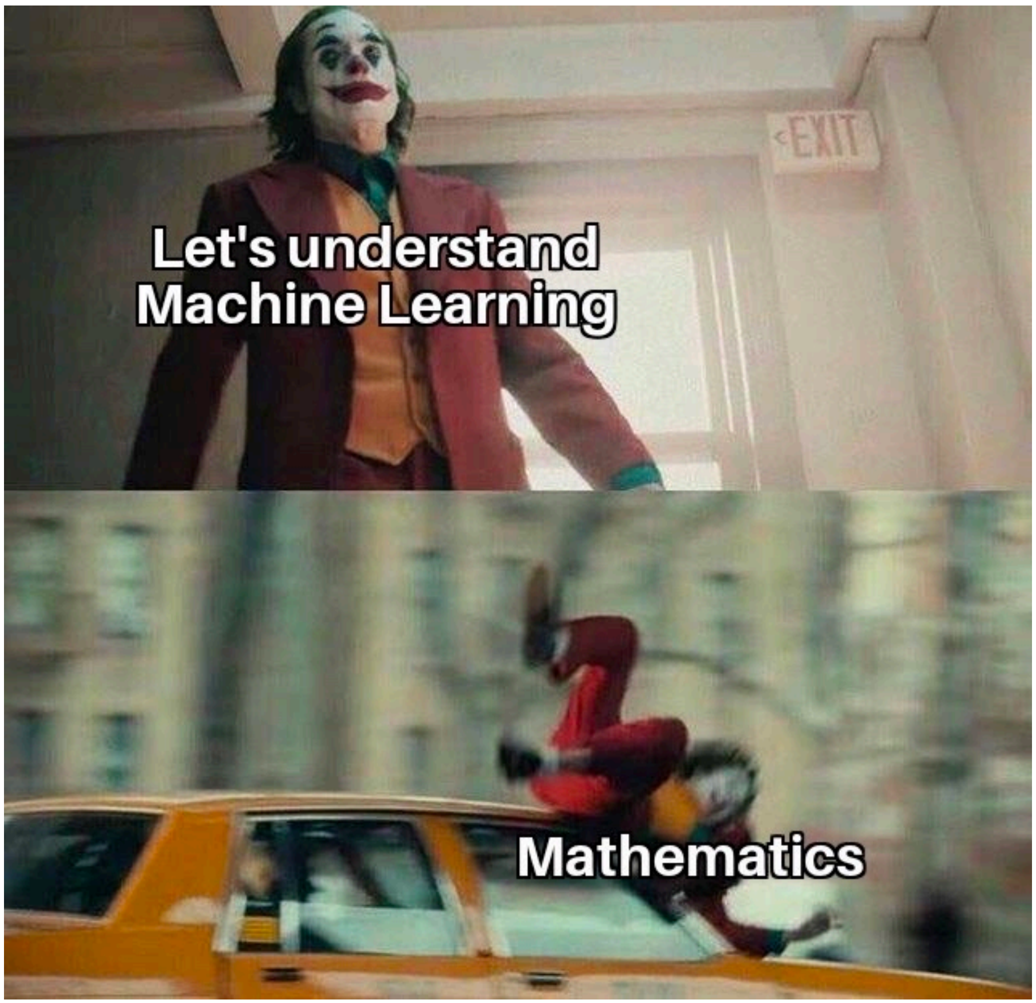
February 18, 2023 DSML : Math for ML.

Linear Algebra : Dot product and Hyperplanes.



Recap:

* The Machine Learning Context : training classifiers.

* New terms :
  (a) Feature
  (b) Label
  (c) Datapoint
  (d) Dataset
  (e) Classifier.

* How to train a classifier? Optimization !!

5 classes to study this !

slope

Lines.
* $y = mx + c \rightarrow$ y-intercept. ← ~~x~~

features.

* $w_1 x_1 + w_2 x_2 + w_0 = 0.$ ←

parameters.

Recap: Let's optimize a classifier by hand.

$x + y = 98 \rightarrow$ the line which.

$$w_1 x_1 + w_2 x_2 + w_0 = 0$$

$\rightarrow$ $w_1 = 1$, $w_2 = 1$, $w_0 = -98$.

x_1 : width of fish

y_1 : length of fish.

x_1	y_1	z_1
62	71	31
66	88	36
80	55	35
99	75	
98	81	
79	94	
51	61	
57	90	
82	67	
64	93	

1. Will the line equation be able to separate these out? No!
2. What can do this? A plane.
3. What is the equation of a plane?

(2 D) : $w_1 x_1 + w_2 x_2 + w_0 = 0$ → General "line" equation.

A 2-D Hyperplane.

(3 D) : $w_1 x_1 + w_2 x_2 + w_3 x_3 + w_0 = 0$.

↓

General "plane" equation

OR.

A 3-D Hyperplane.

(d D) : $w_1 x_1 + w_2 x_2 + \dots + w_d x_d + w_0 = 0$.

↓

General equation for a d-dimensional hyperplane!

Linear Algebra: Vectors.

List, tuple, numpy arrays etc.

What is a vector?

It is a $\underbrace{\text{collection of numbers}}_{\text{array}}$.

How to represent them? Coding

$\bar{x}$ ← the bar denotes a vector.

If $\bar{x}$ has d dimensions, we can write $\bar{x}$ as:

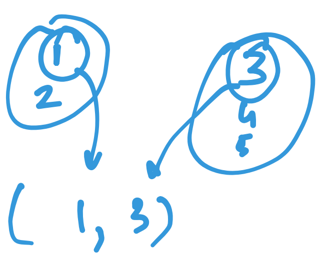
$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

column vector

$$\boxed{\bar{x} \in \mathbb{R}^d}$$

→ It means "belongs to"

$\mathbb{R}$ → The set of all real numbers.



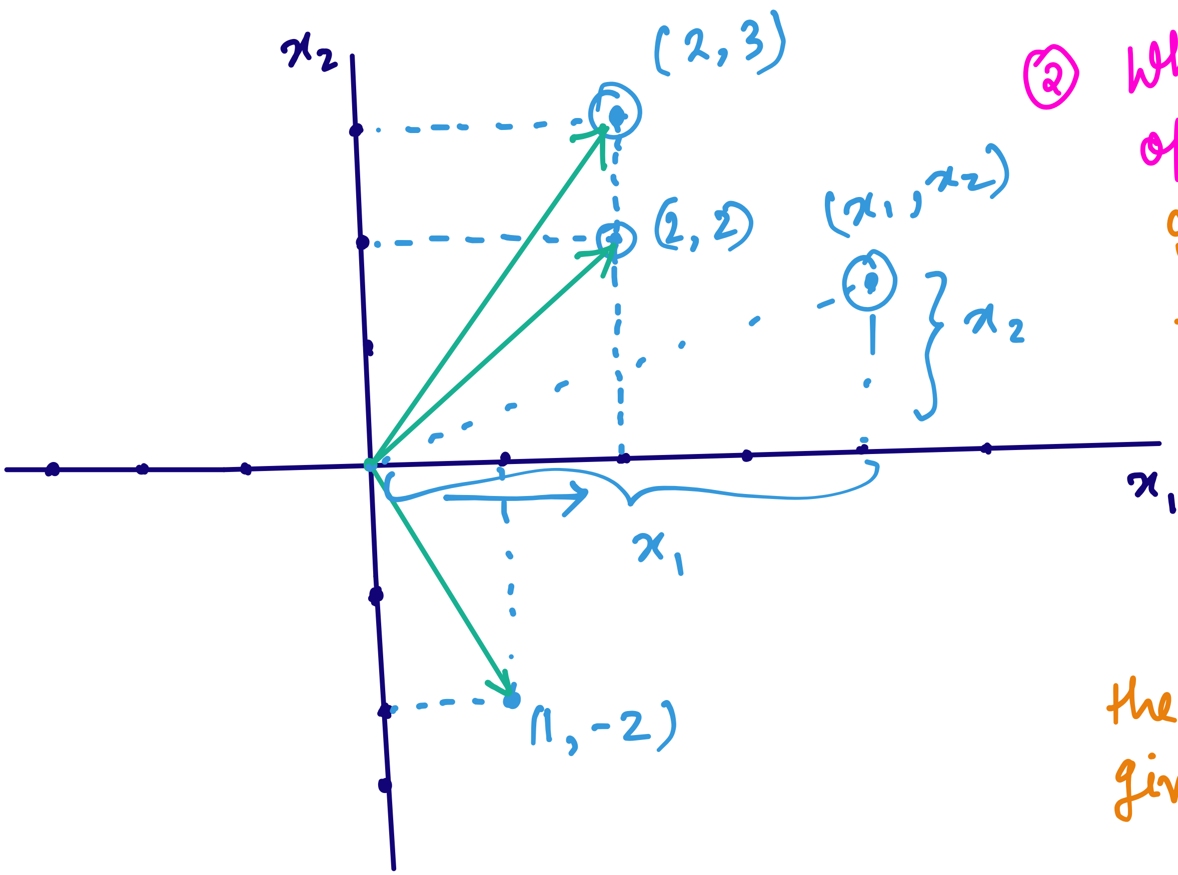
$\bar{x} = [x_1 \quad x_2 \quad x_3 \ldots x_d]$ → row vector.

Linear Algebra: Vectors

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

① How to visualize?

$$\bar{x}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \qquad \bar{x}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \qquad \bar{x}_3 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$



② What is the magnitude of the vector?

It is the length of the line joining the origin to the point.

↓

For 2-D vectors, the magnitude can be given by: $\sqrt{x_1^2 + x_2^2}$

Linear Algebra: Vectors

③ How to get magnitude of the vector in d-dimension?

For 2-D, $\bar{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$\text{magnitude}(\bar{x}) = \sqrt{x_1^2 + x_2^2}.$$

For 3-D, $\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$$\text{magnitude}(\bar{x}) = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

For d-D, $\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_d \end{bmatrix}$

$$\text{magnitude}(\bar{x}) = \sqrt{x_1^2 + x_2^2 + x_3^2 + \cdots + x_d^2}$$

Linear Algebra : Norm of a vector.

* The norm of a vector is its length!!
  ↓
  magnitude.

* We write the norm of a vector using the following symbols:

  norm of $\bar{x} \rightarrow \|\bar{x}\|$

* The formula for the L2 norm of a vector is:

$$\|\bar{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_d^2}$$

L_2 norm similar to euclidean.

we will mostly use this.

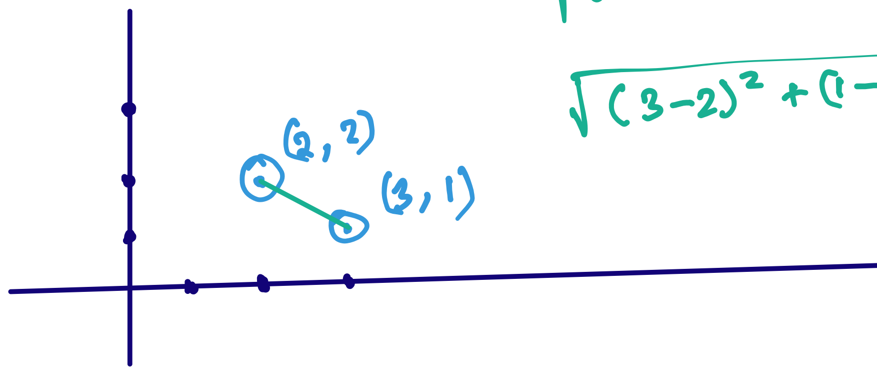
$$\|\bar{x}\| = |x_1| + |x_2| + \cdots + |x_d|$$

L_1 norm, similar to manhattan.

Different types of distances:

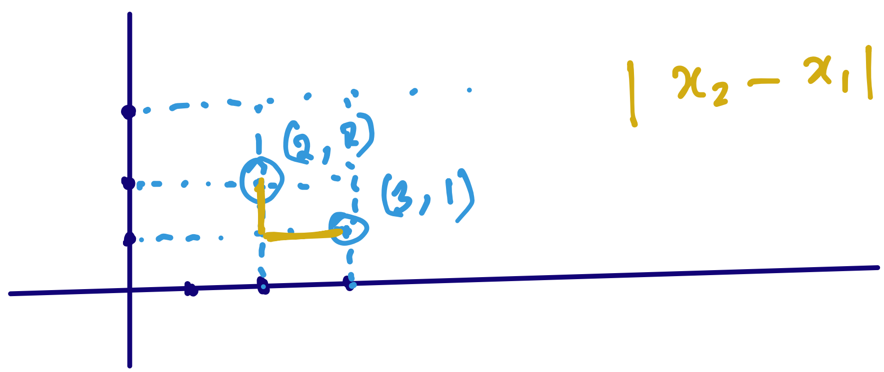
① **Euclidean distance** : Shortest path between two points.

$$\sqrt{(3-2)^2 + (1-2)^2}$$



② **Manhattan distance**: The distance by grid-shaped roads on the 2-D plane.

$$|x_2 - x_1| + |y_2 - y_1|$$



Linear Algebra: Matrix multiplication.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2\times 2} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}_{2\times 2} = \begin{bmatrix} (1\times 5 + 2\times 7) & (1\times 6 + 2\times 8) \\ (3\times 5 + 4\times 7) & (3\times 6 + 4\times 8) \end{bmatrix}$$

$$= \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}.$$

Dimensions of a matrix: number of rows × number of columns.

* To do matrix multiplication, we have to check if: $n = a$. ← this holds.

$$A_{m \times n} \qquad B_{a \times b}$$

Matrix multiplication is possible only if

$\bar{x}$ $\begin{bmatrix} 1 \\ 2 \end{bmatrix}_{2 \times 1}$ $\bar{y}$ $\begin{bmatrix} 3 \\ 4 \end{bmatrix}_{2 \times 1}$

If $\bar{x}$ is a vector, its transpose is written as $\bar{x}^T$

Transpose.

$$\begin{bmatrix} 1 & 2 \end{bmatrix}_{1 \times 2} \begin{bmatrix} 3 \\ 4 \end{bmatrix}_{2 \times 1} = \begin{bmatrix} 1 \times 3 + 2 \times 4. \end{bmatrix}$$

$$3 + 8$$

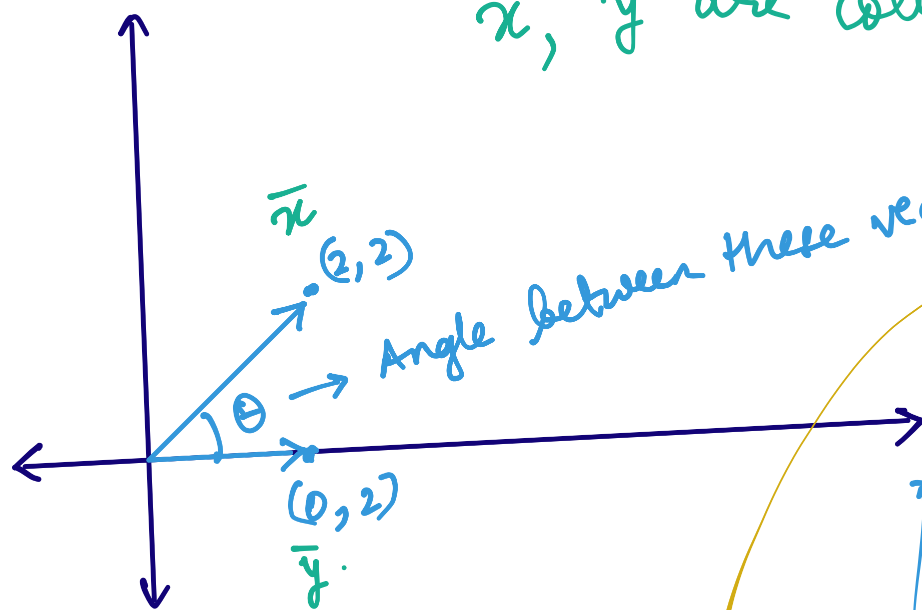
$$= 11$$

This operation is called the dot product.

$\bar{x}^T \bar{y} \rightarrow$ notation for dot-product.

Linear Algebra: Angle between two vectors.

$\bar{x}$, $\bar{y}$ are column vectors.



$\theta \rightarrow$ Angle between these vectors.

This formula works for all dimensions.

$$\cos(\theta) = \frac{\bar{x}^T \bar{y}}{\|\bar{x}\| \cdot \|\bar{y}\|}.$$

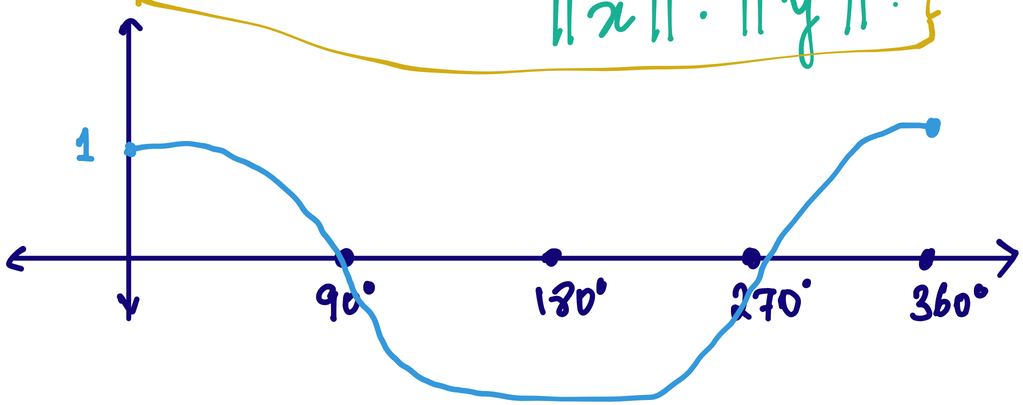
$0^\circ \le \theta \le 90^\circ$

or

$270^\circ \le \theta \le 360^\circ$

cos(θ) will be +ve

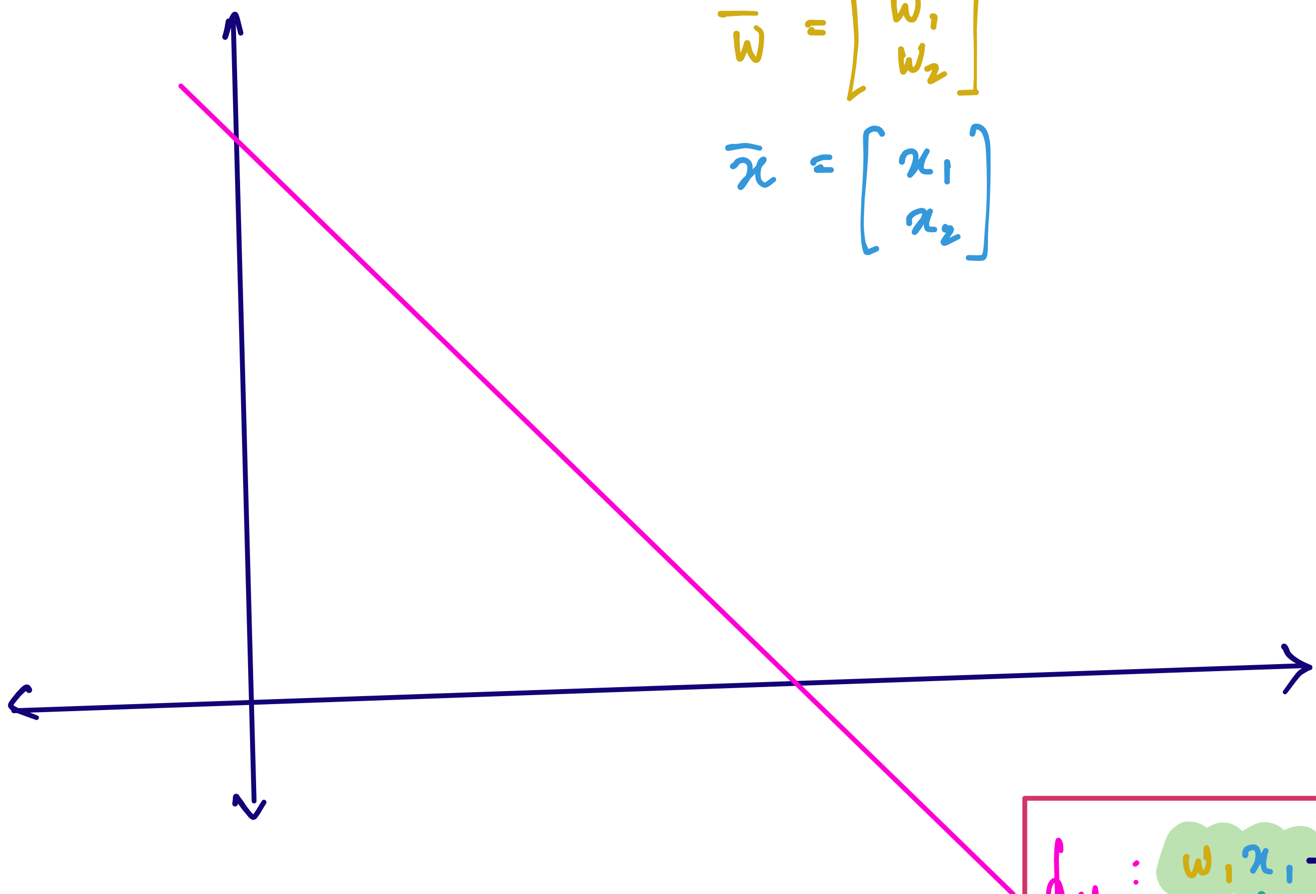
$90^\circ \le \theta \le 270^\circ$
cos(θ) will be -ve.



Crucial connection between Co-ordinate geometry and linear algebra.

$$\bar{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$



$$\bar{w}^T \bar{x} + w_0 = 0.$$

Interpret this as follows.

$$\boxed{l_1 : w_1 x_1 + w_2 x_2 + w_0 = 0.}$$

$$\bar{w}^T \bar{x} = \bar{x}^T \bar{w}$$

Can we do this for the d-dim hyperplane?

(d D) : $w_1 x_1 + w_2 x_2 + \dots + w_d x_d + w_0 = 0.$

$$\bar{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_d \end{bmatrix} \qquad \bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_d \end{bmatrix}$$

→ weight vector.

→ Feature vector.

$\bar{w}^T \bar{x} + w_0 = 0.$

Bias term.

$\bar{a}$

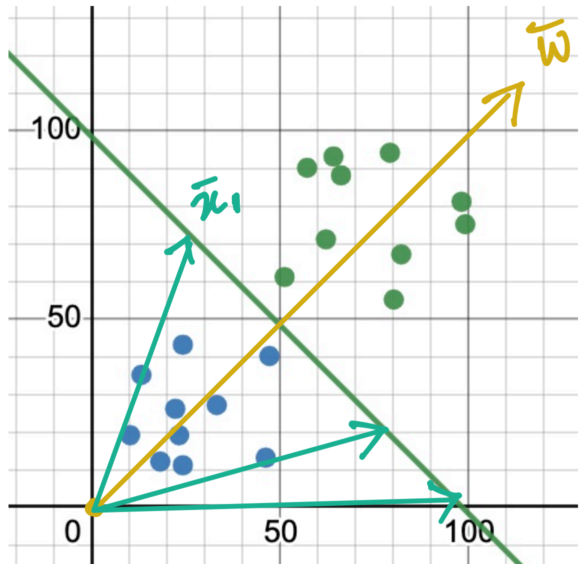
$\bar{w}^T \bar{a} - 98 = 0$

close connection with halfspaces.

$\bar{w}^T \bar{a} = 98.$

$\bar{w}$

The weight vector is $\perp^r$ to the line!!!



$\bar{a} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$x_1 + x_2 - 98 = 0.$

$x_1 + x_2 = 98.$

$\bar{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\quad \bar{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

Cheat Sheet:

① $\bar{x}, \bar{y}$ are vectors.

If their dimension is d, then:

(i) $\bar{x}, \bar{y} \in \mathbb{R}^d$.

(ii) $\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix}$, $\bar{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{bmatrix}$

② $||\bar{x}|| = \sqrt{x_1^2 + x_2^2 + \cdots + x_d^2}$

③ $\bar{x}^T \bar{y} \rightarrow$ Dot product.

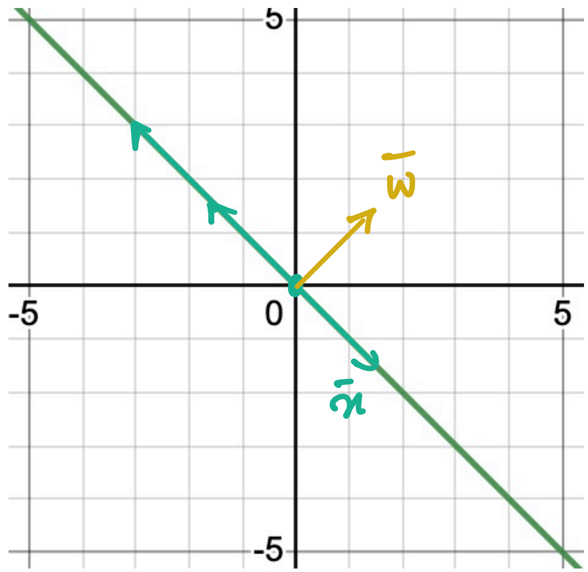
$$= \sum_{i=1}^{d} x_i \cdot y_i .$$

④ If angle between $\bar{x}$ & $\bar{y}$ is θ,

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

Let's say, we don't know $\theta \rightarrow$ angle between $\bar{x}$ and $\bar{w}$. We want to find θ.

$$\boxed{\bar{w}^T \bar{x} = 0.} \quad (1)$$



Substitute (1) in (2), we have

$$\cos(\theta) = \frac{0}{\|\bar{x}\| \cdot \|\bar{w}\|}$$

$$\cos(\theta) = 0.$$

$$\therefore \theta = 90^\circ.$$

$$\boxed{\cos(\theta) = \frac{\bar{w}^T \bar{x}}{\|\bar{x}\| \|\bar{w}\|}} \quad (2)$$