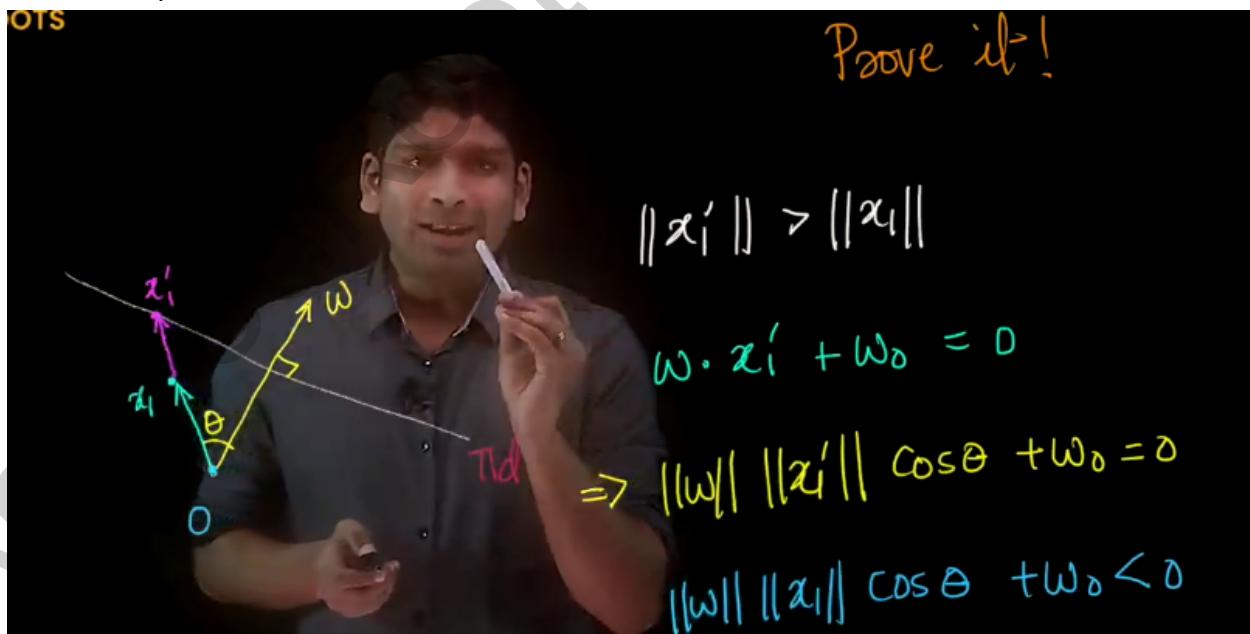
- Now we use the concept of half-spaces to actually classify the points in d dimensional space.
- Given a plane in d-dimensional space and normal vector w as shown above, if we assume it to be a 2D space the line divides the total space into 2 halves everything above the plane and everything below the plane. (+ve half and -ve half). likewise a plane separates the 3D space into two regions.
- The space above the plane in the direction of w is positive half space and space below the plane opposite direction of w is negative half space

At timestamp 4.40 in video



- Let's consider three points x_1 below the plane, x_2 above the plane and x_3 on the plane.
- As shown above for x_2 and all the points above the plane $w \cdot x_2 + w_0 > 0$, and for points below the plane the value is less than 0 and for points lying on the plane it is 0.
- Let's understand each scenario one by one.

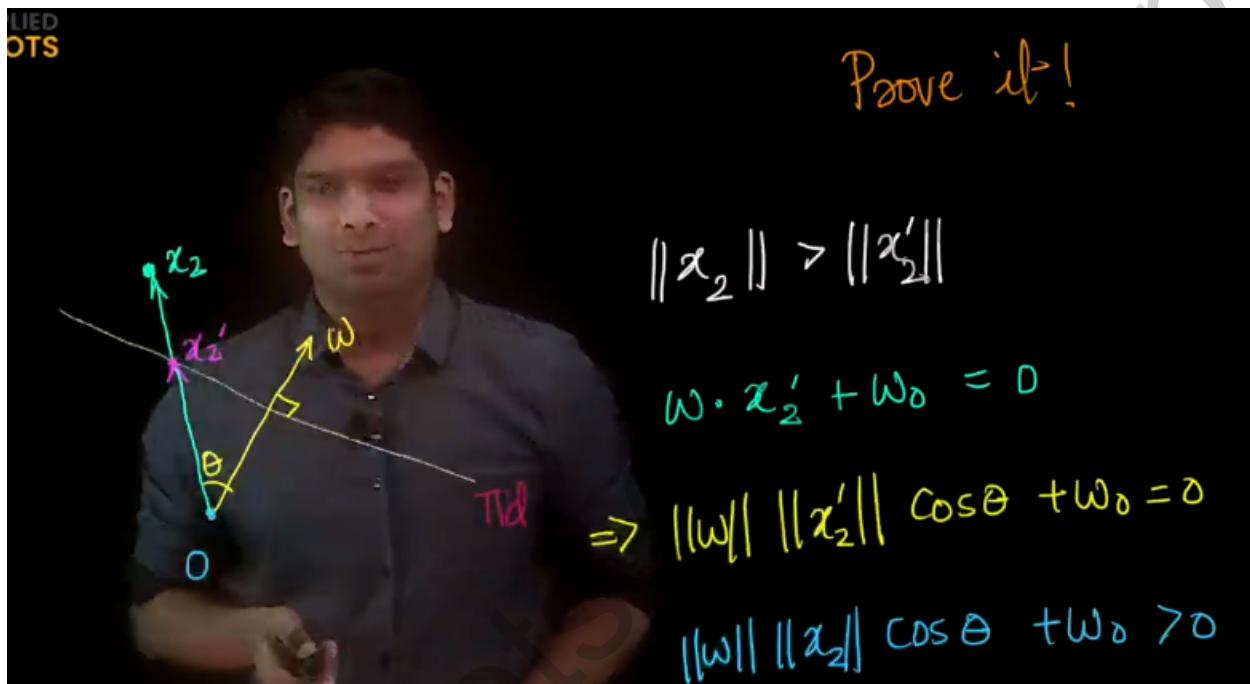
At timestamp 7.27 in video



- As shown above we have our plane and normal vector to the plane w and we have point x_1' below the plane in negative half space.

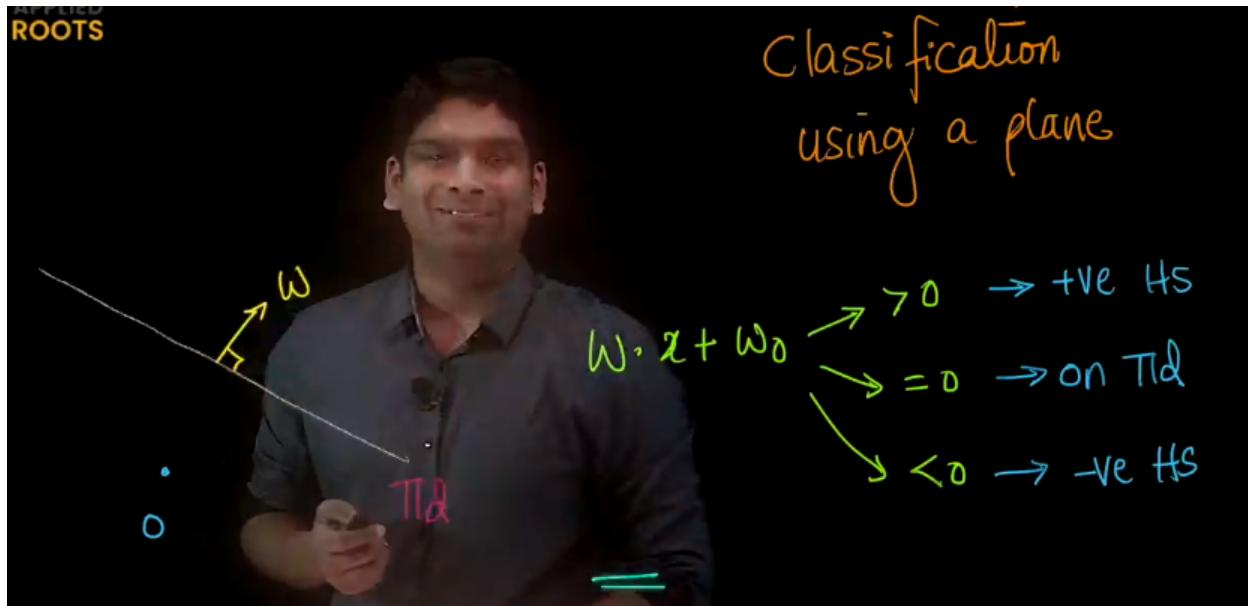
- We extend the vector x_1 slightly till it touches the plane and we call the point x_1' . Now we have two vectors x_1 and x_1' .
- As shown above the magnitude of x_1' is greater than the magnitude of x_1 .
- We know that $w \cdot x_1' + w_0 = 0$ since x_1' lies on the plane. From above equations it's clear that $w \cdot x_1 + w_0 < 0$

At timestamp 8.50 in video



- We now consider a point x_2 in positive half space, let's call the point at which the vector x_2 intersects the plane be x_2' . Now we have two vectors x_2 and x_2' .
- From the above diagram it's clear that length of x_2 is greater than length of x_2' .
- We know that $w \cdot x_2' + w_0 = 0$ since x_2' lies on the plane. From above equations it's clear that $w \cdot x_2 + w_0 > 0$

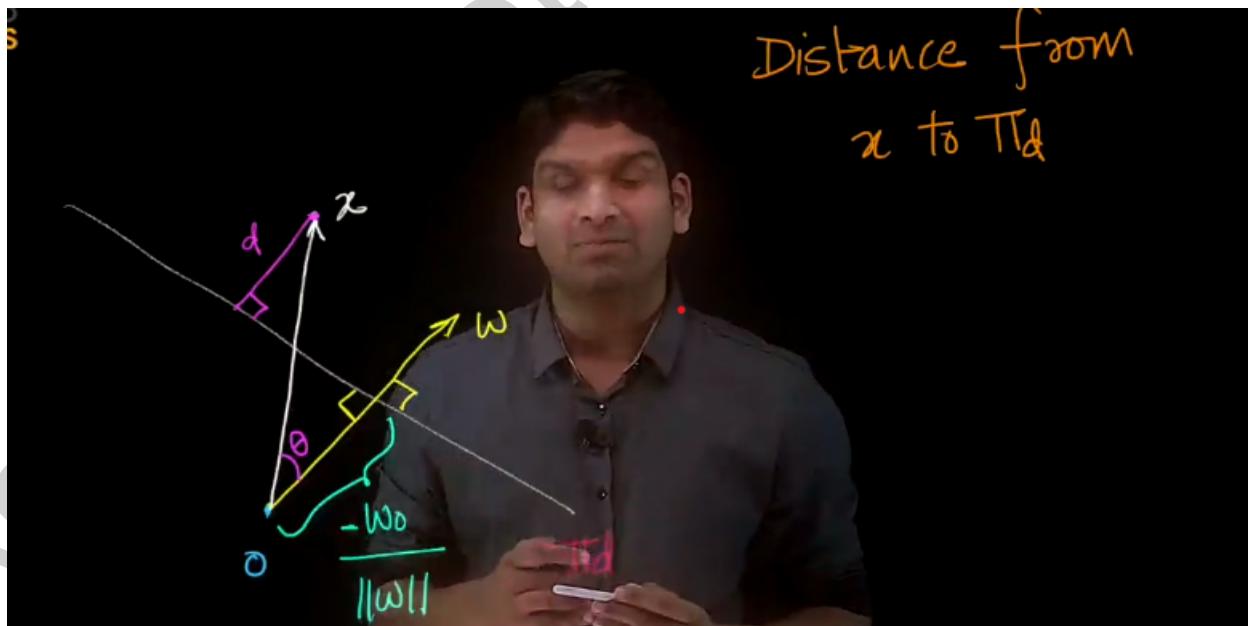
At timestamp 11.53 in video



- To summarize we perform classification using the plane based on above conditions.
- When we use a plane to separate any point in d -dimensional space ,we determine whether the point is above or below the plane.

2.13 Distance of a point from a plane

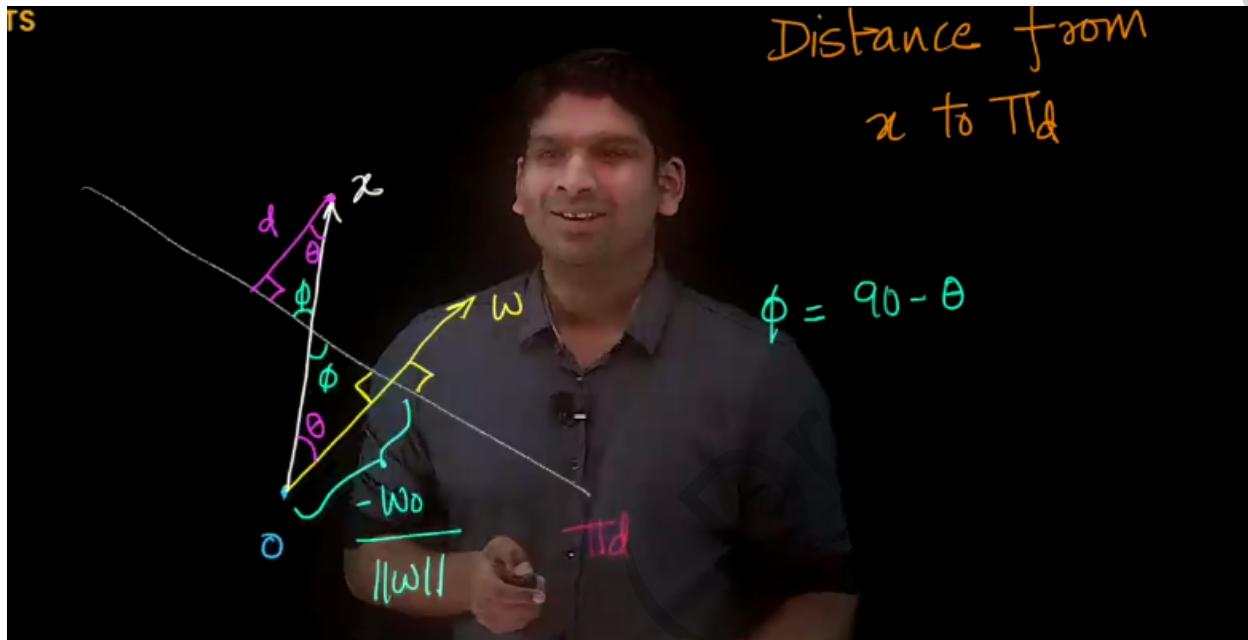
At timestamp 0.13 in video



- Given any point we will find which side of the plane it belongs to and also its important to know how far away it is from the plane.

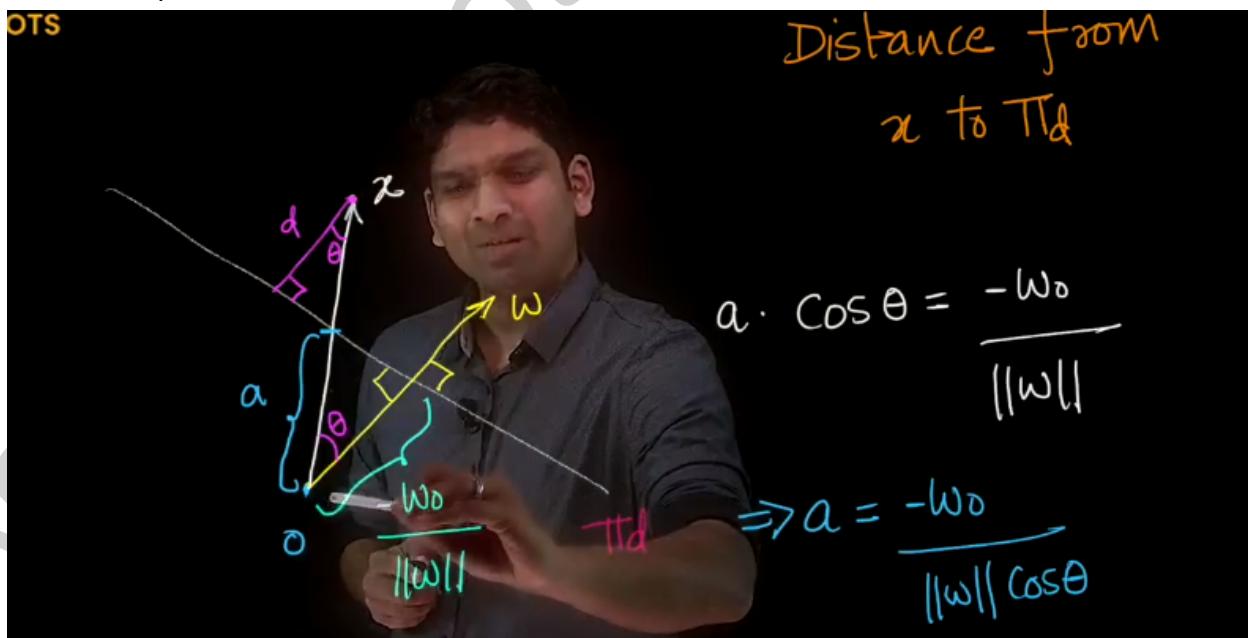
- Given any plane and a point x , we know the normal vector w and also the distance from origin to the plane as shown above. Let's find the distance from point x to the plane
- Distance from plane to the point x is nothing but the length of the line segment that is perpendicular to plane and passing through the point x . (d as shown above)

At timestamp 14.12 in video



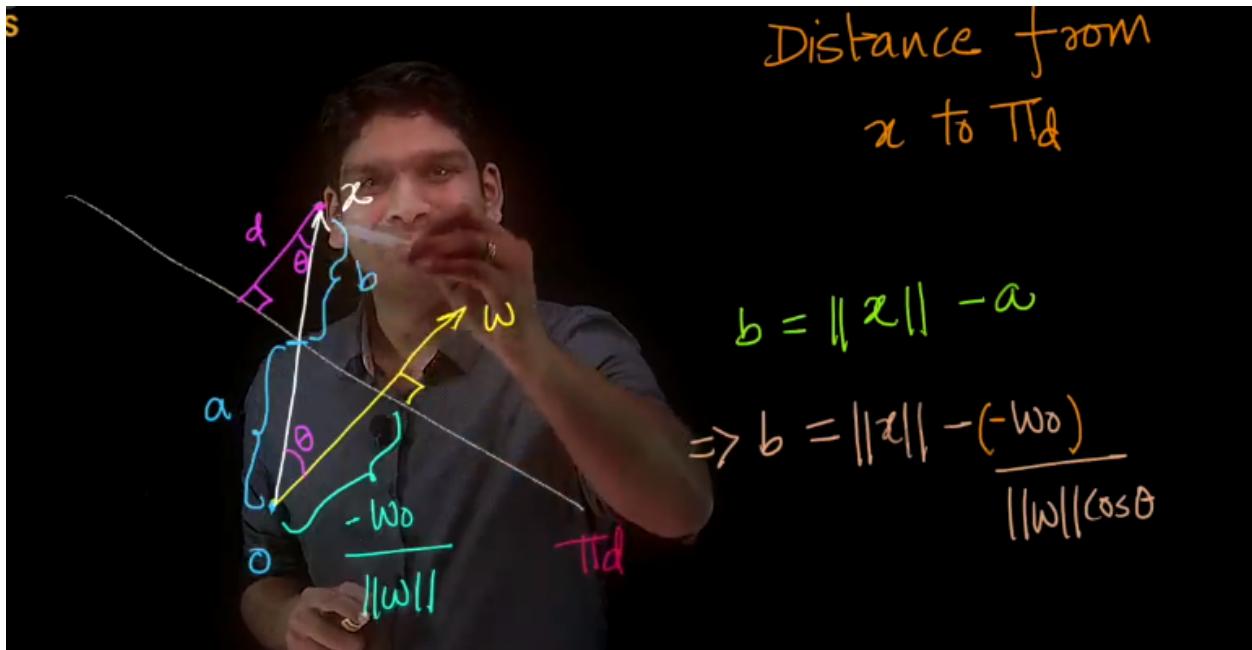
- From the above diagram we can say that $\phi = 90 - \theta$

At timestamp 4.49 in video



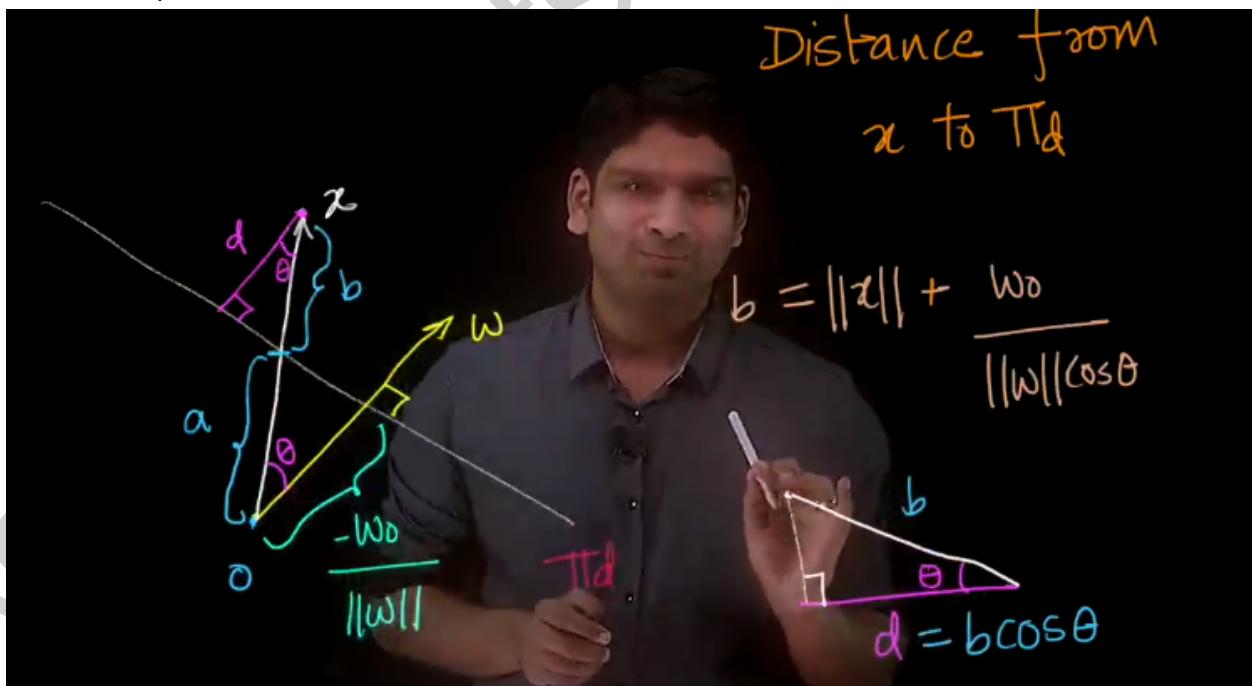
- From the triangle shown above we can calculate a .

At timestamp 5.51 in video



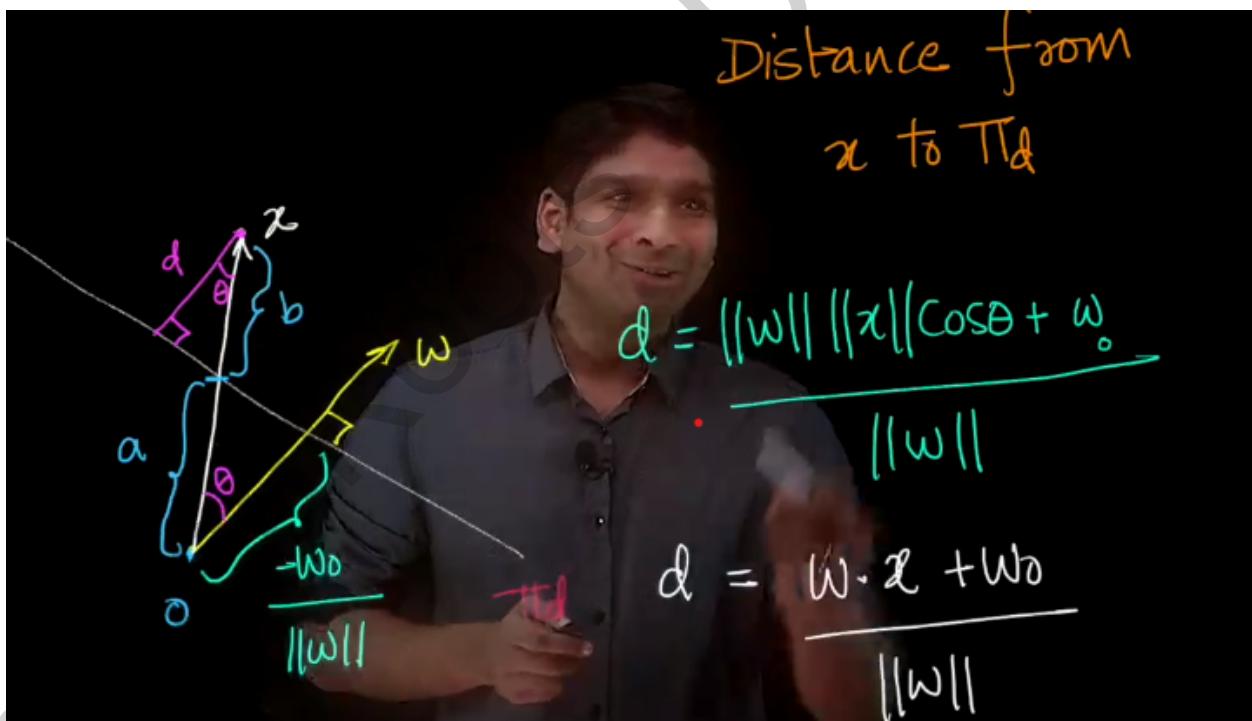
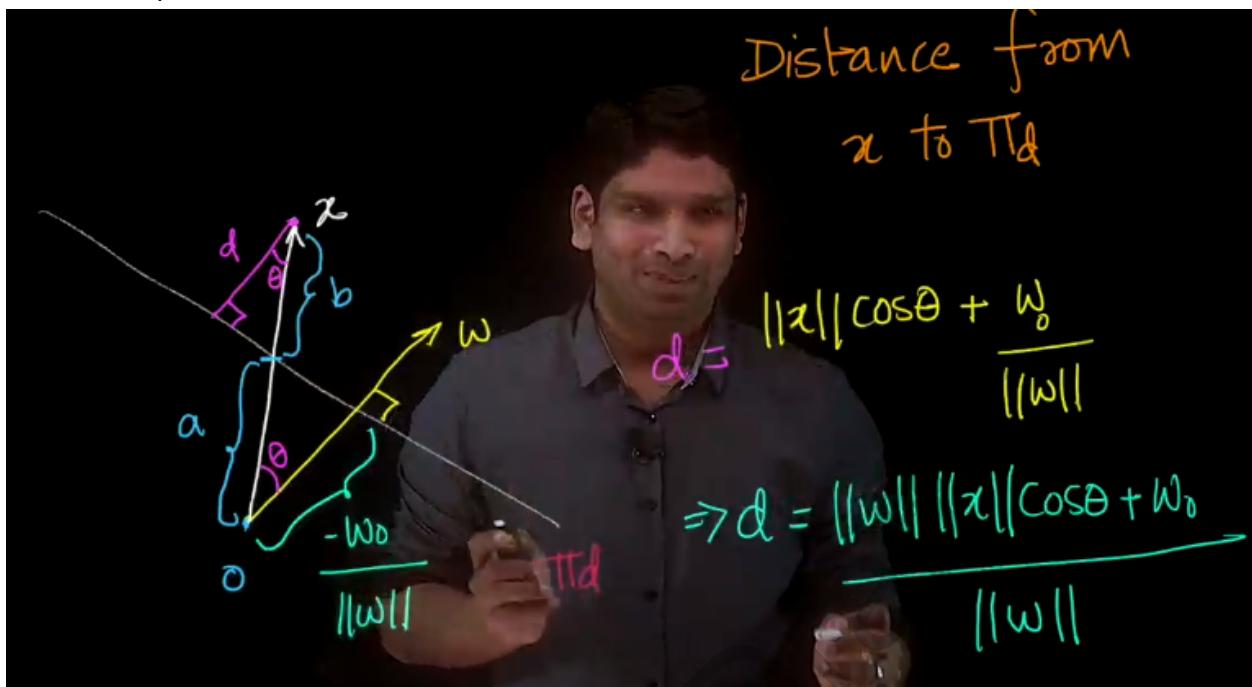
- By subtracting a from the length of vector x we can get b as shown.

At timestamp 7.38 in video



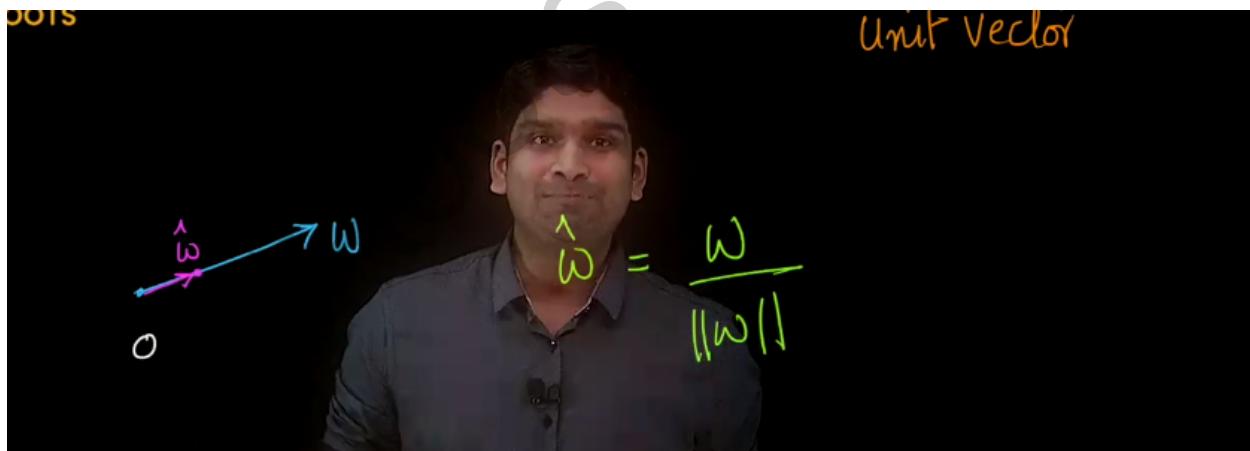
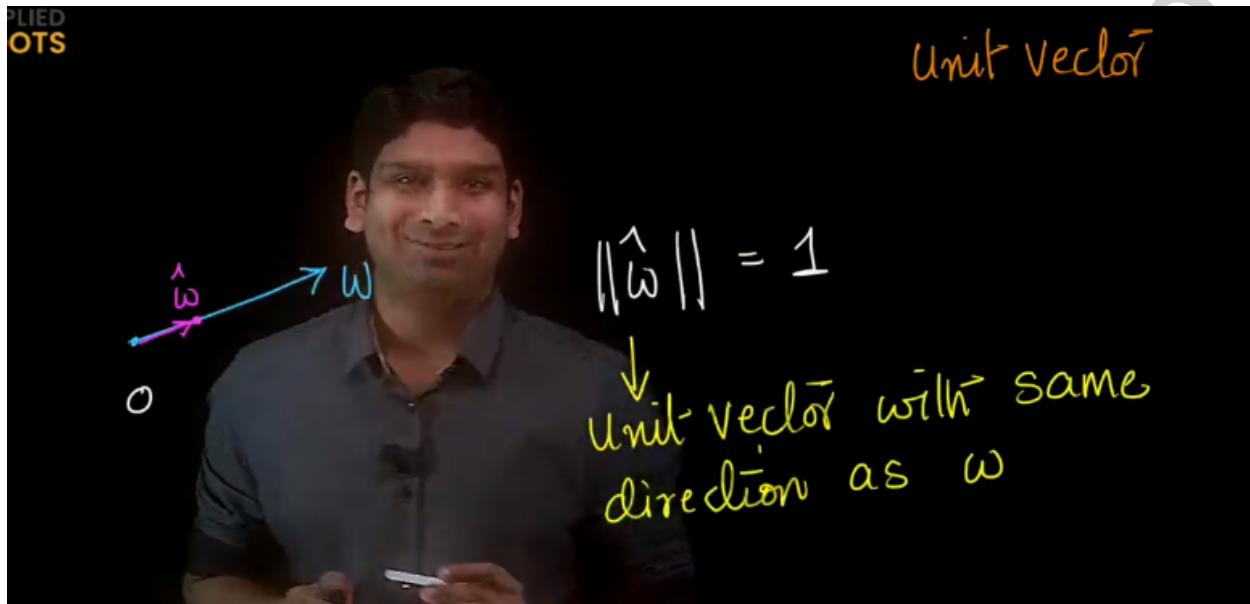
- We can calculate $d = b \cos \theta$ from the above triangle.
- We substitute b value in the above equation and get d value as shown below.

At timestamp 8.53

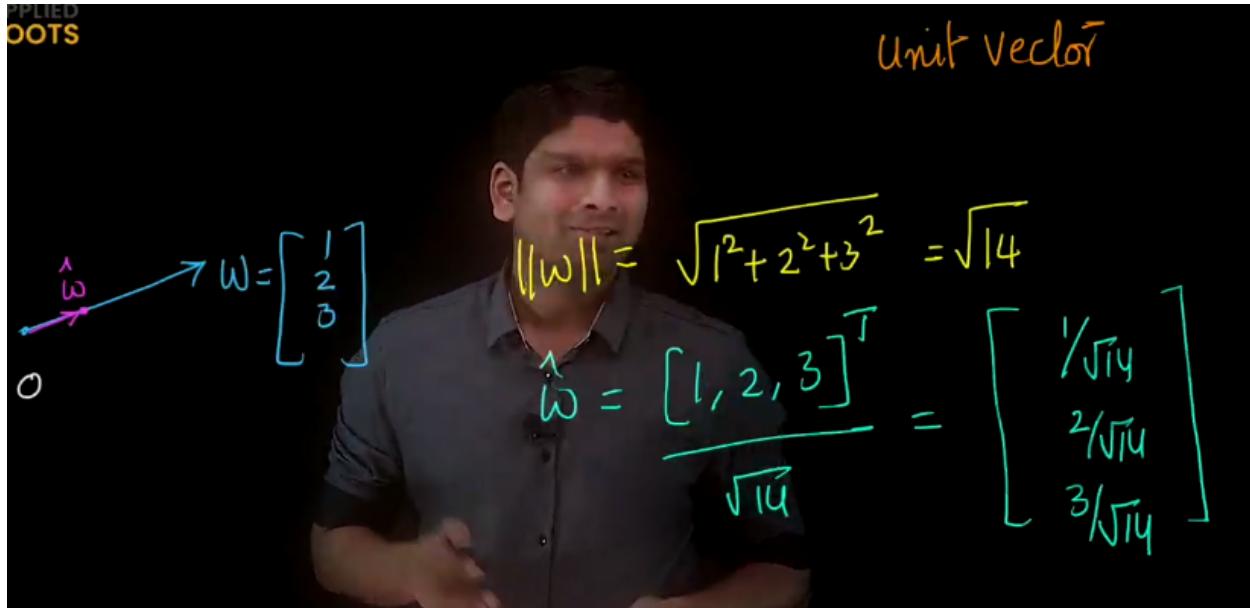


2.14 Projection & arithmetic operations on vectors

At timestamp 0.40 in video

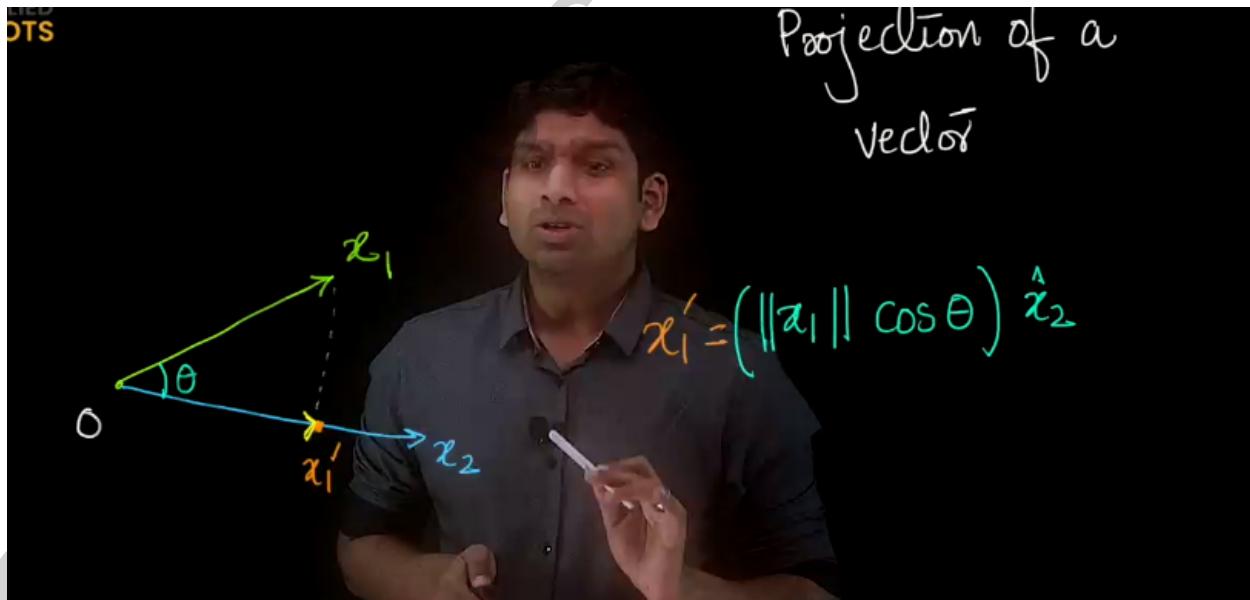


Applied
Mathematics



- Let we have a vector W and its length need not be 1 .If we want to represent a vector whose length is equal to 1 but in the same direction of W we call such a vector as a unit vector.
- When we are given a vector W , we can obtain a unit vector which has the same direction as W as shown above.

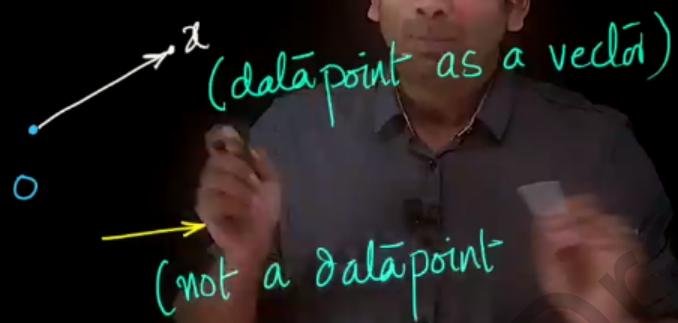
At timestamp 4.6 in video



- Let we have two data points x_1 and x_2 ,if we want to project vector x_1 on to x_2 we can draw a perpendicular line from x_1 onto x_2 and we call it x_1' as shown above.The direction of x_1' is the same as x_2 .

ROOTS

Vectors anywhere
in space



- Whenever we have a datapoint we draw a line from origin to the point and we call it a vector. But vectors can be anywhere like the yellow line shown above and we cannot interpret it as a datapoint.

At timestamp 8.47 in video

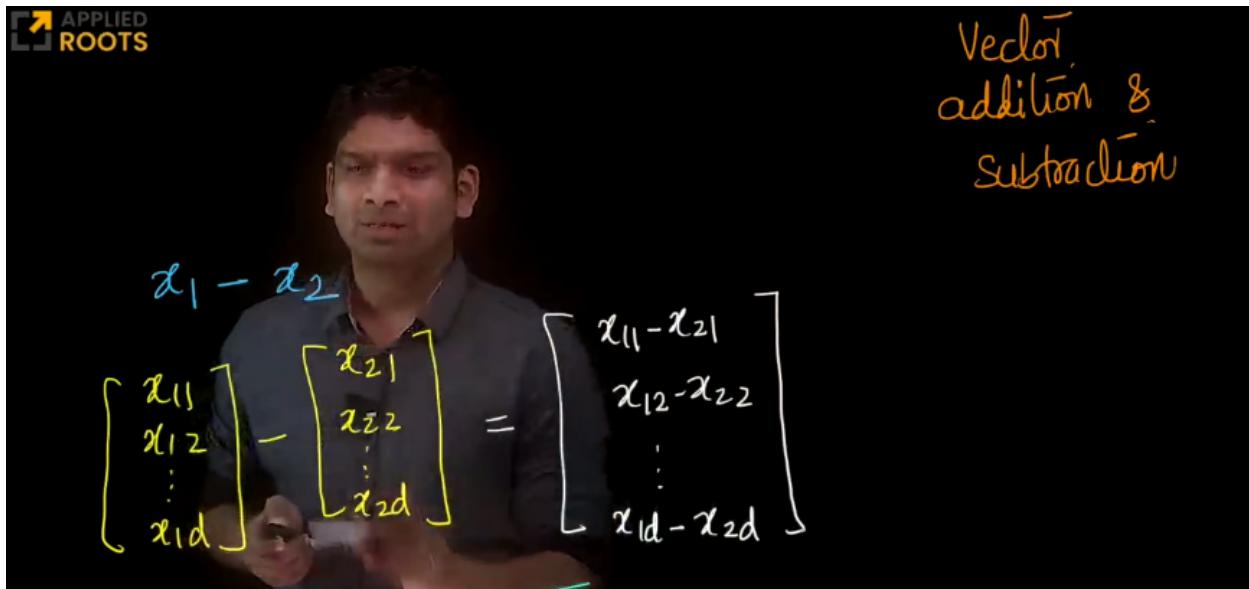
ROOTS

Addition of
vectors

$$\vec{OP} + \vec{PQ} = \vec{OQ}$$
$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$
$$\Rightarrow \vec{PQ} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

- We can perform vector addition as shown above.

At timestamp 10.56 in video



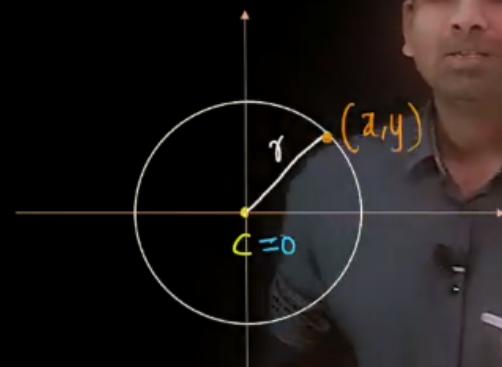
- Vector addition and subtraction be done even in d dimensions as shown .
- Note that for vector addition,subtraction and dot product the vectors should be of the same dimensions.

2.15 Equations of circle, sphere & hypersphere

At timestamp 0.37 in video

LIED
OTS

Circle (2D)

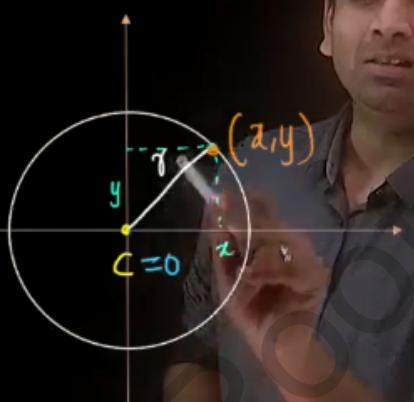


center = origin

r = radius

TS

Circle (2D)

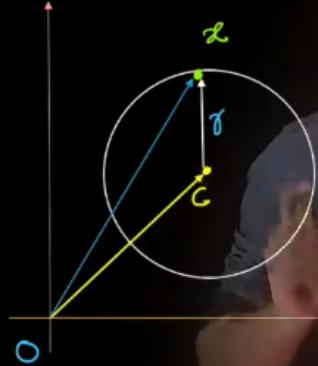


$$\text{dist} = \sqrt{x^2 + y^2} = r \\ \Rightarrow x^2 + y^2 = r^2$$

- Circle in 2D can be represented as shown above with radius r, centre c=0. To represent any circle we need its centre c and radius r.

At timestamp 4.52 in video

circle (2D)



$$c = [c_1, c_2]^T$$

$$\vec{oc} + \vec{cx} = \vec{ox}$$

$$\Rightarrow \vec{cx} = \vec{ox} - \vec{oc}$$

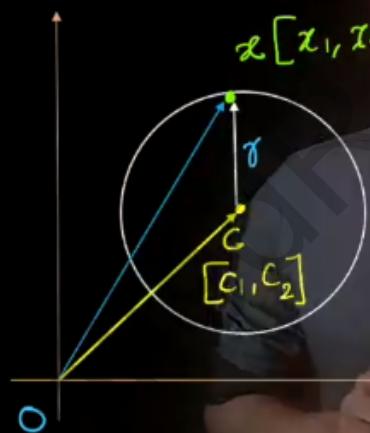
$$\Rightarrow \|\vec{cx}\| = \|\vec{ox} - \vec{oc}\|$$

X NOT $\|\vec{ox}\| - \|\vec{oc}\|$

- From the definition of circle the length from center of circle to any point x is always the same and it is r.
- Using algebra we derived $cx = ox - oc$.

At timestamp 5.33

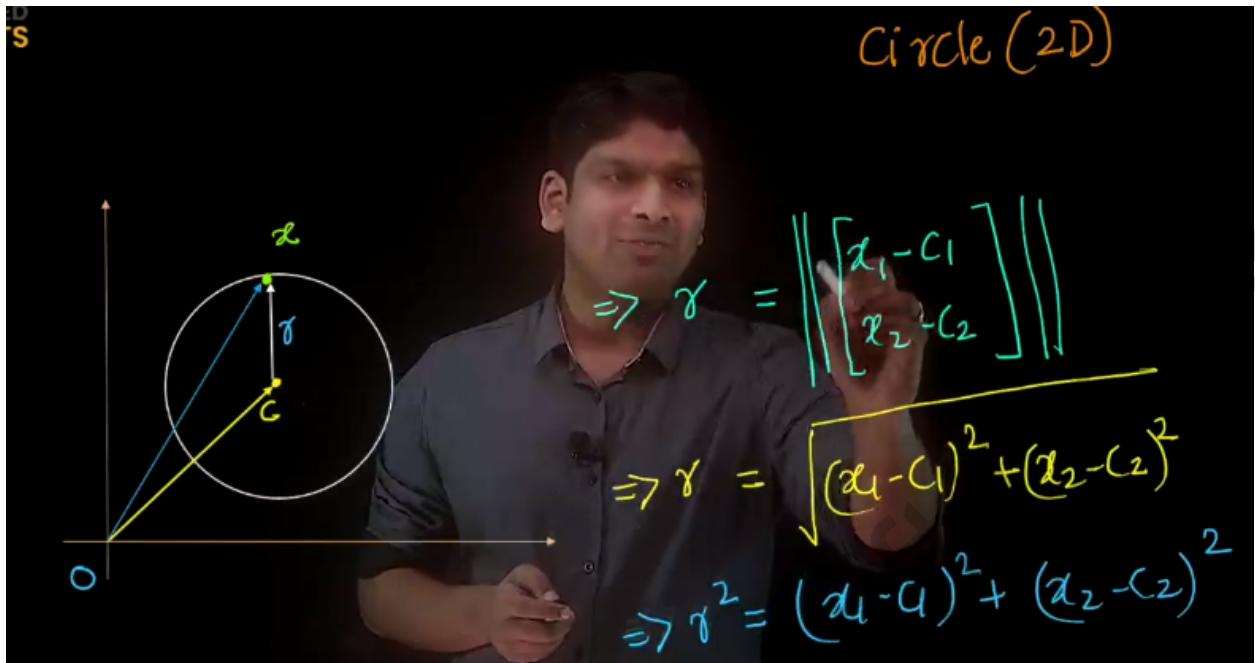
circle (2D)



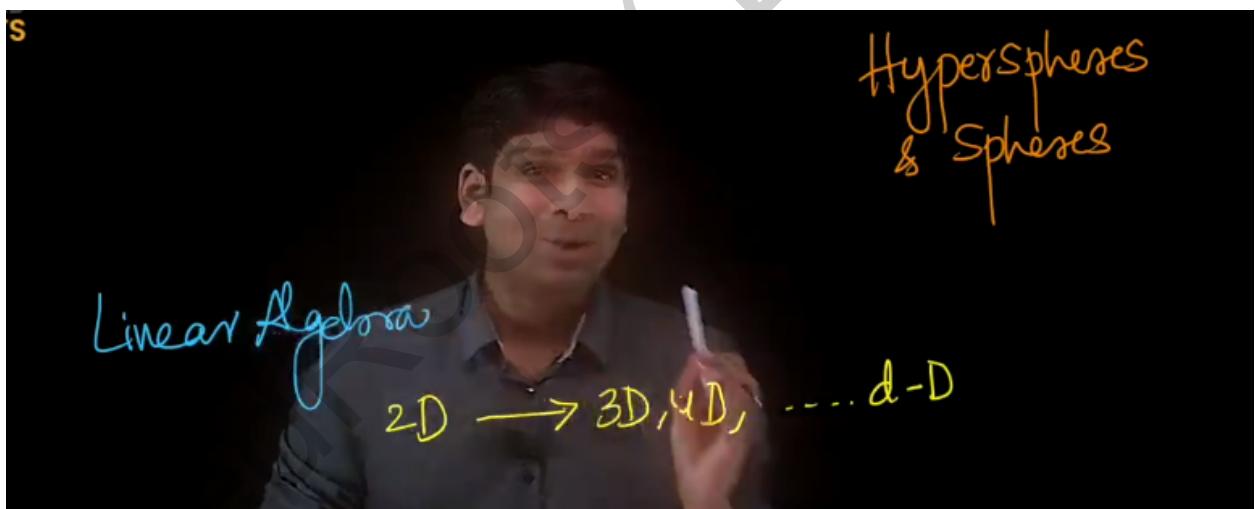
$$\|\vec{cx}\| = \|\vec{ox} - \vec{oc}\|$$

$$\Rightarrow r = \|x - c\|$$

$$= \left\| \begin{bmatrix} x_1 - c_1 \\ x_2 - c_2 \end{bmatrix} \right\|$$



- We already know that the length of cx is radius r. Coordinates of ox are (x_1, x_2) and coordinates of oc are (c_1, c_2) . By doing component wise subtraction of two vectors we get radius r as shown.

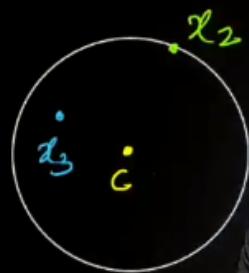


- Whatever we learnt in 2D can be extended to higher dimensions as well.

2.16 Classification using hypersphere

At timestamp 1.04 in video

Classification using Spheres



inside -outside

- Let's understand how we classify or data using spheres and hyperspheres. We assume three points x_1, x_2, x_3 lying outside, on and inside the sphere. Typically we use inside/outside as a mechanism to determine whether a point belongs to which class.

$$\|x_1 - c\| > r$$

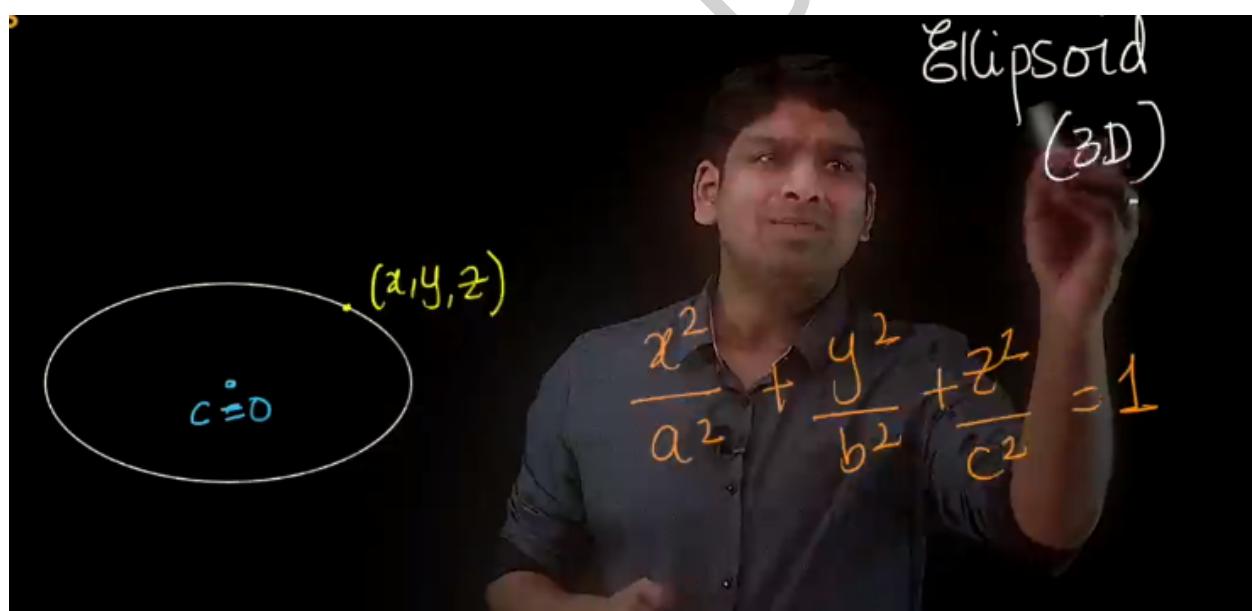
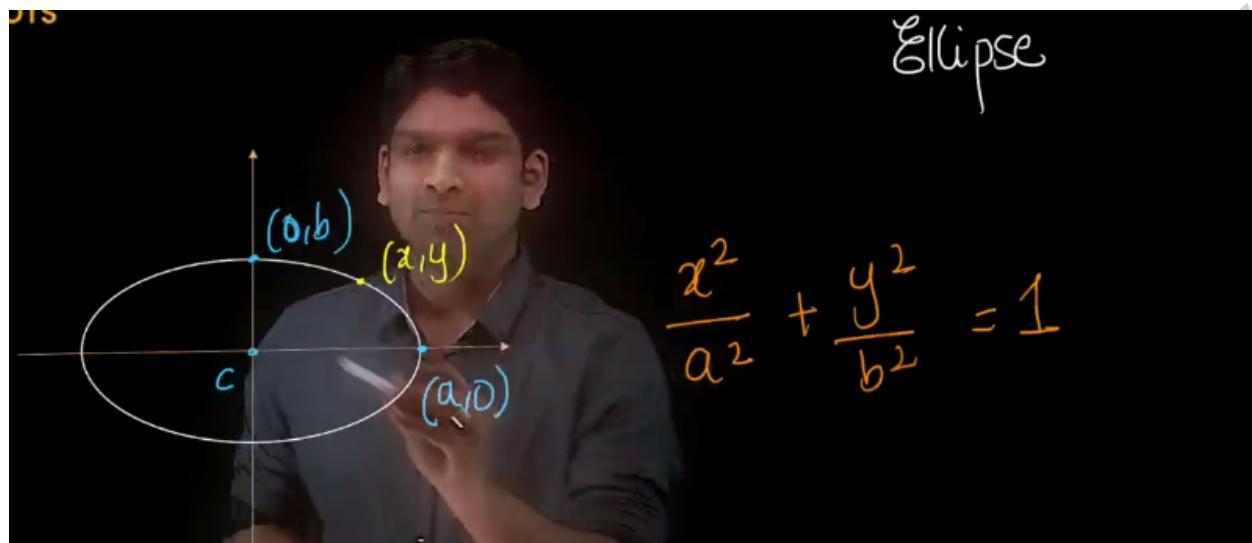
$$\|x_2 - c\| = r$$

$$\|x_3 - c\| < r$$

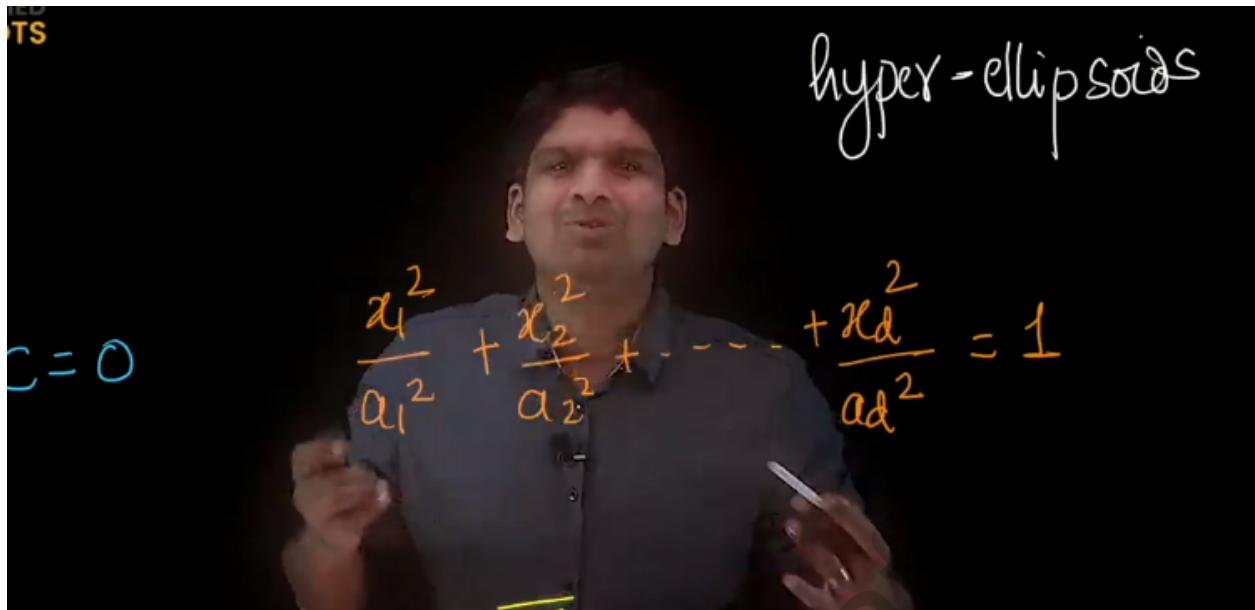
- We have to calculate the length of the vector from centre c to to determine where the datapoint lies. We have to check the conditions shown above.

2.17 Ellipse, Ellipsoid, Hyper-Ellipsoid

At timestamp 1.18 in video



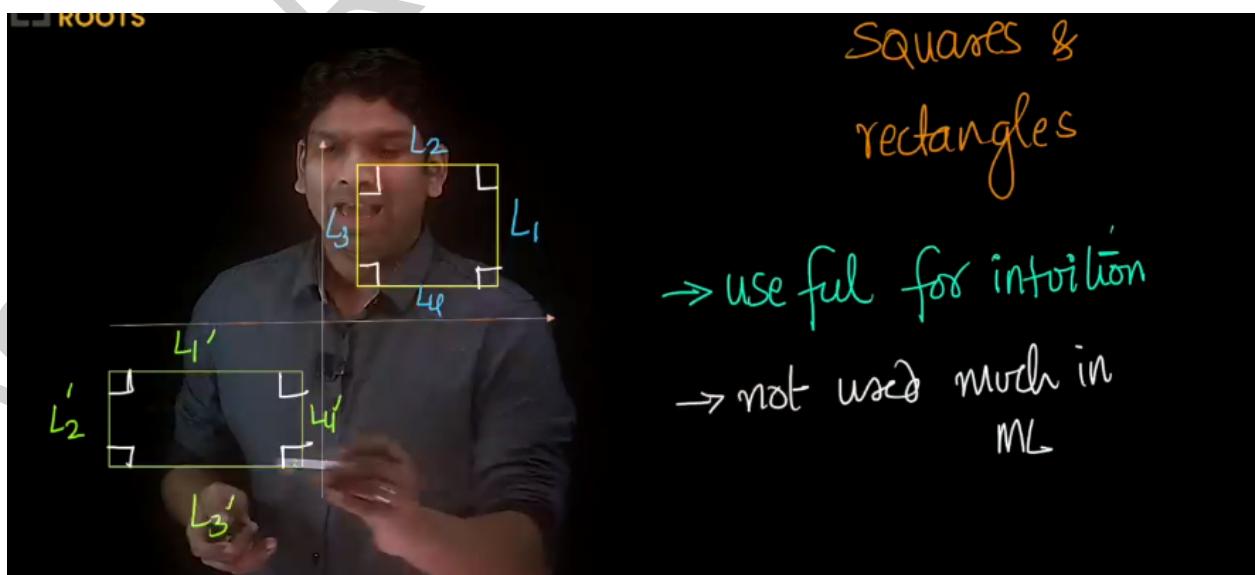
Applied



- We represent an ellipse which is centred at origin as shown above ,we call the x axis as major axis and y axis as minor axis here
- We can extend the same concept to 3D and higher dimensions as shown above.

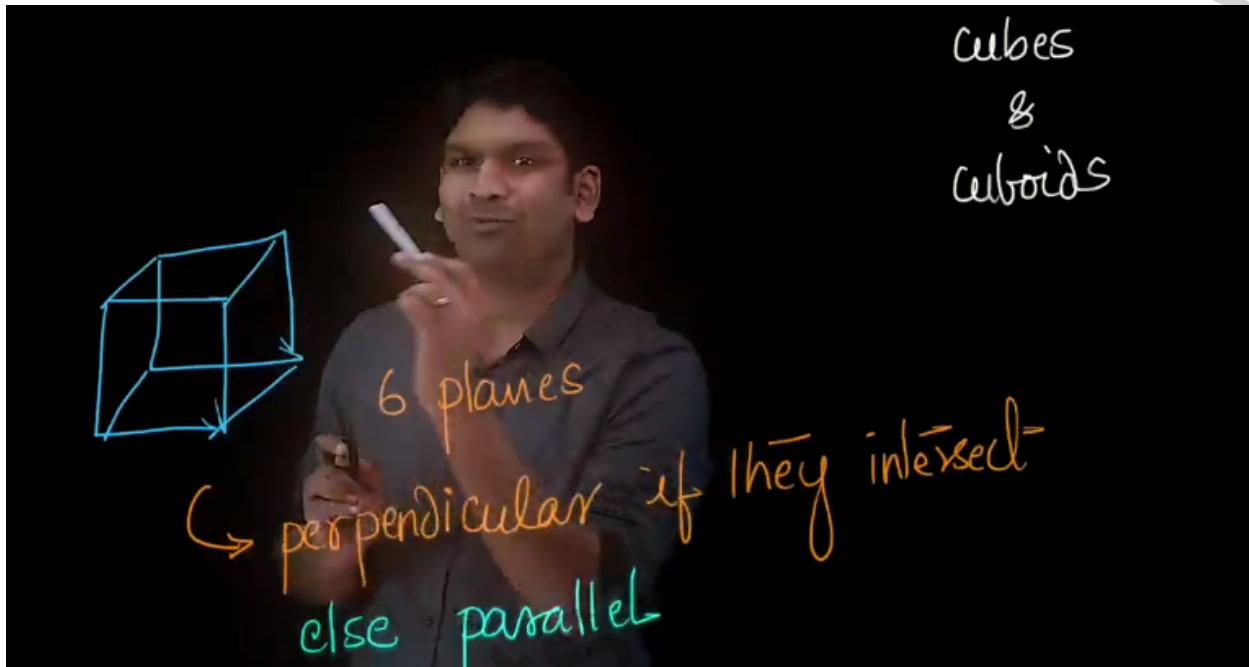
2.18 Squares, Rectangles, Hypercubes and Hyper-cuboids

At timestamp 1.42 in video



- We can represent a square using 4 lines in 2D such that every pair of lines that intersect will be perpendicular and if they don't intersect they have to be parallel to each other .
- Rectangle is similar to a square except that opposite sides will be the same length and all sides need not be of the same length.

At timestamp 2.19 in video



- We can extend the same concept to higher dimensions as well.
- A point lies inside or outside of the rectangle/square if and only if it's x-coordinate lies between the x-coordinate of the given bottom-right and top-left coordinates of the rectangle and y-coordinate lies between the y-coordinate of the given bottom-right and top-left coordinates.