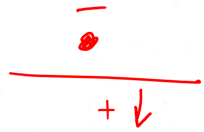
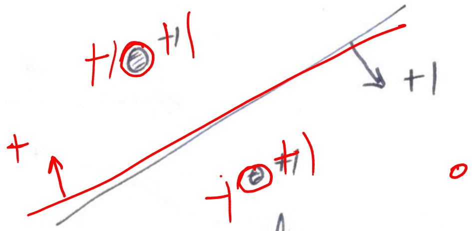
Learnability / Capacity of a binary classifier



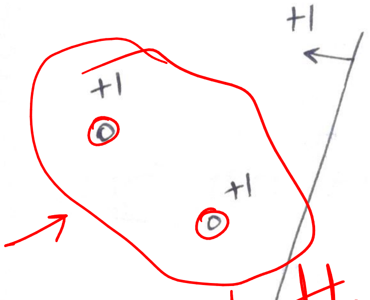
VC dimension (Vapnik Chervonenkis)

Consider 2 points

2D space



H: class of half space



A hyperplane class can shatter the configuration of 2 points

can be shattered.

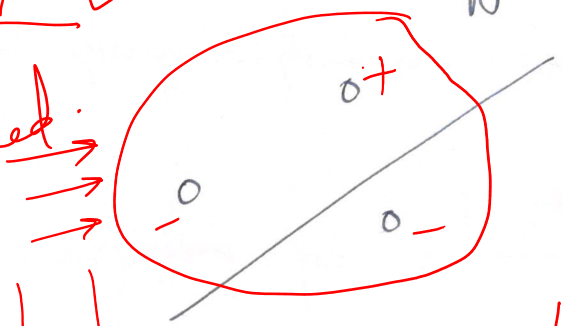
$\mathcal{H}$: set of all hyperplanes.

set of examples

$\mathcal{H}$ can shatter a finite set C iff for any labelling of the set C there is a hypothesis h in $\mathcal{H}$ which can correctly classify it.

eg1	eg2	
0	0	✓
0	1	✓
1	0	✓
1	1	✓

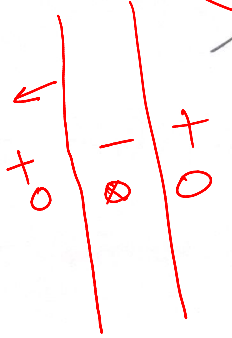
Can be shattered.

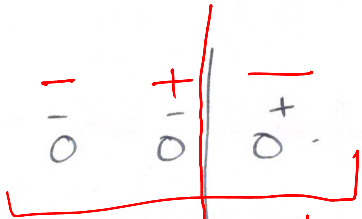


$\mathcal{H}$: half space in 2D space.

can be shattered by $\mathcal{H}$

$\mathcal{H}$: class of hyperplanes in 2D. (straight lines in 2D).





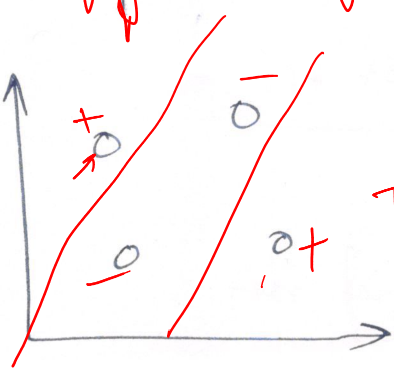
− + − cannot be shattered.

this configuration of 3 points cannot be shattered by any $h \in \mathcal{H}$

linear halfspaces

4 points in 2D space

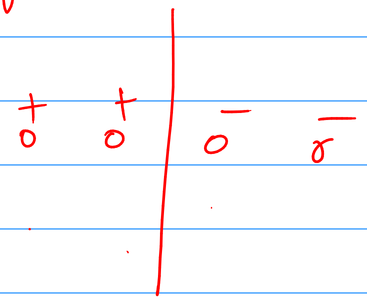
$+ + + +$
$+ + + -$
$+ + - -$

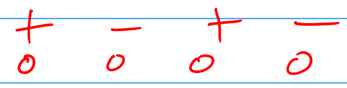


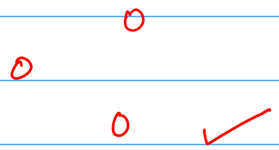
This configuration of 4 points cannot be shattered

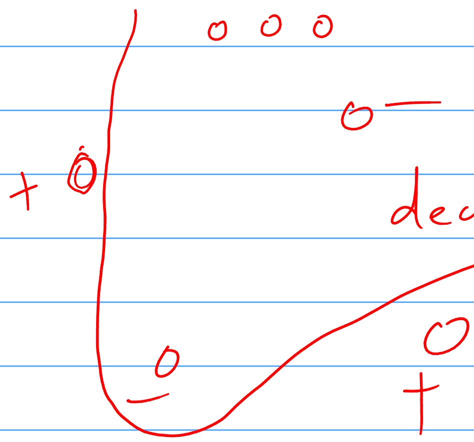
Shattering

1. Fix the examples (configuration) is our choice.
2. Label them in all possible ways.
   and find $h \in H$ which can correctly classify them







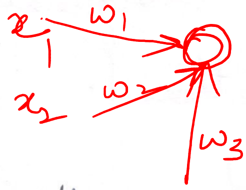


A curved decision boundary will work.

There is no configuration of 4 points which can be shattered

A set of 4 points cannot be shattered by our hypothesis class of straight lines in a 2D space.

A straight line in 2D space has a representation



$$ax + by + c = 0$$

or $w_1 x + w_2 y + w_3 = 0$

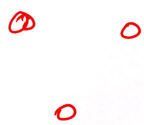
Non-homogeneous representation will have the parameters $\underline{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ and bias w_3.

The non homogeneous space is $\mathbb{R}^2$.
i.e. the hyper plane exists in nonhomogeneous space $\mathbb{R}^2$.

Homogeneous representation has $\underline{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$ i.e. the hyperplane exists in the homogeneous space $\mathbb{R}^3$.

its a property of the $\mathcal{H}$ class.

VC dimension of $\mathcal{H} \equiv$ hyperplanes/half space classifier
in non-homogeneous $\mathbb{R}^2$
or homogeneous $\mathbb{R}^3$
is 3.



i.e. there exists some configuration of 3 points for which
all possible labellings of the examples.
any possible mapping $x \rightarrow y$ has a corresponding h available in our hypothesis class $\mathcal{H}$

In general,
for a homogeneous half space in $\mathbb{R}^d$, the VC dimension is d.

I To show this, we need to first show that our hypothesis class can shatter a set of d points.

II And secondly we need to show that there is no configuration of a set of $d+1$ points that can be shattered.

Proof: $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$ homogeneous.

Let's first come up with a configuration of d points.

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \cdots \quad e_i = \begin{array}{r} \\ \\ \\ \\ i^{th} \rightarrow \\ \\ \\ \end{array}\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad e_d = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Our choice

Consider an arbitrary labelling for these d points.

$$y = \begin{array}{r} e_1 \\ e_2 \\ \\ \\ \\ e_d \end{array}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_d \end{bmatrix} \longrightarrow \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Our hypothesis class $\mathcal{H}$ of half spaces in $\mathbb{R}^d$ has a vector $\underline{w} = \underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{bmatrix}$ which satisfies $\boxed{\langle \underline{w}, \underline{e_i} \rangle = y_i \quad \forall i}$ GT

Correct labelling

$\underline{w} = \underline{y}$

$\Rightarrow$ the set of d points has been shattered.

2^{nd} part

Now consider a set of $d+1$ points in $\mathbb{R}^d$

$\underline{x_1} \quad \underline{x_2} \quad \underline{x_3} \quad \cdots \quad \underline{x_{d+1}}$ our choice

The vectors corresponding to these points are linearly dependent.

$$\sum_{i=1}^{d+1} a_i \underline{x_i} = \underline{0}$$

a_i's are not all zeros.

We group these a_i's into 2 sets

$$I : \{ i : a_i > 0 \}$$

$$J : \{ j : a_j < 0 \}$$

a_1	3.1
a_2	4.5
a_3	−2
⋮	
a_{d+1}	

$$\therefore \sum_{i \in I} a_i x_i + \sum_{j \in J} a_j x_j = 0$$

absolute value

or

$$\sum_{i \in I} a_i x_i - \sum_{j \in J} |a_j| x_j = 0$$

$$\therefore \boxed{\sum_{i \in I} a_i x_i = \sum_{j \in J} |a_j| x_j}$$

valid for any configuration of $d+1$ points

Now consider an arbitrary labelling assigned to these points, $\underline{x_1}$ $\underline{x_2}$ ... $\underline{x}_{d+1}$

All points x_i $i \in I$ get $y_i = 1$ $\quad \therefore \langle \underline{w}, x_i \rangle > 0$

All points x_j $j \in J$ get $y_j = -1$ $\quad \therefore \langle \underline{w}, x_j \rangle < 0$

No $\underline{w}$ can achieve correct classification of $d+1$ points for this labelling.

Assume that there exists a hypothesis $\underline{w} \in \mathcal{H}$ such that $\underline{w}$ correctly classifies all these points.

For the points in set I

$$a_i > 0 \qquad \langle \underline{w}, \underline{x}_i \rangle > 0$$

∵ we insist on correct classification

$$\sum_{i \in I} \underset{+ve}{a_i} \underbrace{\langle \underline{w}, \underline{x}_i \rangle}_{>0} > 0$$

$$\Rightarrow \sum_{i \in I} \langle \underline{w}, a_i \underline{x}_i \rangle > 0$$

$$\Rightarrow \langle \underline{w}, \underbrace{\sum_{i \in I} a_i \underline{x}_i}_{=0} \rangle > 0$$

$$\Rightarrow \langle \underline{w}, \underbrace{\sum_{j \in J} |a_j| \underline{x}_j}_{} \rangle > 0 \quad \Rightarrow 0$$

$$\Rightarrow \underbrace{\sum_{j \in J} \underbrace{|a_j|}_{>0} \underbrace{\langle \underline{w}, \underline{x}_j \rangle}_{<0}}_{} > 0.$$

Can not hold ✗.

Contradiction. ∵ for correct classification $\langle \underline{w}, \underline{x}_j \rangle < 0$

Correct classification of all the points in set J is not possible. (d+1 points).

∴ Our assumption that $\underline{w}$ correctly classifies all points in set I and J has been contradicted

∴ $d+1$ points are not shattered by the hypothesis class $\mathcal{H}$ of linear functions in $\mathbb{R}^d$.