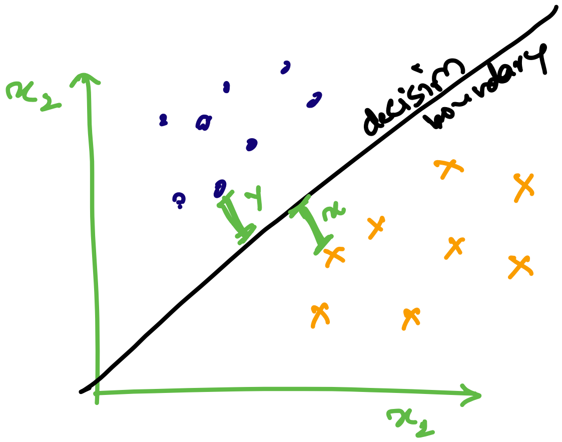
Support Vector Machine (SVM)



Points which are closest to your d.b. should be very far away from each other.

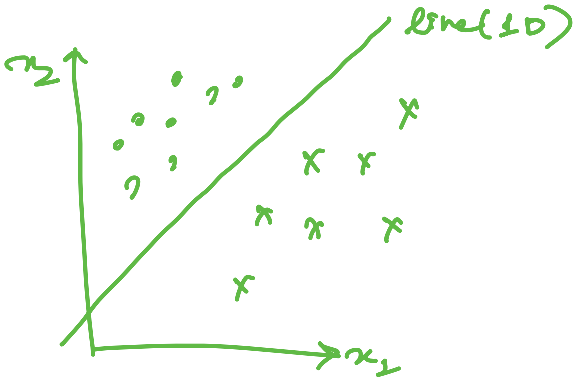
Recap:

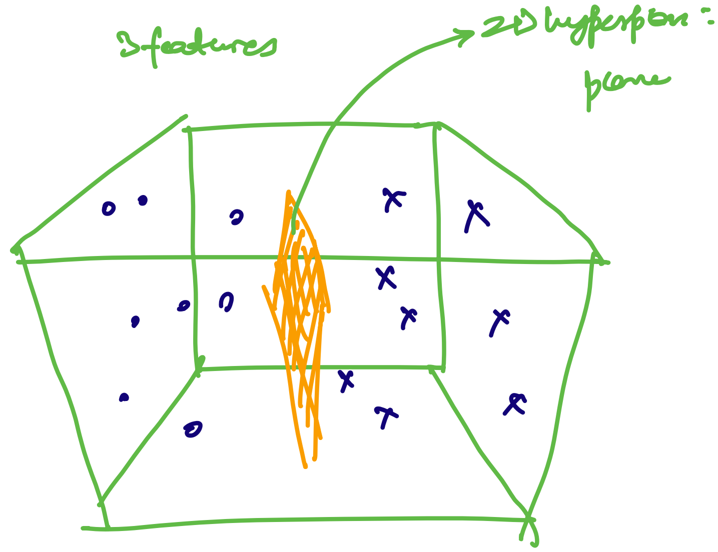
Regression:
Linear Regression

Classification:
Logistic Regression
KNN
Naive Bayes
Decision Trees

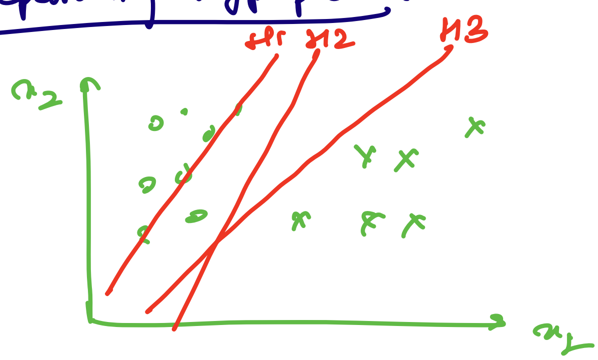
Hyperplane:

n features hyperplane n-1 dimension

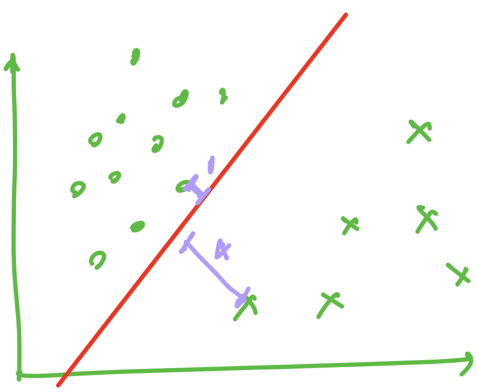


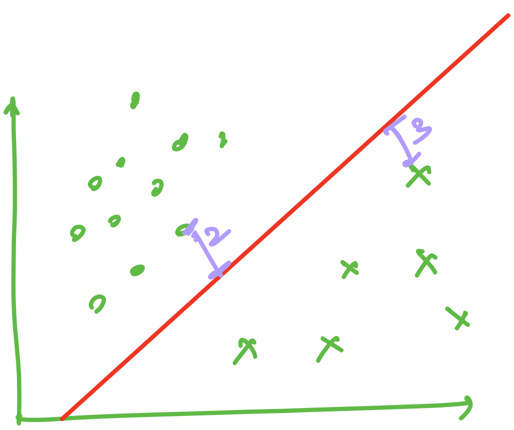


Separating Hyperplane:



H2, H3: Separating Hyperplane
H1 ✗

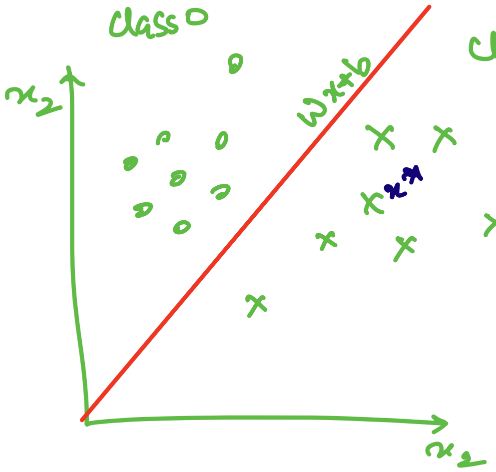




$\min(1, 4) = 1$ $\qquad$ $\min(2, 3) : 2$ ✔

max: 2

Maximise the minimum distance from SH



$y = mx + c$

$y = wx + b$

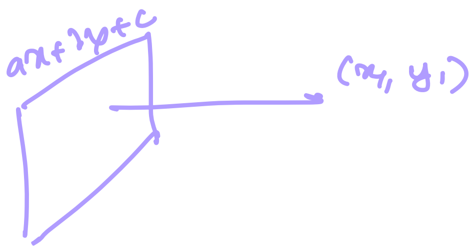
$w_0 + w_1 x_1 + w \ldots w_n x_n$

$b + wx$

$wx^* + b$

$wx^* + b > 0$: Class 1

$wx^* + b < 0$: Class 0



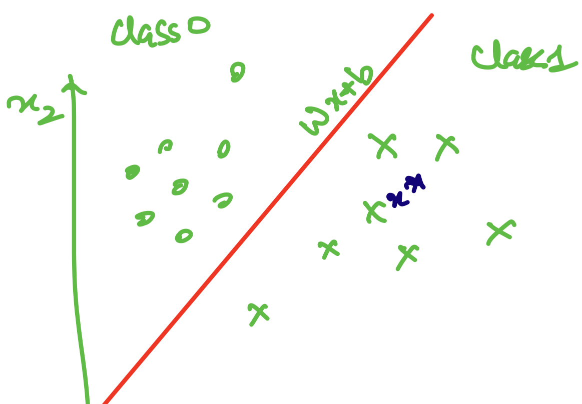
$$\frac{ax_1 + by_1 + c}{\sqrt{a^2 + b^2}}$$

L_2 norm: $(a^2 + b^2)^{1/2}$

L_1 norm: $(a^1 + b^1)^{1/1}$

$\vdots$

L_3 norm: $(a^3 + b^3)^{1/3}$



$wx + b = w_1 x_1 + w_2 x_2 + \underbrace{b}_{\text{bias}} = 0$

$x_1^{(i)}$ $\quad$ $x_2^{(i)}$

$w_1 x_1^{(i)} + w_2 x_2^{(i)} + b$

$\sqrt{w_1^2+w_2^2}$ L_2 norm $\|w\|_2$

formulate objective:

$w_1x_1+w_2x_2+b$

$\begin{bmatrix}w_1\\w_2\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

w^Tx

$\begin{bmatrix}w_1 & w_2\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

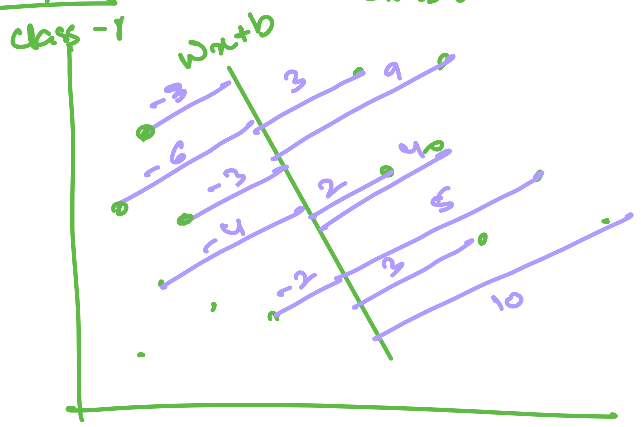
$w_1x_1+w_2x_2$

$X = \{x^1, x^2, x^2 \dots x^m\}$

$Y = \{y^1, y^2, y^3 \dots y^m\}$

Binary classification $y^{(i)} \in \{-1, 1\}$ → class labels

Loss fn:



$$\gamma^{(i)} = \frac{w^Tx^{(i)}+b}{\|w\|_2}$$

→ distance of the point from decision boundary

$$y^{(i)}\gamma^{(i)} = \frac{y^{(i)}(w^Tx^{(i)}+b)}{\|w\|_2}$$

$$\gamma = \min_{i=1\dots m} \gamma^{(i)}$$

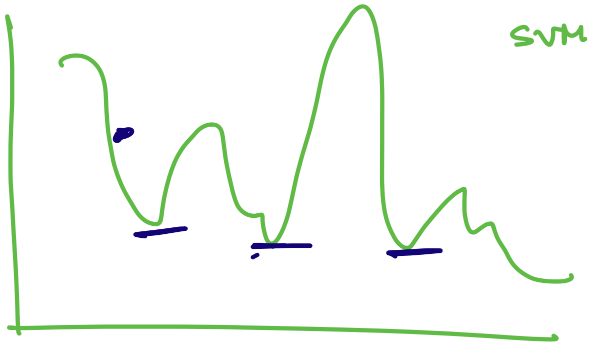
target: max γ

SVM objective

$$\begin{cases} \max \gamma \\ \text{such that } \dfrac{y^{(i)}(w^Tx^{(i)}+b)}{\|w\|_2} \geq \gamma \text{ for all } i = 1 \dots m \end{cases}$$

$\dfrac{w^Tx^{(i)}+b}{\|w\|_2}$ → normalized distance → + , → −

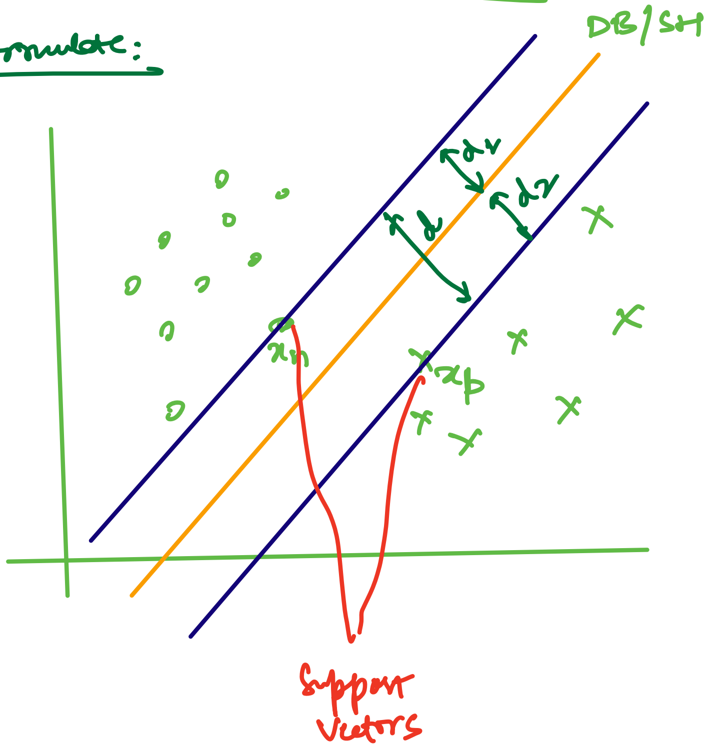
$y^{(i)} \dfrac{w^Tx^{(i)}+b}{\|w\|_2}$ → normalized absolute distance → +ve



SVM objective: Non convex fxn

$$w_1 = w_1 - \eta \frac{\partial L}{\partial w_1}$$

Reformulate:



Re-normalize the data points such that points which are closest are at distance +1 & -1

$$d_1 = \frac{|w^T x_n + b|}{\|w\|_2}$$

$$d_2 = \frac{|w^T x_p + b|}{\|w\|_2}$$

$$w^T x_n + b = -1$$

$$w^T x_p + b = 1$$

$$d_1 = \frac{1}{\|w\|_2}$$

$$d_2 = \frac{1}{\|w\|_2}$$

$$d = d_1 + d_2 = \frac{2}{\|w\|_2}$$

$$\max d \Rightarrow \min \frac{\|w\|_2}{2}$$

SVM S $\min \frac{\|w\|}{2}$

objective: under the condition that all points should have min distance 1.

$$\frac{y^{(i)}(w^T x^{(i)}+b)}{\|w\|_2} \geq \frac{1}{\|w\|_2}$$

SVM objective:

$$\begin{cases} \min \dfrac{\|w\|}{2} \\ \text{such that} \quad y^{(i)}(w^T x^{(i)}+b) \geq 1 \end{cases}$$

$\|w\|_2 = \sqrt{w_1^2+w_2^2}$

$\|w\|_2^2 = w_1^2+w_2^2$

$\|w\|_2^2 = w^T \cdot w$

$$w^T \cdot w = [w_1\ w_2 \dots w_n] \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

$= \underline{w_1^2+w_2^2} + w_3^2 \dots w_n^2$

SVM objective:

$$\begin{cases} \min \dfrac{\|w\|_2^2}{2} \\ \text{such that} \quad y^{(i)}(w^T x^{(i)}+b) \geq 1 \\ \qquad \forall i \in \{1 \dots m\} \end{cases}$$

PYQ: