SVM

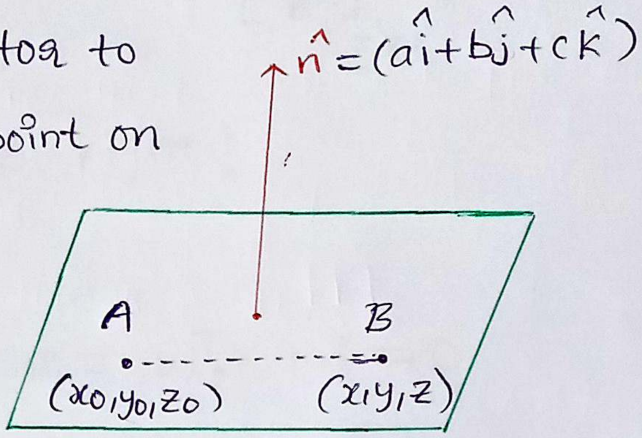
- Support Vector Machines!

* Before going to actual concept let's understand basic Mathematics!

Equation of plane:-

* Given a normal vector to the plane, and a point on it (A)

* B is any random point



How we will find the Eq^n?

It's very simple right! Define a vector $\overline{AB}$,

Since $\overline{AB}$ lies in the plane then $\overline{AB}$ is $\perp r$ to $\hat{n}$ right!

$\therefore$ Eq^n of plane $= \hat{n} \cdot \overline{AB}$ (dot-product) $= 0$

$$\Rightarrow (a\hat{i}+b\hat{j}+c\hat{k}) \cdot ((x-x_0)\hat{i}+(y-y_0)\hat{j}+(z-z_0)\hat{k}) = 0$$

$$\Rightarrow ax+by+cz-\underbrace{(ax_0+by_0+cz_0)}_{}=0$$

$$\Rightarrow ax+by+cz+d=0$$

$$\boxed{ax+by+cz+d=0}$$

Can we able to write the Eq^n of plane in Matrix form?

yes, we can!

Let $w = \begin{bmatrix} a \\ b \\ c \end{bmatrix}_{3\times1}$ and $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{3\times1}$

$w^T X = ax + by + cz$ right!

Now,

Equation of plane = $w^T . X + d = 0$

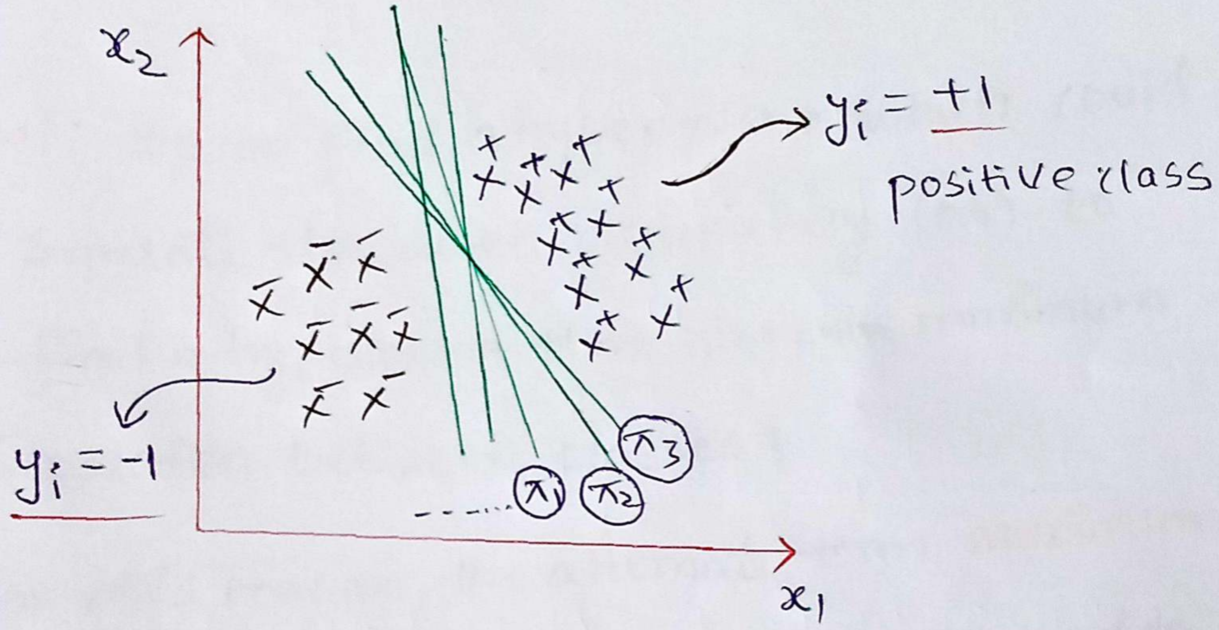
V. Imp:- Since $\hat{n} = a\hat{i} + b\hat{j} + c\hat{k}$ is normal vector to plane $\Rightarrow$ w is normal vector to plane.

* Since the dimension of W and X is '3' we got a plane! If the dimension is more than '3' we call it has

HYPER PLANE

Hope you understood the definination of Hyperplane!

Now Let's Dive into SVM:-



* we are assuming the given data is Linearly Separable. which means we were able to classify data using linear equation

Linear equation:- If the degree of each term of an equation is equal to 1 then we say its linear equation.

Exs:- $ax^1 + by^1 + cz^1 + \dots = 0$ ✓

$ax^2 + bx^3 = 0$ ✗

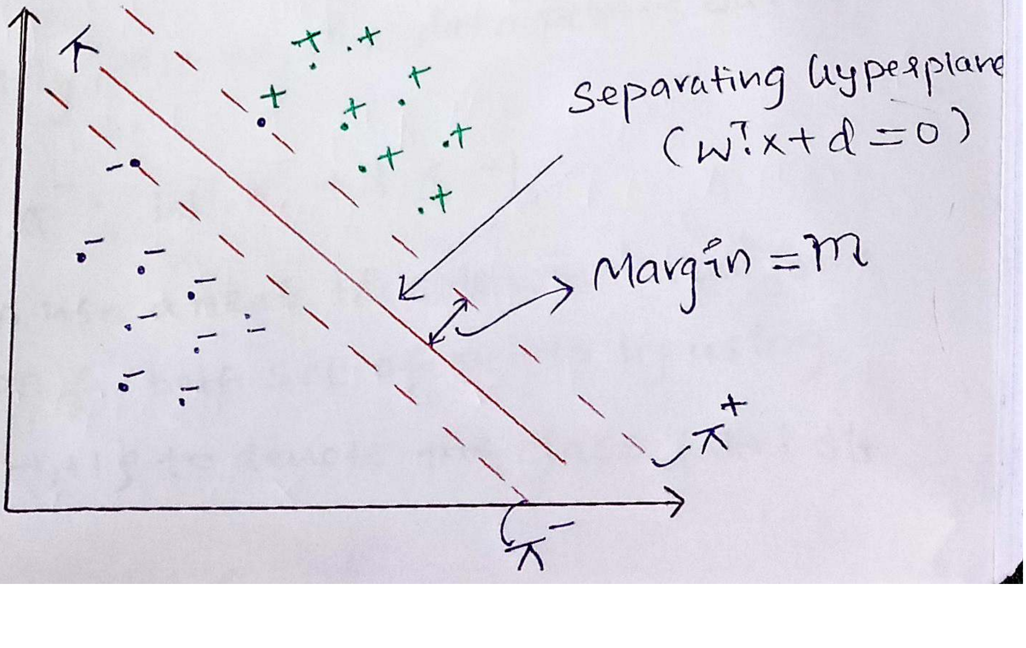
* clearly, we can define 'n' no.of Hyper planes ($\pi_1, \pi_2, \pi_3 \ldots \pi_n$) to seperate the data. But we want More Accurate and Robust one!

How to find it?

-Don't worry SVM helps us!

Goal: Is to find a hyperplane which could seperate the data accurately (or) to find a hyperplane that gives the maximum separation between classes!

* For this reason, the alternate term maximum margin classifer is also sometimes used to referred to as the SVM.



Problem Formulation :-

1. The Geometric Approach

The equation of Hyper plane :

$$\pi : W^T . x + d = 0$$

$$W = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

* As we are looking to maximize margin between postive and negative data points, we would like this positive data points to satisfy the following constraint

$$\pi^+ : W^T . x_i^+ + d \geq +1$$

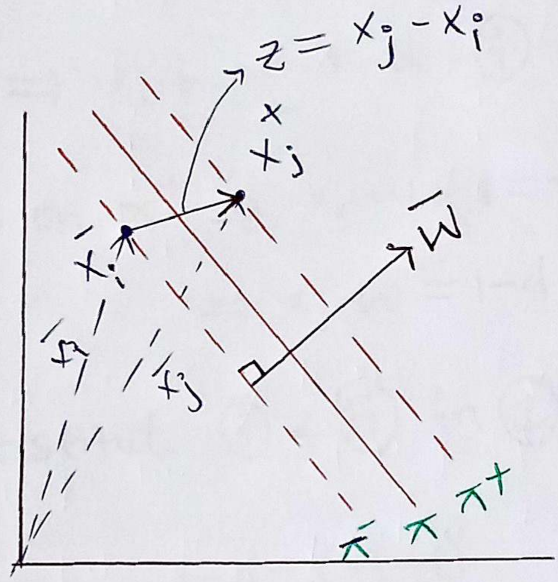
Simiarly, the negative data points satisfy:

$$\pi^- : W^T . x_i^- + d \leq -1$$

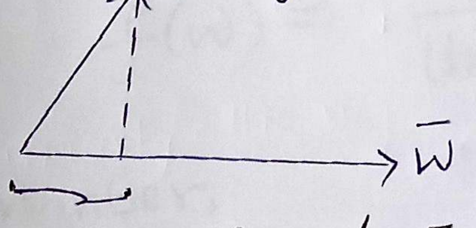
We can use a neat trick to write a uniform equation for both set of points by using $y_i \in \{-1, +1\}$ to denote the class label of data point x_i:

$$\boxed{y_i(w^T \cdot x_i + d) \geq +1}$$

* We aim to maximize the distance between π^+ and π^- which mean maximizing the margin.



* since we can move the vectors freely:



projection of z on $\bar{w}$

clearly, the projection of z on $\vec{w}$ gives the ~~perpendicular~~ Distance between π^+ and π^-

$$f(w) = \frac{(x_j - x_i) \cdot \bar{w}}{||w||}$$

$$= \frac{x_j \cdot \bar{w} - x_i \cdot \bar{w}}{||w||} \quad -(1)$$

x_i lies on π^- $\Rightarrow$ $x_i \cdot w + d = -1$ ($\because$ point lying on line)

$\Rightarrow x_i \cdot w = -1 - d \quad -(i)$

& x_j lies on π^+ $\Rightarrow$ $x_j \cdot w + d = +1$

$\Rightarrow x_j \cdot w = 1 - d \quad -(ii)$

Now substitute (i) & (ii) in (1) we get

$$f(w) = \frac{1 - d - (-1 - d)}{||w||}$$

$$f(w) = \frac{2}{||w||}$$

Remember,

$$\text{Goal} = \max_{w,d} f(w)$$

subjected to constraint "

$$y_i (w \cdot x_i + d) \geq 1$$