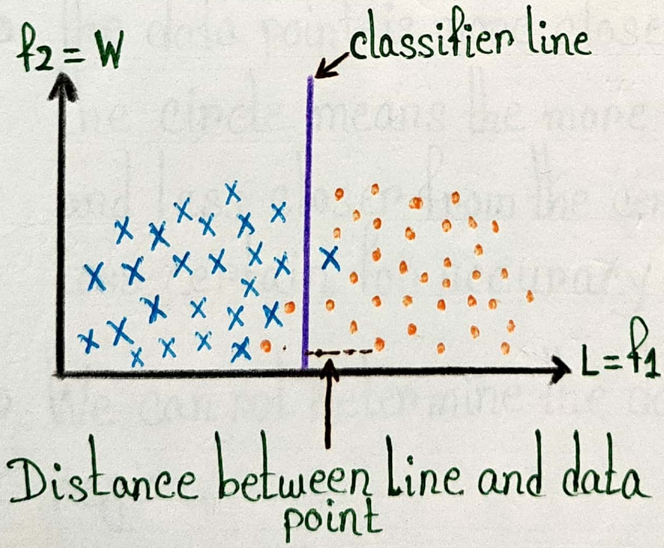
Mathematical Concepts For Data Science

In geometry, a line is a one-dimensional figure that has length but no width. It is made up of a set of points that extend infinitely in opposite directions. A line is determind by two points in a two-dimensional plane.

• Line for Classification



Legend of Plot

- L = Length
- W = Width
- f_1 = Feature 1
- f_2 = Feature 2
- X = Type 1
- • = Type 2

• How line works

① Line decides the side of each data point.

② The more distance between line and data points means the more accuracy and less distance determine less accuracy.

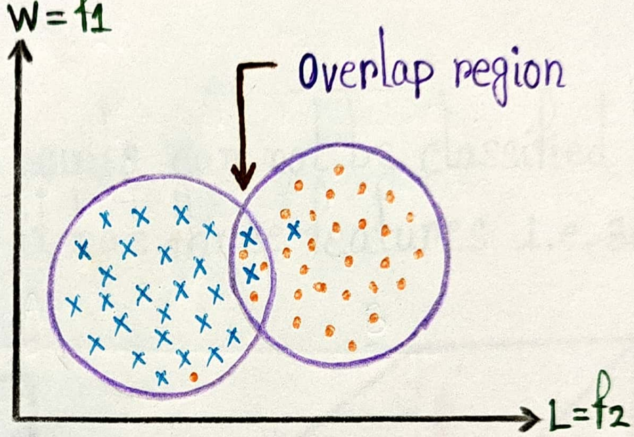
What is a model (M) ?

Here it is the line. The line classify the data points.

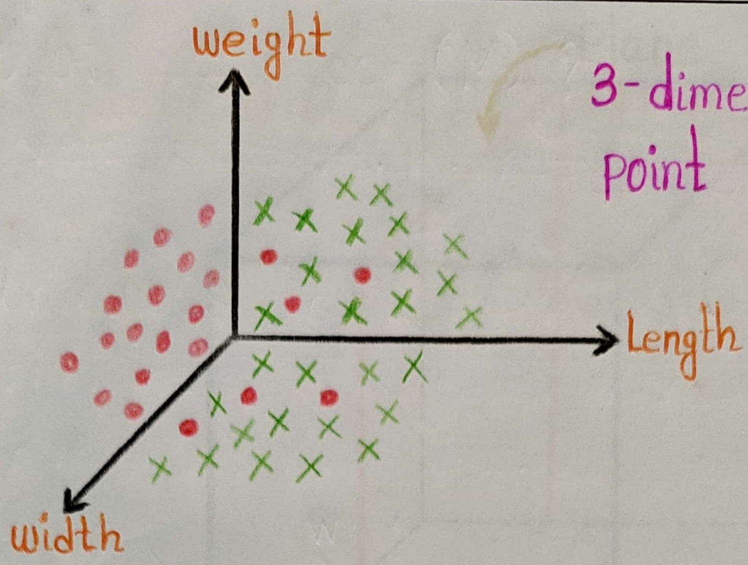
What does "building" a model mean ?

Find the separating line, which can determine the side of the data points.

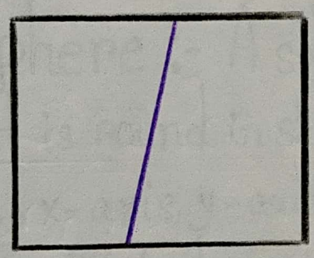
In this graph the model is circle.



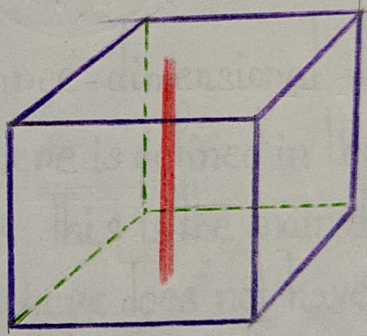
- The data point is more closer from the center of the circle means the more certain for accuracy and less closer from the center of the circle means less certain for accuracy.
- We can not determine the data points in the overlap region.



The above data points can not be classified by a line. Because it has three features i.e. 3d.



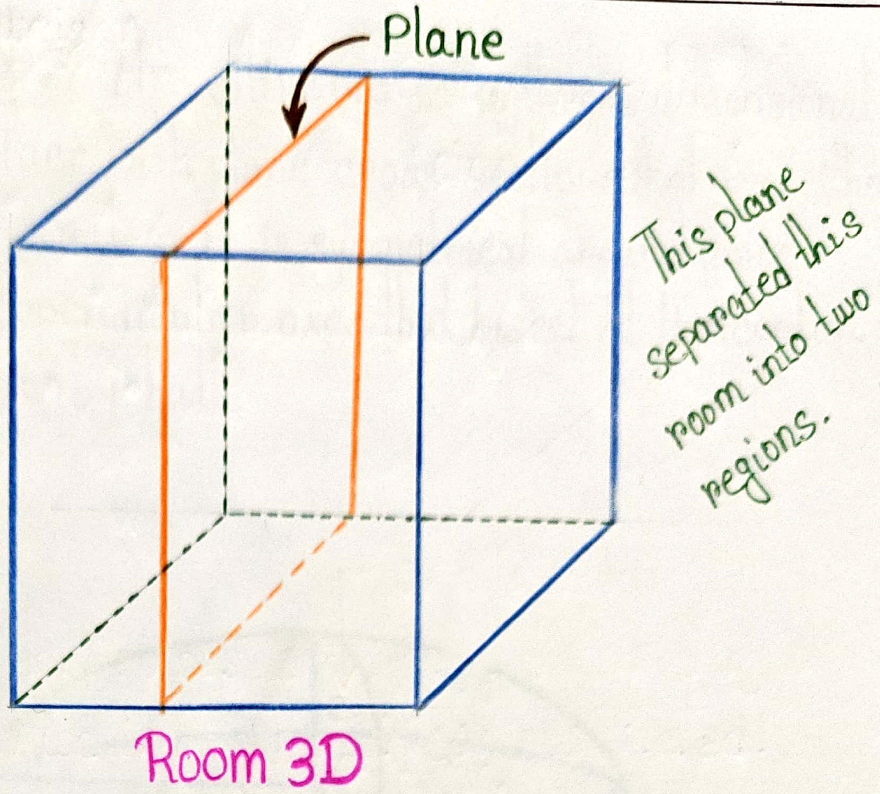
Wall 2D



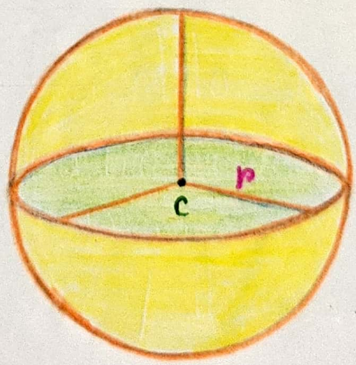
Room 3D

A line is sufficient to split into two region for 2D space. But a line is not sufficient to separate the 3D space. 3D space requires a diffrent model like plane / sphere / ellipsoid to separate into two region.

Plane :- A plane is a flat surface that extends into infinity. It is also known as a two-dimensional surface.



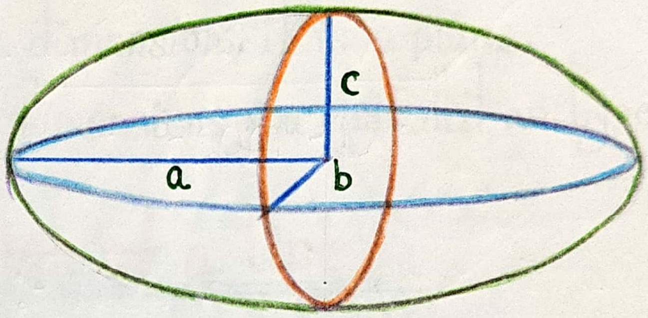
Sphere :- A sphere is a three-dimensional object that is round in shape. The sphere is defined in three axes, i.e., x-axis, y-axis and z-axis. This is the main difference between circle and sphere. A sphere does not have any edges or vertices, like other 3D shapes. Ex- Football.



All three axes of the above sphere are on same distance from the center point.

Ellipsoid :- An ellipsoid is a three-dimensional closed surface with plane cross-sections that are either ellipses or circles. It is symmetrical around three mutually perpendicular axes that bisect at the center.

Example - Rugby ball.



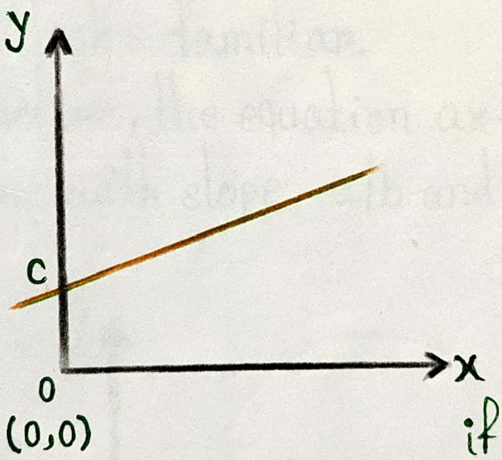
Three axes of the above ellipsoid have not same distance from the center point of ellipsoid.

Hyper-Plane

Hyperplane is used for higher dimension (>3-dimension). A hyperplane is a generalization of a plane.

- In one dimensions, a hyperplane is called a point.
- In two dimensions, it is a line.
- In three dimensions, it is a plane.
- In more dimensions you can call it an hyperplane.

Equation of a line



$y = mx + c$

m → slope

c → y-intercept

if $x = 0$; $y = c$

y = How far up.
x = How far along.
m = Slope or gradient (how steep the line).
c = The y-intercept, where the line crosses y axis.

(m = Change in y / change in x)

Equation of the form $ax+by+c=0$, will represent a straight line. Here a, b, c are arbitrary constants (a and b can not be both 0), and x and y are variables (which represent the coordinates of points on the line).

$$ax+by+c=0$$

or

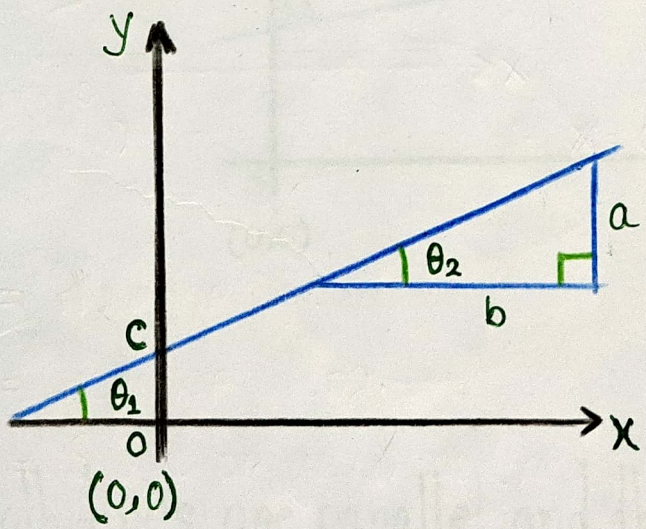
$$y=(-a/b)x+(-c/b)$$

By putting $-a/b=m$ and $-c/b=c$, the above equation becomes.

$$y=mx+c$$

This looks familiar.

Therefore, the equation $ax+by+c=0$ represents a line with slope $-a/b$ and y-intercept $-c/b$.



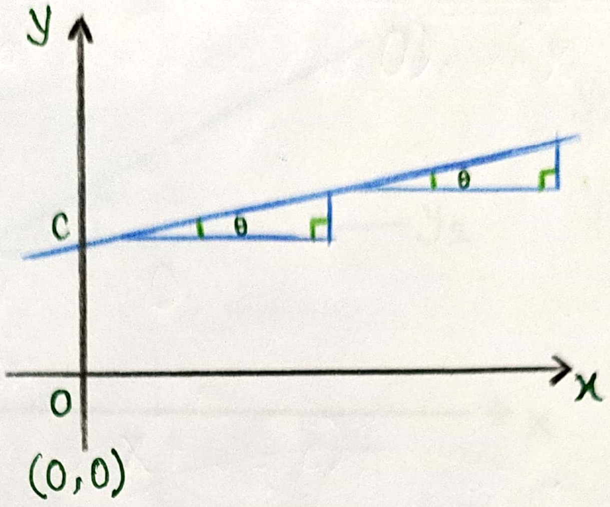
$$y=mx+c$$

$$m=\tan\theta=\frac{a}{b}$$

θ = Theta (slope)

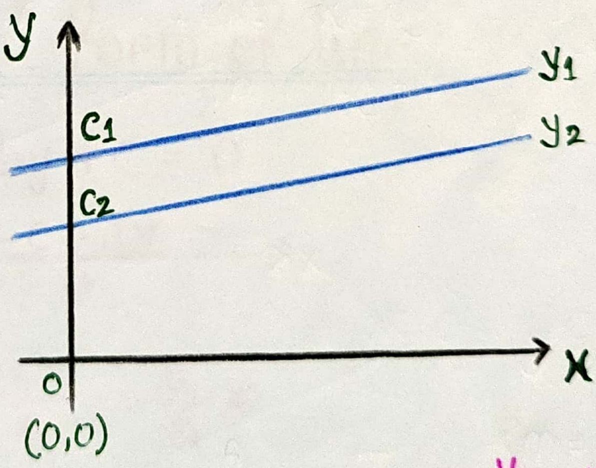
Both the θ_1 and θ_2 are same slope and same angel.

- ## Same slope everywhere



Same slope is everywhere in the above line.

- ## Unique c

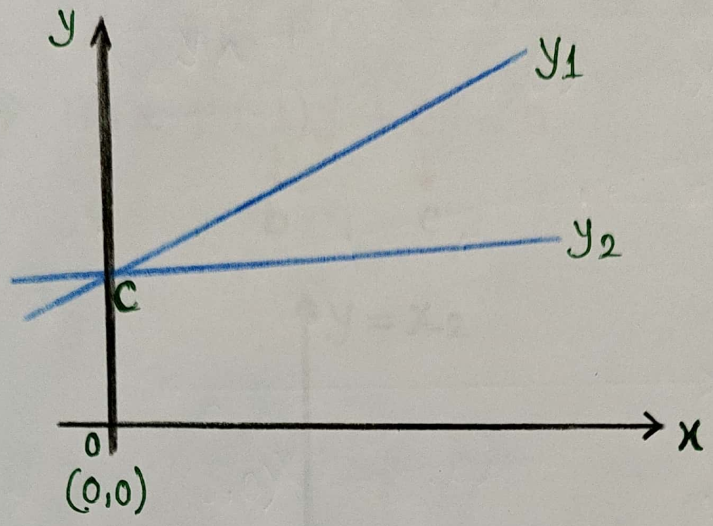


$$y_1 = mx + c_1$$

$$y_2 = mx + c_2$$

Both lines are parallel and the slopes are same. Their intercepts are different. Here m is same but c is different for both line(y_1, y_2).

Unique m



$y_1 = m_1 x + c$

$y_2 = m_2 x + c$

The two line have same intercept c. But they have different slope (m).
y_1 line has larger slope than y_2 line.

General form of line

$$ax + by + c = 0$$

$$\Rightarrow y = \frac{-c - ax}{b}$$

$$\Rightarrow y = \frac{-c}{b} - \frac{a}{b}x$$

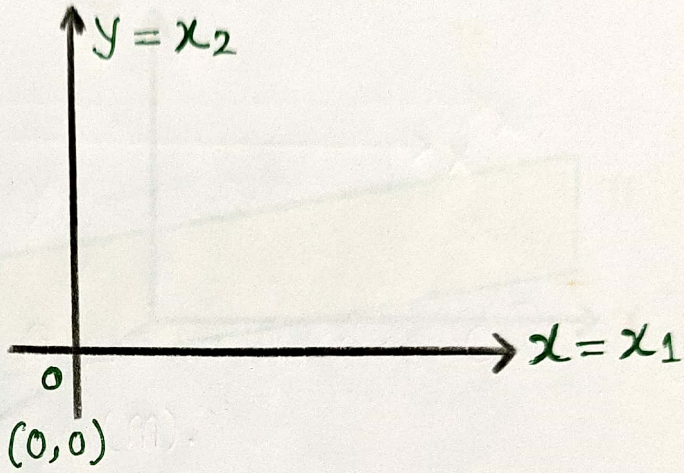
Now we can say $y = \underbrace{\frac{-c}{b}}_{\text{y-intercept}} + \underbrace{\left(-\frac{a}{b}\right)}_{\text{Slope}}x$

General form to slope intercept is shown in above example.

Slope intercept to General Form

$$y = mx + c$$

$$\Rightarrow \underset{a}{m}.x + \underset{b}{(-1)}y + \underset{c}{c} = 0$$



$$ax + by + c = 0$$

$$\Downarrow$$

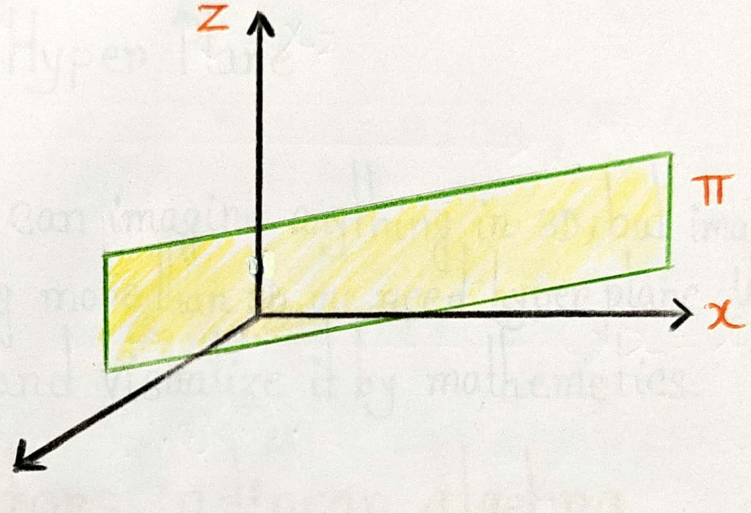
$$w_1x_1 + w_2x_2 + w_0 = 0$$

Variable x and y are replaced by x_1 and x_2. a, b and c are replaced with w_1, w_2 and w_0. This equation is very usefull for multidimension and machine learning model.

Equation of plane & hyper plane

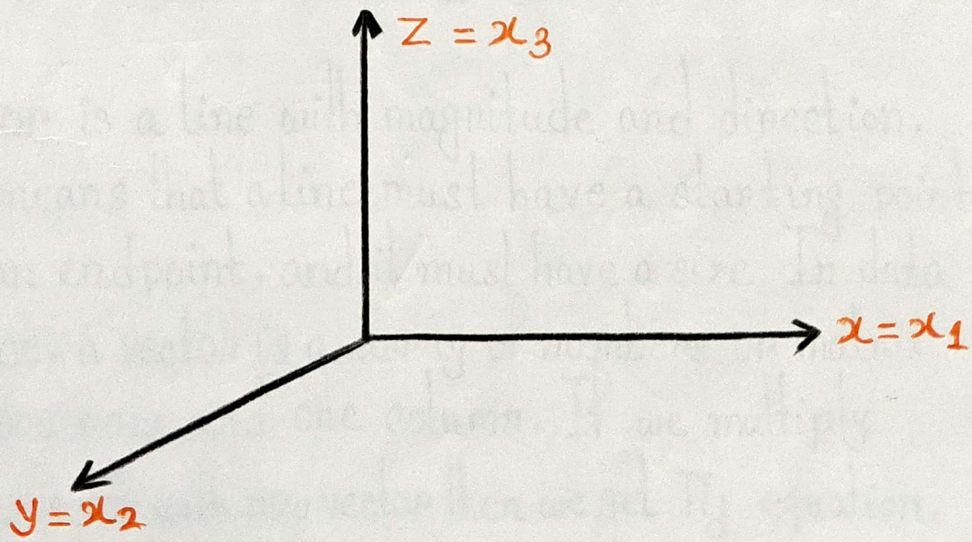
π : $ax + by + cz + d = 0$ → 3D equation

π → plane; x, y, z → 3 axes



π : $w_1x_1 + w_2x_2 + w_3x_3 + w_0 = 0$ [w = a variable]

In the above equation, variable a, b, c and d are replaced by variable w_1, w_2, w_3 and w_0. Variable like x, y and z are replaced by x_1, x_2 and x_3.



$L : w_1x_1 + w_2x_2 + w_0 = 0 \rightarrow$ 2D equation [L = Line]

$\Pi : w_1x_1 + w_2x_2 + w_3x_3 + w_0 = 0 \rightarrow$ 3D equation

$HP : w_1x_1 + w_2x_2 + \ldots\ldots + w_dx_d + w_0 = 0 \rightarrow$ more than 3D

$\downarrow$

Πd d-D = d-dimensions

HP = Hyper Plane

Human can imagine anything in 3D, but imagine anything more than 3D we need hyper plane. We can study and visualize it by mathemetics.

• Vectors in linear algebra

$\Pi : w_1x_1 + w_2x_2 + \ldots\ldots + w_dx_d + w_0 = 0$

$$\begin{bmatrix} w_1 & w_2 & w_3 & \ldots\ldots & w_d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} + w_0 = 0$$

Column vector

Row vector

Vector is a line with magnitude and direction. This means that a line must have a starting point and an endpoint, and it must have a size. In data science, a vector is a array of numbers or matrix with one row and one column. If we multiply Column vector with row vector then we get Π_d equation.

$\begin{bmatrix} 2.8 \\ 9 \\ 5 \end{bmatrix} \rightarrow$ Vector is an array of numbers.

$5.6 \rightarrow$ Scalar is a single number.

$$W = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_d \end{bmatrix}$$

$\rightarrow$ d- dimensional vector.

$\rightarrow W \in R^d$ (mathematical notation).

$\in$ = belongs to / an element of

R= Real number , It is a number that can be used to measure a continuous one-dimensional quantity such as a distance, duration or temperature.

R^d = Sets of real number.

w belongs to R^d when w is vector of d-dimension where each dimension is a real number. R is a real number of d-dimension.

- By default a vector is a column vector.

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{bmatrix} \rightarrow W^T = \begin{bmatrix} w_1 & w_2 & \cdots & w_d \end{bmatrix}$$

T= Transpose

Above column vector transposed into row vector.

Transpose of a matrix

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

Here each row becomes one columns.

$$\Pi : \underbrace{[w_1\ w_2\ w_3 \ldots\ldots w_d]}_{w^T} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}}_{x} + w_0 = 0$$

Now we can write below formula:

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

Note: w and x should be same dimensinal vectors.

We can not multiply below vectors.

$$\underbrace{[2\ \ 3\ \ 4]}_{\text{3 components}} \qquad \begin{bmatrix} 1 \\ 2 \\ 5 \\ 6 \end{bmatrix}$$

→ Not valid
→ 4 components

The above column vector and row vector can not be multiplied becouse thier dimension is different.

$$\pi : w^T x + w_0 = 0$$

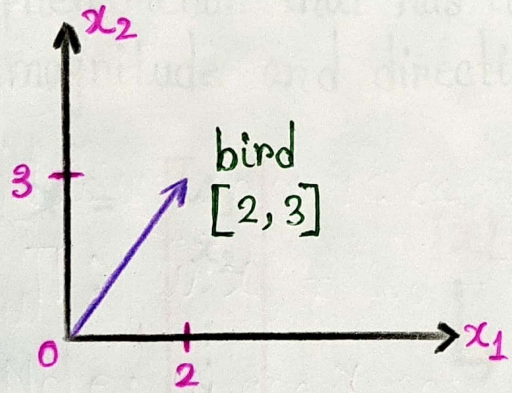
w → d-dim, x → d-dim

If
$d = 2 \Rightarrow$ Line
$d = 3 \Rightarrow$ Plane
$d \geq 4 \Rightarrow$ Hyperplane

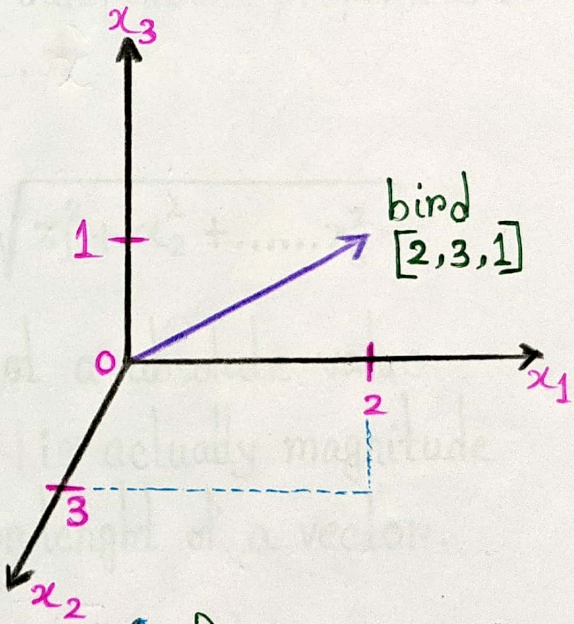
(dim = dimension)

If d change then it would be line, plane and hyperplane. We can use the above formula for every dimension.

- **d-dim points as vector**

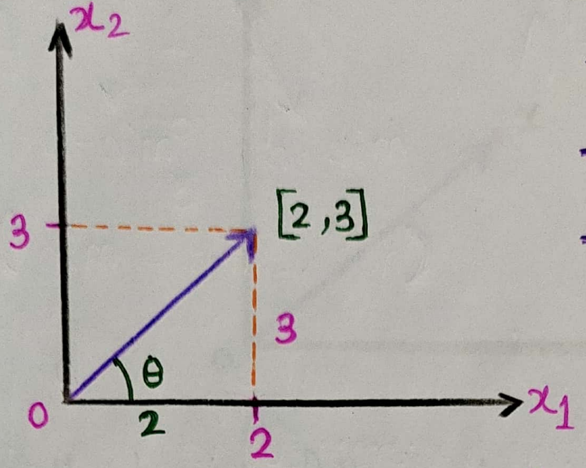


(2 features/2D)



(3 features / 3D)

Vectors in physics



→ O - Origin
→ direction (θ)
→ magnitude = lenght
$= \sqrt{2^2 + 3^2}$

θ = Theta

(asper pythagoras theorem)

As per physics, a vector is a quantity or phenomenon that has two independent properties : magnitude and direction. ★

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

↓

d-dimension Vector

$$|x| = \sqrt{x_1^2 + x_2^2 + \ldots\ldots x_d^2}$$

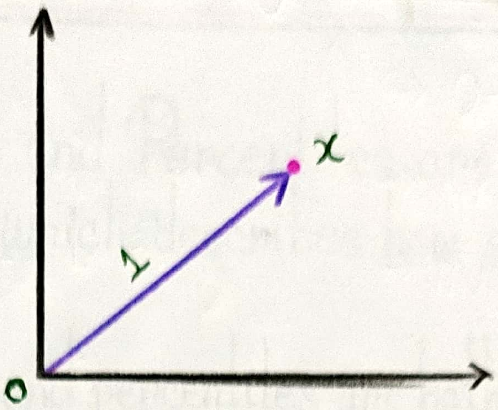
↳ Not a absolute value
$|x|$ is actually magnitude or lenght of a vector.

To avoid confusin, we can use this notation:

$|x| \rightarrow \|x\|$

★ Vector start from O (origin). We can get direction using the angle (θ) between vector and x-axis.

Unit vector



$\|x\| = 1$

If the lenght of the vector is 1 then it is called Unit vector.

Dataset of Flower

f_1	f_2	f_3	f_4	f_5	Type
←	←	x_1	→	→	y_1
←	←	x_2	→	→	y_2
←	←	⋮	→	→	⋮
←	←	⋮	→	→	⋮
←	←	x_n	→	→	y_n

$\in$ = Belongsto
$\forall$ = For all
R^d = Sets of real number

$x_i \in R^5 \quad \forall i : 1 \rightarrow n$ → x_i belongsto R^5 for all ($\forall$) i in range of 1 to n.

$y_i \in \{c_1, c_2\}$ ← sets → Each y_i belongs to c_1 or c_2 in set (c_1, c_2 refer the type of x_i).

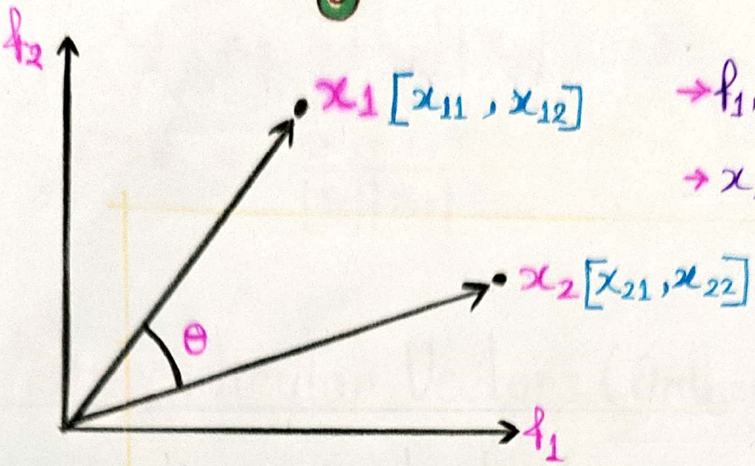
- Above data set can be express in below equation.

$D = \{(x_i y_i);\ x_i \in R^d,\ y_i \in \{c_1, c_2\}\}$

↓

Dataset

Angle Between Vectors



→ f_1, f_2 are axes.

→ x_1, x_2 are vectors.

x_1 is the data point and it has two coordinate. x_{11} is a coordinate of f_1 axis and x_{12} is a coordinate of f_2 axis. x_2 is similar like x_1.

- **Dot product** : This is the operation between two vectors.

$x_1 \cdot x_2 = |x_1| \, |x_2| \cos\theta$ ← geometric way.

θ = Theta is angle between two vectors.

$x_1 . x_2 = x_1^T x_2$ ← linear algebric way.

$$\Rightarrow x_1^T x_2 = |x_1| |x_2| \cos\theta$$

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

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

$$\Rightarrow \cos^{-1}\left(\frac{x_{11} x_{21} + x_{12} x_{22}}{\sqrt{x_{11}^2 + x_{12}^2} \sqrt{x_{21}^2 + x_{22}^2}} \right) = \theta$$

In concise way we can write

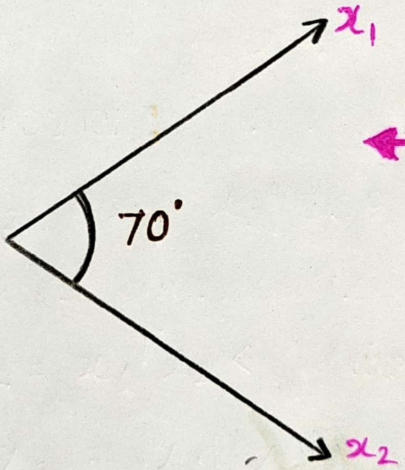
$$x_1 \cdot x_2 = |x_1||x_2| \cos\theta$$

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

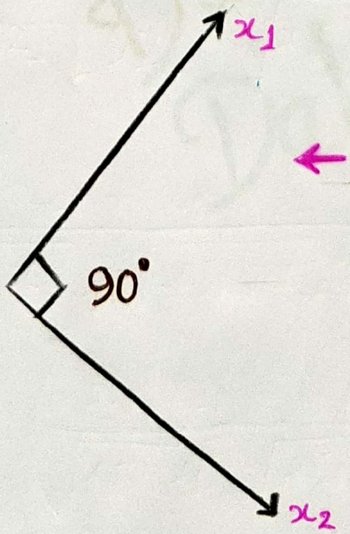
$x_1 \in R^d$
$x_2 \in R^d$

- ## Perpendicular Vectors (Orthogonal):

In geometry, perpendicular lines are defined as two lines that meet or intersect each other at the right angles (90°).



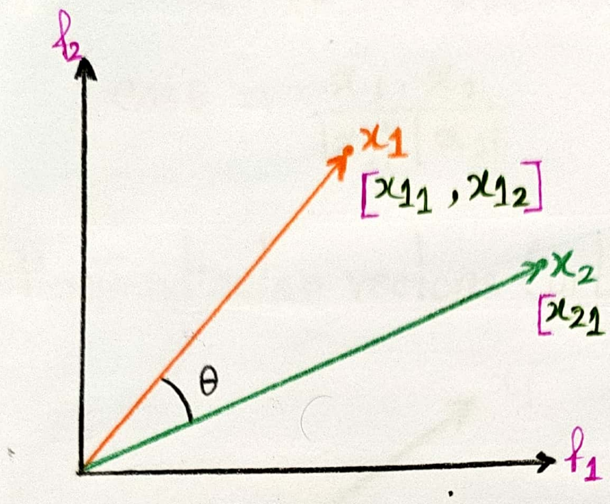
← This is not a perpendicular.



← This is a perfect perpendicular.

$$x_1 \cdot x_2 = |x_1||x_2| \cos 90^\circ$$
$$= 0$$

Angel between vectors



⇒ Use f_1, f_2 for axis

⇒ x_1 & x_2 for vector

x_1 is datapoint. x_1 has two coordinate. x_{11} is a coordinate of f_1 axis and x_{12} is a coordinate of f_2 axis. x_2 is similar like x_1.

Dot product

This is a operation between two vectors.

$$x_1 \cdot x_2 = |x_1|\,|x_2| \cos\theta \leftarrow \text{geometric way}$$

θ = Theta is angel between two vectors.

$$x_1 \cdot x_2 = x_1^T x_2$$

$$\Rightarrow x_1^T x_2 = |x_1|\,|x_2| \cos\theta$$

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

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

$$\Rightarrow \cos^{-1}\left(\frac{x_{11} x_{21} + x_{12} x_{22}}{\sqrt{x_{11}^2 + x_{12}^2} \; \sqrt{x_{21}^2 + x_{22}^2}}\right) = \theta$$

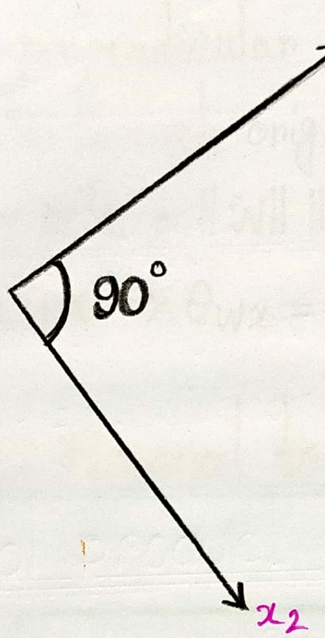
Scanned with OKEN Scanner

In concise way we can write

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

$$\cos\theta = \frac{x_1 \cdot x_2}{|x_1||x_2|} \qquad x_1 \in R^d, \quad x_2 \in R^d$$

Perpendicular vectors (orthogonal)



$$x_1 \cdot x_2 = |x_1||x_2| \cos 90^\circ = 0$$

★ Dot product is equal to zero.

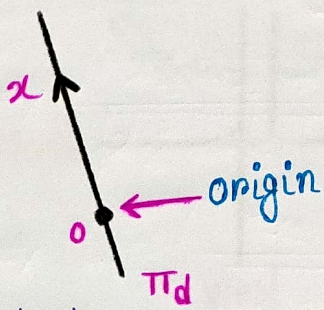
• Equation of a plane

$$\Pi_d = W^T x + W_0 = 0$$

Π_d ↙ d-dim; $W^T x$ ↓ vectors

Plane through origin

Let $x = 0 = [0\,0\,0 \ldots\ldots 0]^T$



d-dim space

$$[W_1\, W_2 \ldots W_d] \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + W_0 = 0$$

(Product of this two vectors is 0)

$\Rightarrow W_0 = 0$

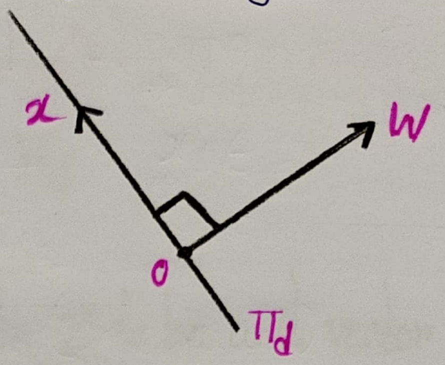
If plane is passing through the origin then W_0 will be 0.

$W^T x = 0$

$\Rightarrow W.x = 0$

$\downarrow$

$W \perp x$



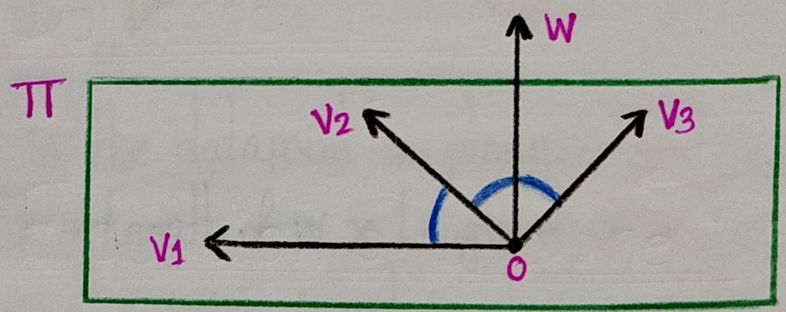
W is perpendicular on every vector on the plane and also through origin.

$W.x = W^T x = \|W\| \|x\| \cos\theta = 0$

$W_1 x$

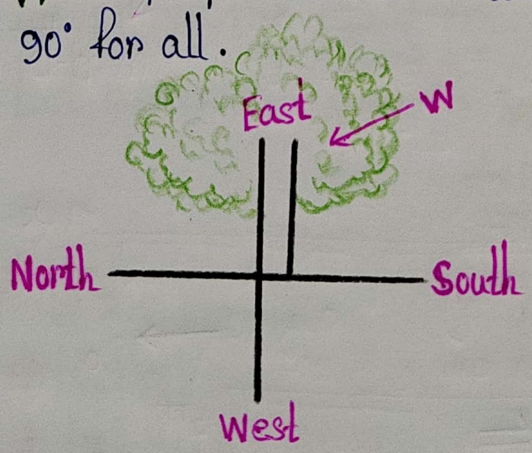
$W \perp x \Rightarrow \theta_{Wx} = 90^\circ$

- **W : Normal to π**



3-D plane

W is perpendicular to all three vectors and angel is same 90° for all.



Imagine this tree is W and the W will be always perpendicular to all four direction (four vectors) with 90° angel.

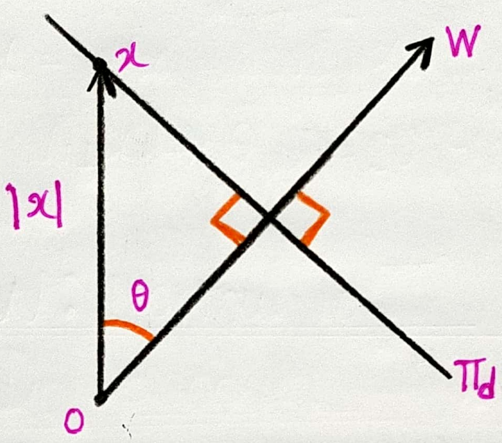
Length of W

$|W| = 1$ (unit vector)

W : unit normal to π_d
(convention)

It is not necessary to W to be unit vector always.

Plane not passing through origin



$$W^T x + W_0 = 0$$
$$\downarrow$$
$$\neq 0$$

x is the datapoint on plane.

$W^T x$ = Length of W x Length of x x cos of x.

Here the origin is not in plane so $W^T x$ will not be equal to 0. And W_0 will not also be equal to compensate or prove the equation.

Let assume, $\underbrace{W^T x + W_0}_{L.H.S} = \underbrace{0}_{R.H.S}$

$(2) + (-2) = 0$

$= 0$

* $W^T x + W_0 = 0$

$\Rightarrow |W||x| \cos\theta + W_0 = 0$ If W : unit normal (perpendicular)

$\Rightarrow |x| \cos\theta + W_0 = 0$