Recap

15 October 2025 19:41

$$\underbrace{ax + by}\underbrace{+c} = 0$$

$$\underbrace{w_1 x + w_2 y}_{\text{weights}} + \underbrace{w_0}_{\text{bias}} = 0$$

$$\left\{ w_1 x_1 + w_2 x_2 + w_3 x_3 + w_4 x_4 \ldots\ldots w_0 = 0 \right\} \Rightarrow \text{hyperplane}$$

$$\left\{ \begin{array}{l} a \cdot b \Rightarrow \text{directional similarity} \\ a \cdot b = \|a\| \, \|b\| \cos\theta \end{array} \right\}$$

i) Hands-on

$\underbrace{\text{Maths on ML}}$

[Purely Math ⇒ Interview]

ii) Take doubts

Matrix Multiplication

18 August 2025 20:12

$$\underset{M1}{\begin{array}{c} V_1 \Rightarrow \\ V_2 \Rightarrow \end{array}\begin{bmatrix} 1 & 3 \\ 4 & 5 \end{bmatrix}} \times \underset{M2}{\begin{array}{cc} V_3 & V_4 \\ \begin{pmatrix} 2 \\ 8 \end{pmatrix} & \begin{pmatrix} 7 \\ 5 \end{pmatrix} \end{array}} = \begin{bmatrix} V_1 \cdot V_3 & V_1 \cdot V_4 \\ V_2 \cdot V_3 & V_2 \cdot V_4 \end{bmatrix}$$

$$\begin{bmatrix} 2+24 & 7+15 \\ 8+40 & 28+25 \end{bmatrix}$$

(No. of columns) ☆= (No of rows)

$$[\ \ ]_{m \times n} \times [\ \ ]_{a \times b} = [\ \ ]_{m \times b}$$

rule → $n = a$

$$\begin{bmatrix} 1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$a \cdot b$

$[a^T \cdot b]$

$$\begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$ } (follow rule for Matrix Multiplication)

Transpose :

Swapping of rows and columns is called a transpose

$$\begin{bmatrix} \begin{pmatrix} 1 \\ 4 \end{pmatrix} & \begin{pmatrix} 2 \\ 6 \end{pmatrix} \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 4 \\ 2 & 6 \end{bmatrix}$$

Dot Product of Vector with itself

15 October 2025 19:25

$$1 \quad \begin{bmatrix} 4 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 4 \end{bmatrix} \Rightarrow 32$$

Which of two give a higher dot product?

$$2 \quad \begin{bmatrix} 4 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ 9 \end{bmatrix} \Rightarrow 28 + 36 = 64$$

But by looks of it which two are more similar?

2

Why?

Magnitude is driving the dot product

$\Downarrow$

wrong interpretation.

[Neutralize the effect of Magnitude]

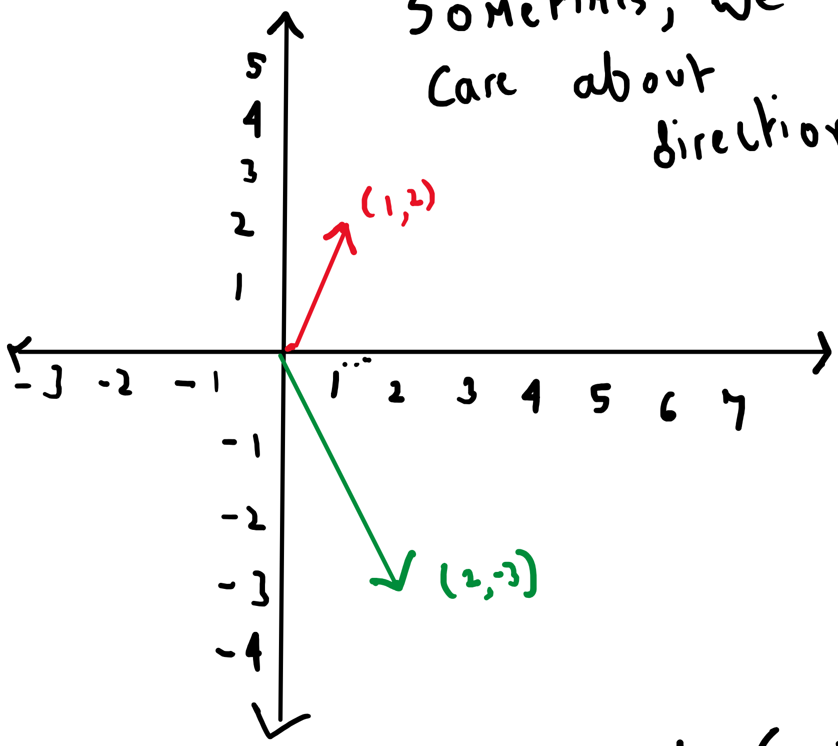
{Unit vectors}

Vector with only directional component

Magnitude = 1

Unit vector

19 August 2025 08:03



Sometimes, we only care about direction

$V_1 = \begin{bmatrix}1\\2\end{bmatrix}$

$V_2 = \begin{bmatrix}2\\-3\end{bmatrix}$

$$\frac{1}{\sqrt{5}}\begin{pmatrix}1\\2\end{pmatrix} =$$

What are magnitudes of the 2 vectors?

$$V_1 \begin{bmatrix}1\\2\end{bmatrix} \qquad \|V_1\| \Rightarrow \sqrt{1^2+2^2} = \sqrt{5}$$

$$\left\{ V_2 \begin{bmatrix}2\\-3\end{bmatrix} \right\} \qquad \|V_2\| \qquad \sqrt{2^2+(-3)^2} = \sqrt{13}$$

$$\|V_1\| \times V_1 = \left(\frac{1}{\sqrt{5}}\right)\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}1/\sqrt{5}\\2/\sqrt{5}\end{bmatrix} = \hat{V}_1 \Bigg\} \quad \frac{1}{\sqrt{13}}\begin{bmatrix}2\\-3\end{bmatrix} = \begin{bmatrix}\frac{2}{\sqrt{13}}\\ \frac{-3}{\sqrt{13}}\end{bmatrix}$$

What if I want to transform the vector such that the direction stays intact but magnitude becomes 1?

Why? Because sometimes we don't care about magnitude, and only care about direction.

represented by $\wedge$ on top

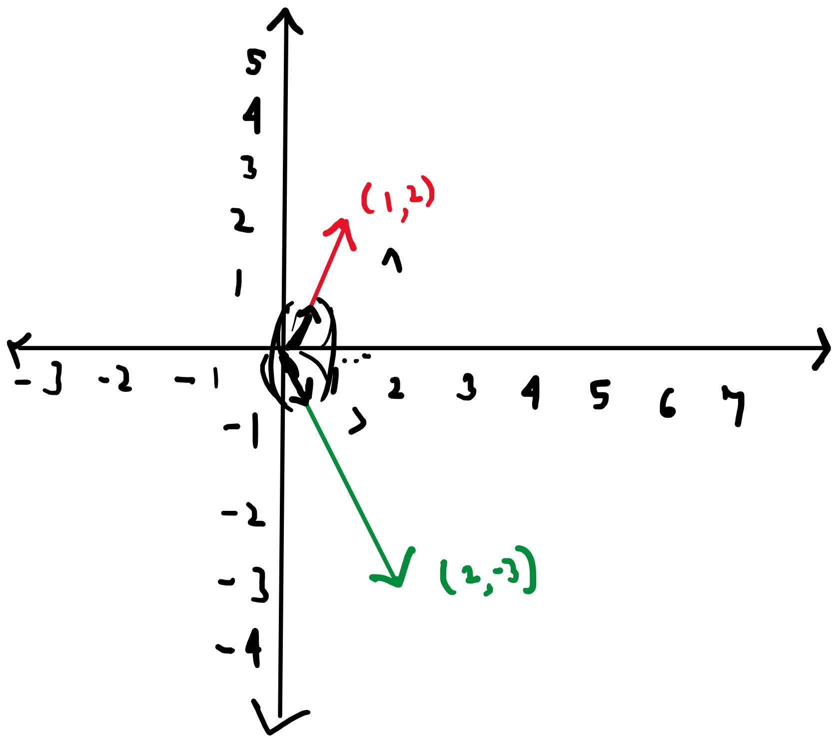
$\vec{w} \Rightarrow$ original

$\hat{w} \Rightarrow$ unit

$$\hat{w} = \frac{1}{\{\|w\|\}} \times \vec{w}$$

$$\hat{w} = \frac{1}{\|w\|} \times \vec{w}$$

What exactly am I doing here?



$$\hat{v} \quad \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} =$$

$$\sqrt{\left(\frac{1}{\sqrt{5}}\right)^2 + \left(\frac{2}{\sqrt{5}}\right)^2}$$

$$= \sqrt{\frac{1}{5} + \frac{4}{5}}$$

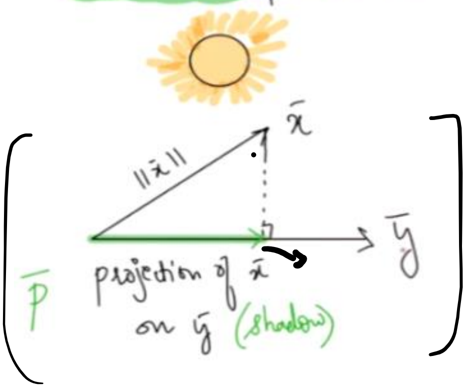
$$= \sqrt{\frac{5}{5}} = \sqrt{1}$$

$\left[\text{Unit vector}\right]$ $= 1$

If you only care about direction

$\left[\hat{a} \cdot \hat{b}\right]$ $\dfrac{a \cdot b}{(\hat{a} \cdot \hat{b})}$ → unit vectors

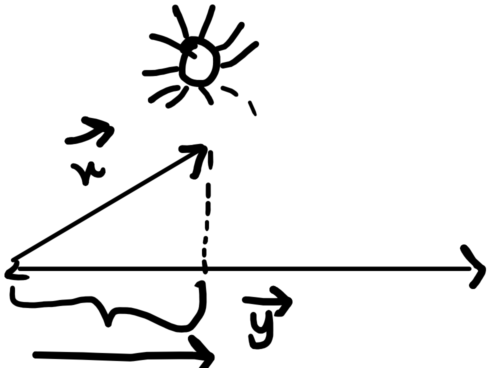
Projection of a Vector



$\begin{pmatrix} 2 \\ 4 \end{pmatrix} \quad \begin{pmatrix} 5 \\ 2 \end{pmatrix}$

$$\bar{x}\cdot\bar{y} = \bar{x}^T\bar{y}$$

$$\cos\theta = \frac{\bar{x}^T\bar{y}}{\|\bar{x}\|\,\|\bar{y}\|}$$



(length or shadow)

[Projection of a Vector]

$$\cos\theta = \frac{b}{h} = \frac{\text{length of Projection}}{\|x\|}$$

$$\|x\| \cos\theta = \text{length of projection} \Rightarrow 1$$

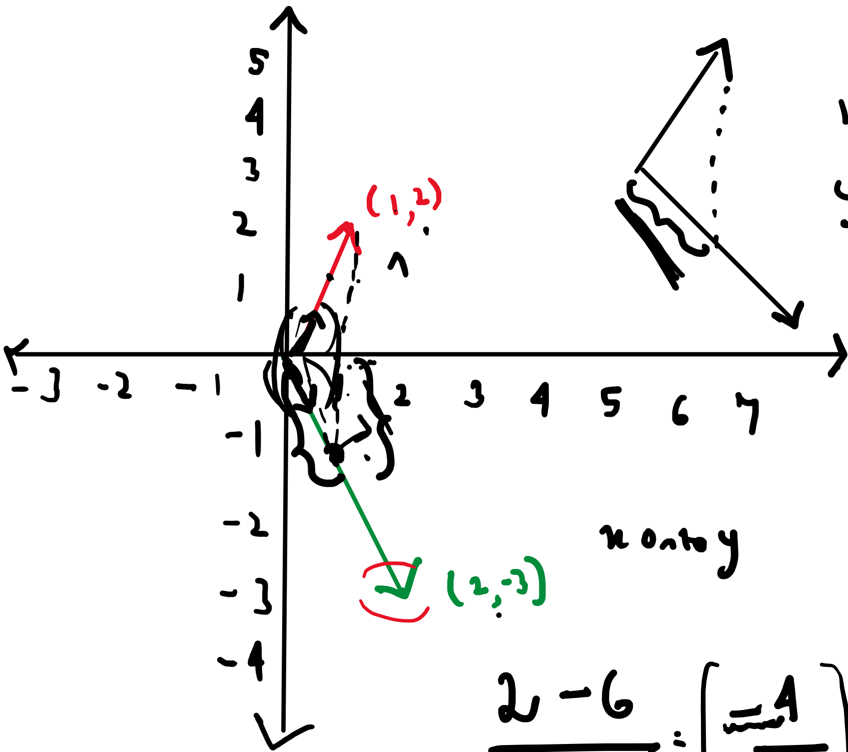
$$x\cdot y = \|x\|\,\|y\|\cos\theta$$

$$\left(\cos\theta = \frac{x\cdot y}{\|x\|\,\|y\|} \Rightarrow 2\right)$$

$$\cancel{\|x\|} \times \frac{x\cdot y}{\cancel{\|x\|}\,\|y\|} = \left[\frac{x\cdot y}{\|y\|} = \text{length of the projection on shadow}\right]$$

Projection of x onto $y = \dfrac{x \cdot y}{\|y\|}$:

Projection of y onto $x = \left(\dfrac{x \cdot y}{\|x\|}\right)$



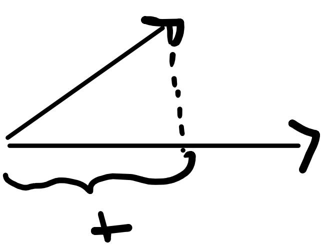
$x\,(1 \quad 2)$

$y \cdot (2 \; -3)$

x onto $y \quad \dfrac{x \cdot y}{\|y\|}$

$\dfrac{2-6}{\sqrt{12}} = \left(\dfrac{-4}{\sqrt{13}}\right) \qquad \dfrac{(-)4}{\sqrt{13}}$

$\left(\dfrac{4}{\sqrt{13}}\right)$



(PCA) (SVM)

[Dot product + projection]

(ML Algorithms)

Cosine Similarity

15 October 2025 19:35

$$\underbrace{\dot{A}\cdot\dot{B}} = \underbrace{\|\dot{A}\|}\ \underbrace{\|\dot{B}\|}\ \underbrace{\cos\theta}$$

$-1 = \cos\theta$

$\theta = 180°$

$$\left\{ \left(\frac{\dot{A}\cdot\dot{B}}{\|\dot{A}\|\ \|\dot{B}\|} \right) = \text{cosine similarity} \right\}$$

$-1 \Rightarrow$ opposite directions

to

$1 \Rightarrow$ Same direction.

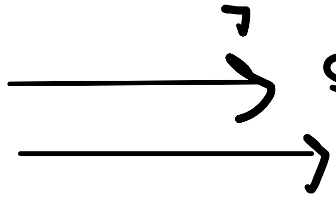
$\cos\theta = 1$

$\underbrace{\theta = 0°}$

$\Downarrow$

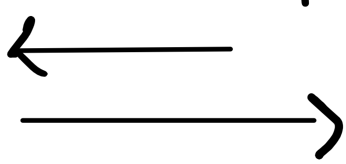
Same directions.

[Cosine similarity] $\Rightarrow$ Directional similarity



Same direction. $\Rightarrow 0°$ $\quad 1$

cos



opposite direction

$\Rightarrow 180°$

cos

-1

Role of Similarity (Interview important)

15 October 2025 19:29

Role of similarity in ML?

Let us say instagram knows you and I are similar. How would it use this info?

[Same reels, same ads, same content, Suggesting friends.]

Two popular similarity metrics we have studied:

i) Eucledian distance → Magnitude → Closer → similar

ii) Cosine similarity → angle ⇒ direction

Customer	Groceries	Electronics	Clothing
A 4:1:2	400	100	200
B 4:1:2	40	10	20
C	200	250	50

⇒ {A and C}

Which one to use here?

(a) A–B

$$d_E(A,B)=\sqrt{(400-40)^2+(100-10)^2+(200-20)^2}$$

Step-by-step:

- $(400 - 40)^2 = 360^2 = 129{,}600$
- $(100 - 10)^2 = 90^2 = 8{,}100$
- $(200 - 20)^2 = 180^2 = 32{,}400$

Sum = 129,600 + 8,100 + 32,400 = 170,100

$$d_E(A,B)=\sqrt{170{,}100}\approx 412.4$$

✅ **Euclidean distance (A,B) ≈ 412.4**

(b) A–C

$$d_E(A,C)=\sqrt{(400-200)^2+(100-250)^2+(200-50)^2}$$

A and C

Step-by-step:

- $(400 - 200)^2 = 200^2 = 40{,}000$
- $(100 - 250)^2 = (-150)^2 = 22{,}500$
- $(200 - 50)^2 = 150^2 = 22{,}500$

Sum = 40,000 + 22,500 + 22,500 = 85,000

$$d_E(A,C)=\sqrt{85{,}000}\approx 291.5$$

✅ **Euclidean distance (A,C) ≈ 291.5**

(a) A–B

Dot product:

$$A\cdot B=(400\times 40)+(100\times 10)+(200\times 20)=16{,}000+1{,}000+4{,}000=21{,}000$$

Magnitudes:

$$\|A\|=\sqrt{400^2+100^2+200^2}=\sqrt{160{,}000+10{,}000+40{,}000}=\sqrt{210{,}000}\approx 458.26$$

$$\|B\|=\sqrt{40^2+10^2+20^2}=\sqrt{1{,}600+100+400}=\sqrt{2{,}100}\approx 45.83$$

Cosine similarity:

$$\cos(A,B)=\frac{21{,}000}{458.26\times 45.83}=\frac{21{,}000}{21{,}014.6}\approx 0.9993$$

✅ **Cosine similarity (A,B) ≈ 0.999 → almost identical pattern** ≈ 1

(b) A–C

Dot product:

$$A\cdot C=(400\times 200)+(100\times 250)+(200\times 50)=80{,}000+25{,}000+10{,}000=115{,}000$$

Magnitude of C:

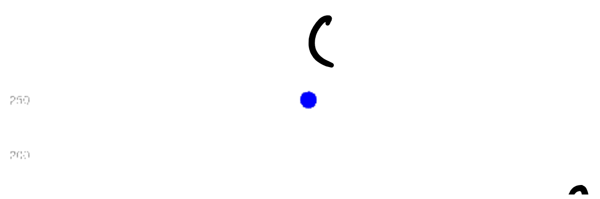
$$\|C\|=\sqrt{200^2+250^2+50^2}=\sqrt{40{,}000+62{,}500+2{,}500}=\sqrt{105{,}000}\approx 324.04$$

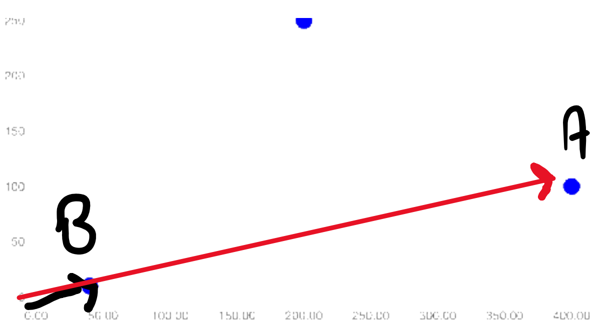
Cosine similarity:

$$\cos(A,C)=\frac{115{,}000}{458.26\times 324.04}=\frac{115{,}000}{148{,}464.9}\approx 0.7745$$

✅ **Cosine similarity (A,C) ≈ 0.77 → moderate pattern overlap**

[A and C ⇒ Euclidean
A and B → cosine] Depends !!





When will you use Eucledian vs Cosine ?
Magnitude of spending on Minutes of some reel or video ⇒ Eucledian

Reccomend products or reels based on Qualitative
Cosine similarity

Consider a large equation

$$\left[\left\{ \underbrace{\left[w_1 x_1 + w_2 x_2 + w_3 x_3 \cdots\cdots w_{10} x_{10} \right]}_{} + w_0 = 0 \right\} \right]$$

What we call this in ML lingo? $\Rightarrow$ [equation of hyper plane]

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ \vdots \\ \vdots \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ \vdots \end{bmatrix}$$

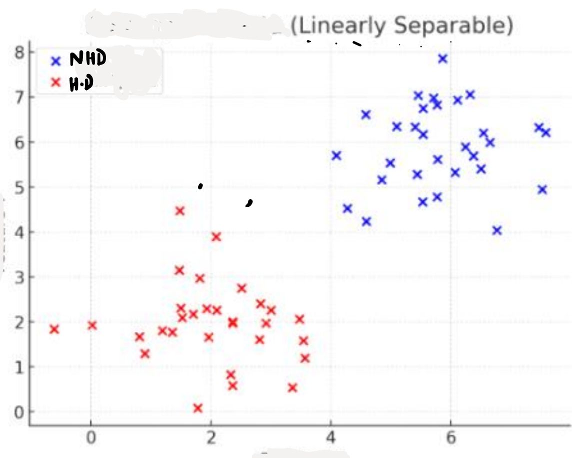
feature vector

$$w_1 x_1 + w_2 x_2 + w_3 x_3 + w_4 x_4 \cdots\cdots + w_0$$

[weight vector] [vector of x]

$$\left[\vec{w}^T \cdot \vec{x} + w_0 = 0 \right]$$

$[\quad] \rightarrow \quad \uparrow [\quad]$



y

Exercise hours

Sleep hours x

$$w^T x + w^0$$
$\Downarrow$
feature vector.

feature vector change or same?

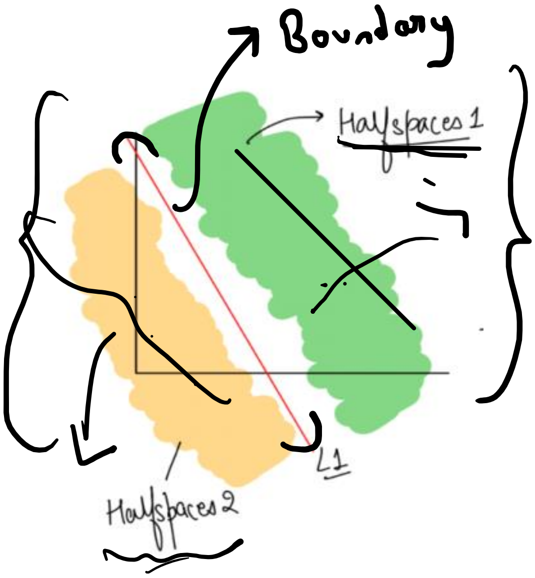
features

$w_1 x_1 + w_2 x_2 + w_3 x_3 \cdots\cdots$

Sleep duration + BMI + age

Half spaces

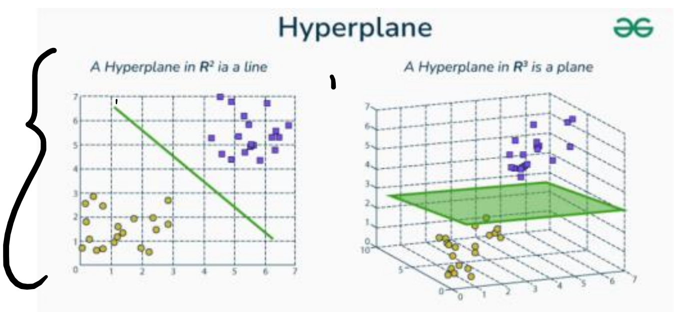
24 August 2025 20:37



Boundary

A line divides space into 2 portions

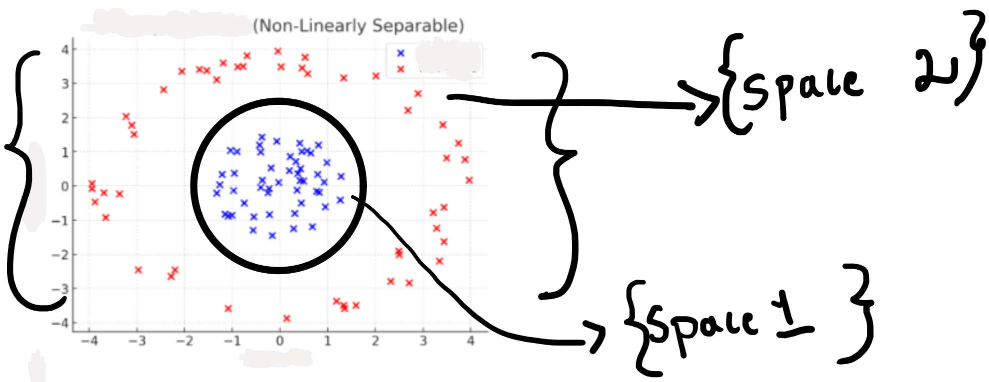
Each portion is called half space



Similarly in 3D we can divide the space into 2 parts using a Plane

- One which lines above the plane
- other is one that lies below the plane

Do non-linear Models have a concept of half space?



Half Space - where is the point?

21 August 2025 21:09

Healthcare case ⇒ you have built an ML model to classify whether a person is under (risk of Heart Disease) using Sleep hours and Alcohol drinks.

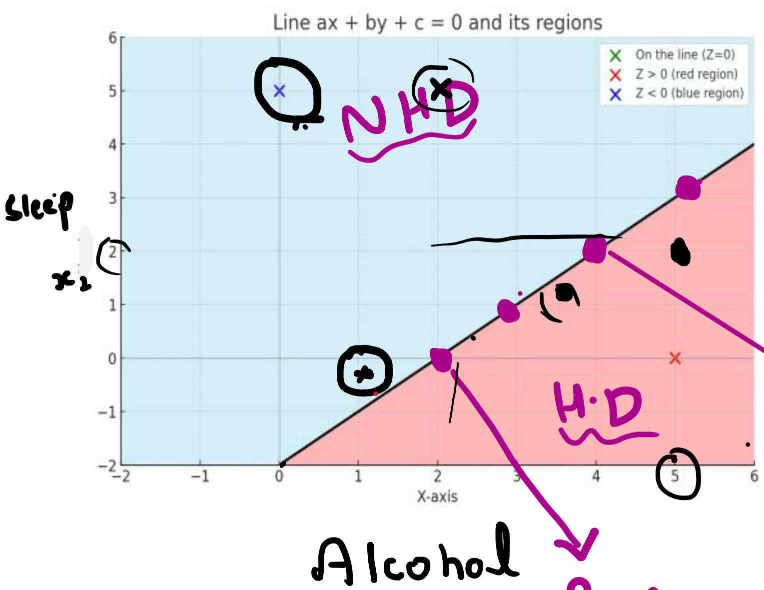
{ML algorithm has found the boundary [$x_1 - x_2 - 2 = 0$] } ?

equation

Extent of disease

Alcohol

Sleep



Bad ⇒ 50-50

[What is one combination of x_1, x_2 which satisfies the equation?]

$$x_1 - x_2 - 2 = 0$$

$(4, 2)$
$x_1 \quad x_2$

$$[4 - 2 - 2 = 0]$$

$x_1 \quad x_2$

How about $(5, 2)$?

$$\underbrace{x_1 - x_2 - 2}$$

$1 > 0$ or < 0

$> 0 \Rightarrow$ H.D

How about $(2, 5)$?

$x_1 - x_2 - 2$

$2 - 5 - 2$

$-3 - 2 = -5$

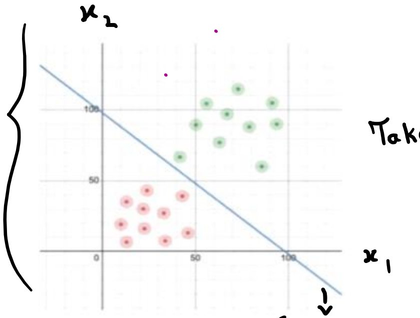
$$\underbrace{-5 < 0}_{\text{NHD}}$$

What does the sign (+ or -) of the result tell us?

Signed distance

Relationship between the Weight vector and the boundary

21 August 2025 20:30



(Boundary)

(x_1, x_2)

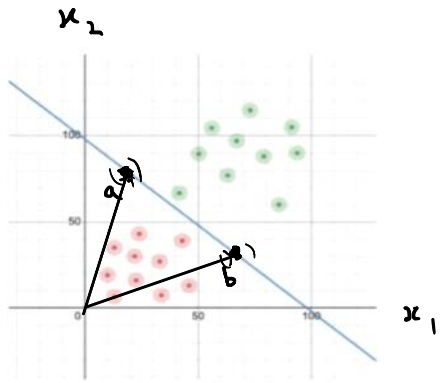
Take a boundary line $(w_1 x_1 + w_2 x_2 + w_0 = 0)$

What is the vector form?

$$\{W^T x + W_0 = 0\}$$

feature vector

What does $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ represent?



Take 2 points

$\cdot$ $a \Rightarrow \begin{pmatrix} x_{1a} \\ x_{2a} \end{pmatrix}$

and

$\cdot$ $b \Rightarrow \begin{pmatrix} x_{1b} \\ x_{2b} \end{pmatrix}$

such that they both lie on the line

Since they lie on the line,

(1) $W^T a + W_0 = 0$ (2) $W^T b + W_0 = 0$

(subtract (1) and (2))

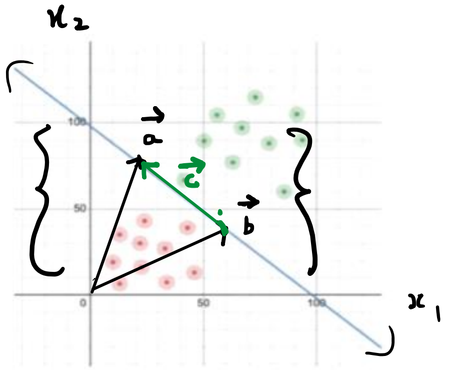
$$[(W^T a + W_0) - (W^T b + W_0) = 0]$$

$$W^T a + W_0 - W^T b - W_0 = 0$$

$$(W^T a - W^T b = 0)$$

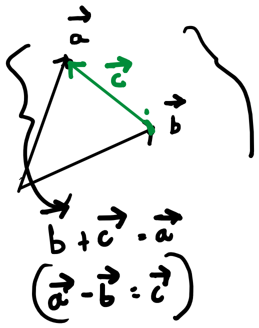
$$(W^T \cdot (a - b)) = 0$$

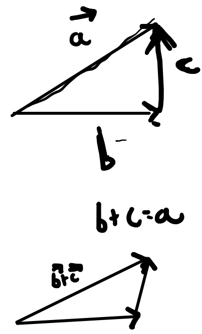
Let us say a and b are $\in$ feature vectors.



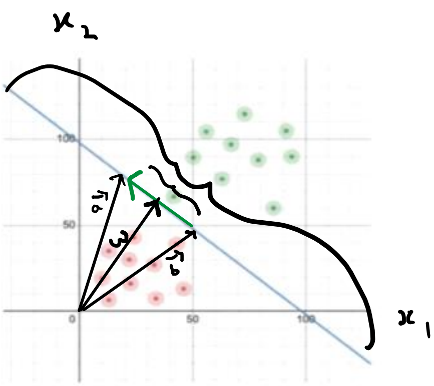
$$\vec{b} + \vec{c} = \vec{a}$$

$$\vec{a} - \vec{b} = \vec{c}$$





$$W^T \cdot C = 0$$



Through vector addition / subtraction, C has to be on the boundary line

So the boundary line is just the extension of $\vec{c}$?

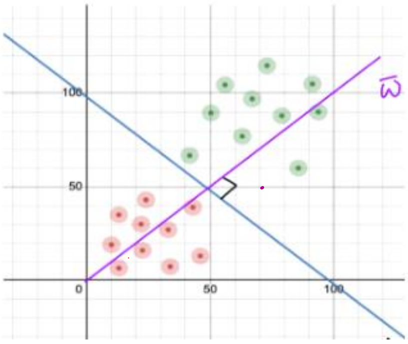
$$W^T(a-b) = 0$$

becomes

$$W^T \cdot c = 0$$

If dot product $= 0$

What is angle b/w them?

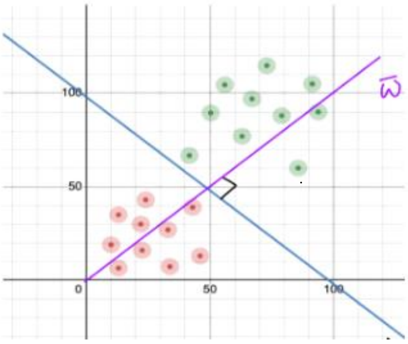


Different Boundaries will have different weight vectors

$w_1x_1 + w_2x_2 + w_0 = 0$

$\begin{bmatrix} 2 \\ 3 \end{bmatrix} \leftarrow \underbrace{2x_1 + 3x_2}_{} = 0$

$\begin{bmatrix} 3 \\ 5 \end{bmatrix} \leftarrow 3x_1 + 5x_2 = 0$



Angles can help identify the halfspace of a point.

Important property of weight vector:

Always points towards positive half space.

[SVM, logistic regression property]

$$\left\{\begin{array}{ccc} x_1 & x_2 & y \\ 1 & 2 & 1 \\ 3 & 4 & 0 \\ 2 & 2 & \end{array}\right\}$$

$$\left\{\begin{array}{l} x = 2 \\ x = -2 \end{array}\right\}$$

$x = 2$ ✓

$x = -2$

$(x_1 - x_2 = 4$ ✓

$$x = 2$$

$$[x_1 - 2 = 0]$$

$$\left\{\begin{array}{l} x_1 = 2 \\ x_2 = 4 \\ x_1 - x_2 = 2 \end{array}\right\}$$

(Equation) or (Boundary)

lines

x_1	x_2	Labels (y)	
1	2	0	-1 ✓
3	1	1	1 ✓
0	0	0	-2 ✓
4	0	1	2 ✓

optimal line

$$\left\{\begin{array}{l} x_1 = 0.5 \\ x_2 = 1 \\ x_1 = 2 \checkmark \\ (x_2 = x_1 + 2) \end{array}\right\}$$

(Boundary lines)

$$(x_1 - 2) = -1$$

$$x_1 = 2$$

$$(x_1 - 2 = 0)$$