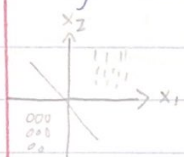
02/07

Support Vector Machine (SVM)

Assume linear separability

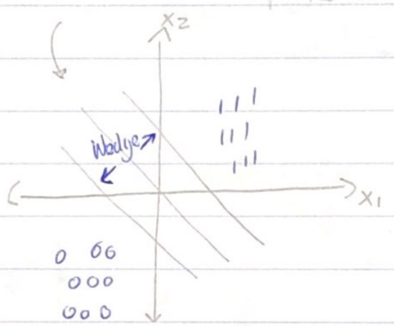
$Y = \{0, 1\}$

Binary Clasification model

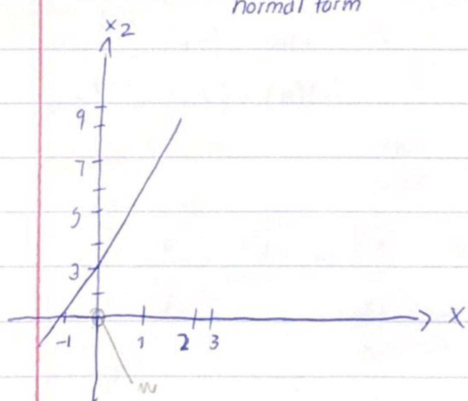


° Wedge space in the lines.

We want a line that separates bouth spaces evenly



$$\mathcal{H} = \{ \mathbb{1}\, \underbrace{\vec{w} \cdot \vec{x} - b}_{\text{Hesse normal form}} \geq 0 \;:\; \vec{w} \in \mathbb{R}^p,\, b \in \mathbb{R} \}$$



$$l:\ x_2 = 2x_1 + 3$$

$$\Rightarrow\ 2x_1 - x_2 + 3 = 0$$

$$= \underbrace{\begin{bmatrix} 2 \\ -1 \end{bmatrix}}_{\vec{w}} \cdot \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\vec{x}} - \underbrace{3}_{\vec{b}} = 0$$

Where is $\vec{w}$ on the plot? "Normal vector" it is $\perp$ (perpendicular) to the hiperplane

$$\|\vec{w}\| := \sqrt{\sum_{i=1}^{p} w_i^2} = \sqrt{\vec{w} \cdot \vec{w}}$$

norm

$$\vec{w}_0 := \frac{\vec{w}}{\|\vec{w}\|}$$ the normal vector w/ lenght.

let $\vec{z} = \alpha \vec{w_0}$

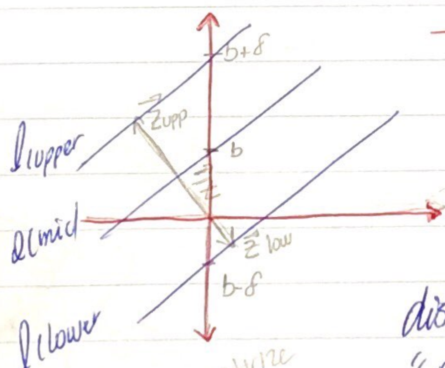
$$\|\vec{z}\| = \sqrt{\sum_{i=1}^{p} (\alpha w)^2} = \sqrt{\alpha^2 \sum_{i=1}^{p} w_{0,i}^2} = |\alpha| \|w_0\| = |\alpha|$$

Find $\vec{z}$ which starts from origin to l:

$$\vec{w} \cdot \vec{z} - b = 0$$

$$\vec{w}(\alpha \vec{w_0}) - b = 0 \Rightarrow \alpha \frac{\vec{w}\vec{w}}{\|\vec{w}\|} = b \Rightarrow \alpha \frac{\|\vec{w}\|^2}{\|\vec{w}\|} b \Rightarrow$$

$$\alpha = \frac{b}{\|\vec{w}\|}$$



$\rightarrow$ $\alpha_u : \vec{w} \cdot \alpha \vec{w_0} - (b+\delta) = 0$

$$\Rightarrow \alpha_u = \frac{b+\delta}{\|w\|}$$

$$\alpha_l = \alpha_l = \frac{b-\delta}{\|w\|}$$

distance between upper & lower line, the "Margin"

$$\text{Margin} = \alpha_u - \alpha_l = \frac{b+\delta}{\|w\|} - \frac{b-\delta}{\|w\|} = \frac{2\delta}{\|w\|}$$

$\rightarrow \frac{2}{\|w\|}$

everything is overparametrize

$$\vec{w} \cdot \vec{x} - b = 0$$

$$c\,\vec{w} \cdot \vec{x} - cb = 0$$

set $\delta = 1$

Because of that we will set one parameter to be what we want.

$\boxed{\frac{2}{\|w\|}}$ $\rightarrow$ 2 over size or $\|w\|$

• We want to maximize the margin.

Minimize this to max

Maximaze margin $\Leftrightarrow$ minimize $\|w\|$

We want to create
the wedge the biggest possibe
we have express the wedge $= \frac{2}{\|\vec{w}\|}$

All $y_i = 1$ must have their x_i's above $L+1$.

$$\forall i \quad y_i = 1 \quad \vec{w} \cdot \vec{x_i} - (b+1) \geq 0$$
$$\Rightarrow \vec{w} \cdot \vec{x_i} - b \geq 1$$

trick multiply both sides by $\left(y_i - \frac{1}{2}\right)$

$$\Rightarrow \left(y_i - \tfrac{1}{2}\right)(\vec{w} \cdot x_i - b) \geq y_i - \tfrac{1}{2}$$

Since $y_i = 1 \Rightarrow \underbrace{\left(y_i - \tfrac{1}{2}\right)}_{\text{intentionally}}(\vec{w} \cdot x_i - b) \geq \frac{1}{2}$

All $y_i = 0$ next have their y_i's both $L > 1$

$$\vec{w_i} \cdot \vec{x_i} - (b-1) \geq 0$$
$$\Rightarrow \vec{w} \cdot \vec{x} - b + 1 \geq 0$$
$$\vec{w} \cdot \vec{x} - b \geq -1$$

Multiply both sides by $y_i = -\frac{1}{2}$

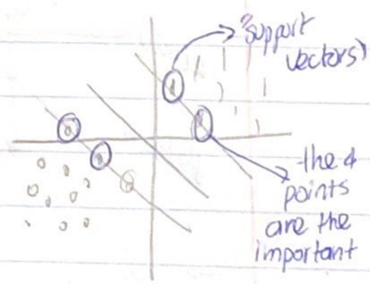
$$\left(y_i - \tfrac{1}{2}\right)(\vec{w} \cdot \vec{x} - b) \geq -y_i + \tfrac{1}{2}$$

Since $y_i = 0 \Rightarrow (y_i - 1/2)(\vec{w} \cdot \vec{x} - b) \geq 1/2$

$$\forall i \quad (y_i - 1/2)(w \cdot \vec{x_1} - b) \geq \tfrac{1}{2}$$

this do that

lower line on 0
upper line on 1's



this will make
all the 0 go
up and
all 1's go down.

american innovations

$$\min_{\|w\|} \text{ s.t. } \forall i \ (y_i - \tfrac{1}{2})(\vec{w}\,\vec{x}_i - b) \geq \emptyset$$

solve $\vec{w}$, b.

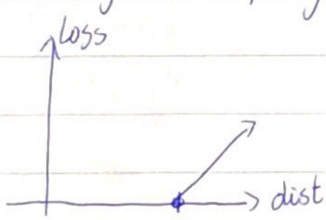
for svm is we allow violation

so if $i = .49 \neq 0.5$ yet we take it $\Rightarrow$ violation

$$H_i := \max\left\{\emptyset, \tfrac{1}{2} - (y_i - \tfrac{1}{2})(\vec{w}\cdot\vec{x}_i - b)\right\}$$ $\Rightarrow$ if condition satisfy error $\emptyset$

if this gives us error, the error is what we know how much we are missing

hinge error / hing loss



$\Rightarrow$ if $(y_i - \frac{1}{2})(\vec{w}\cdot\vec{x}_i - b) = \frac{1}{2} + d \not> \frac{1}{2}$

$$H = \max\left\{\tfrac{1}{2} - (\tfrac{1}{2} + d), 0\right\}$$
$$= \max\{+d, 0\} = \emptyset \quad d$$

Sum of hinge Error

$$SHE = \sum_{i=1}^{n} H_i = \sum_{i=1}^{n} \max\Big\{\emptyset, \underbrace{\tfrac{1}{2} - (y_i - \tfrac{1}{2})(\vec{w}\cdot\vec{x}_i - b)}_{\text{How far away from what I want.}}\Big\}$$

we can do sum or average $\Rightarrow$ AHE

A = minimize SHE

Find $\vec{w}$, b s.t. AHE is minimal.

listen lecture again

We want both min error (distance of the distance from line)
and maximun margin:
Vaphik (1963) proposed:

margin
$$\underset{\vec{w}, b}{\arg\min} \{ \underbrace{AHE}_{\text{mini average error}} + \lambda \underbrace{\|w\|^2}_{\text{max wedge}} \}$$

$\lambda \rightarrow$ hyperparameter / tunning parameter has to be explicit in both it is my choice.

From this

$\hookrightarrow g = A(\mathbb{D}, \mathcal{H}, \lambda)$

$\mathcal{Y} = \{0, 1 \ldots L\}, \ L > 2$

nominal

{Red, Green, blue, ...}

Mode ???

Null model: $g = \text{Mode}(y)$

$d > 0$
$d(x,y) = d(y,x)$
$d(x,x) = 0$

Consider the Model (nirest neighbor algorithm)

$$g(x_*) = y_i \text{ s.t. } i = \underset{i=1..n}{\arg\min} \{ d(\vec{x}_*, \vec{x}_i) \}$$

d → this is hipper parameter it
a distance function

it has to be > 0

Create to put x_* there

represents data
$x_* \rightarrow$ new data point

→ prediction new data poin

thing Bob

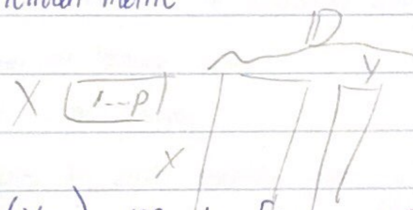
⇒ y_i that correspond x_i closser to x_*

$$d(\vec{X}_*, \vec{X_i}) = \sqrt{\sum_{i=1}^{p} (X_{*,j} - X_{i,j})^2}$$

Encliden metric

$\mathcal{H} = \{ \ \}$?

$A = \{ . . \}$



$g(X_*) = \text{Mode}[Y_1 - - Y_K]$ → new hipper parameter

where $Y_1 ... Y_K$ correspond to the K clossest observation in $\mathbb{D}$ to X_*.

K – new neighbor (KNN)

dict $a = first

$a → new

[1] " first