Linear Algebra - 2

17 March 2025 10:48

→ Summary of Hyperplanes (with equation)

If we have two features & we call them x & y then the classifier/ boundary/ seperator/ hyperplane will be of the form:

$$ax + by + c = 0$$

Features

constants.

→ What if we have 3 features?

Names of the features = x, y, z

" " " constants = a, b, c & d

eq^n : $ax + by + cz + d = 0$

→ This will become complicated for higher no. of features

∴ Let's call our feature $x_1, x_2, x_3, ----$ &

let's call our constants $w_1, w_2, w_3, ----$. (Why w? because these constants are also called 'weights')

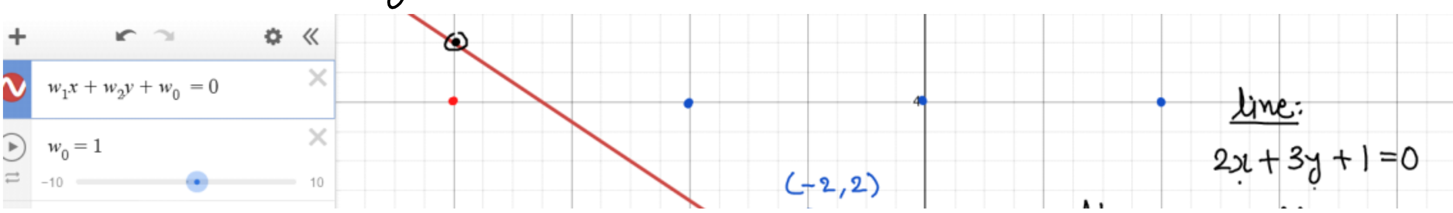
∴ The hyperplane with 2 features : $w_1x_1 + w_2x_2 + w_0 = 0$

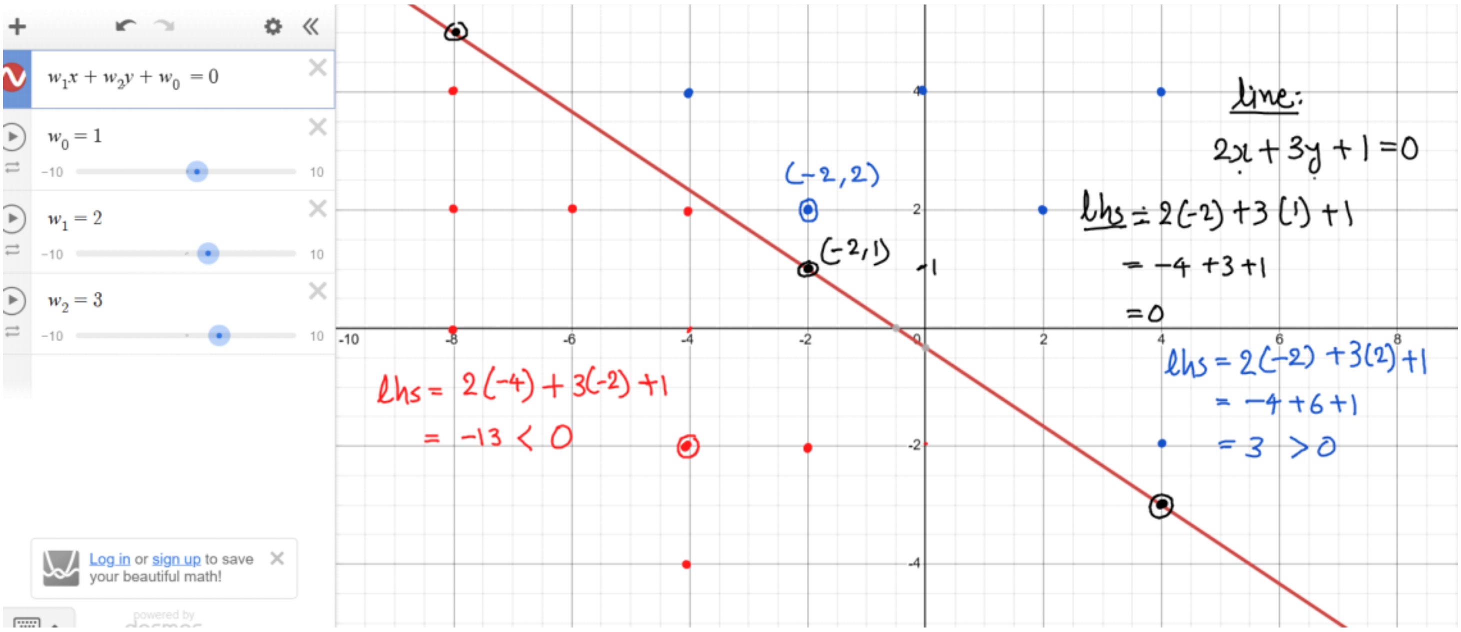
" " " 3 " : $w_1x_1 + w_2x_2 + w_3x_3 + w_0 = 0$

⋮

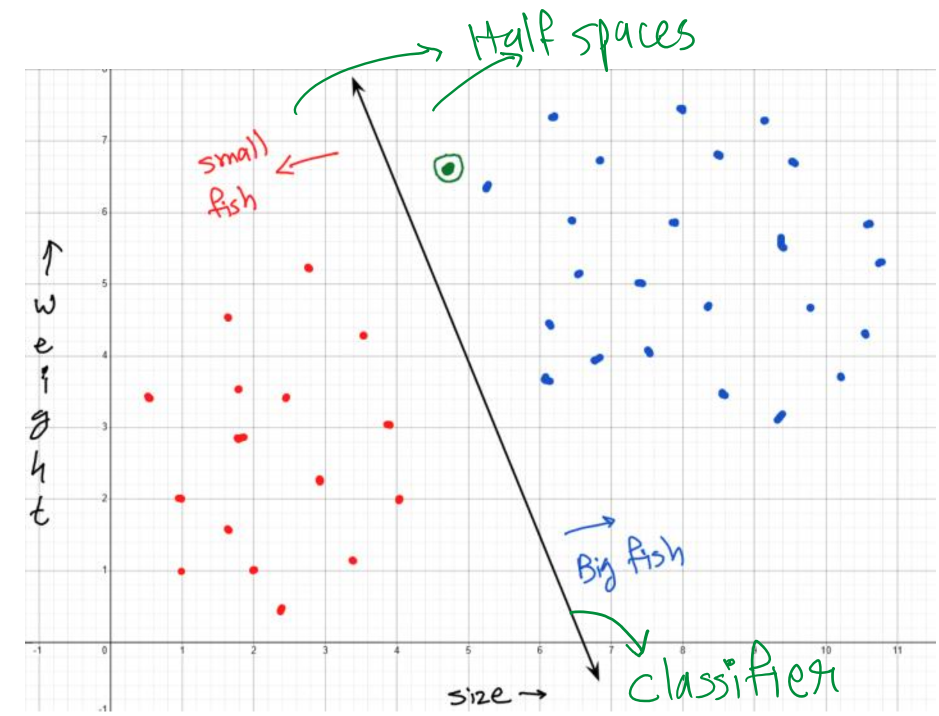
∴ The hyperplane with n features : $w_1x_1 + w_2x_2 + \_ \_ . + w_nx_n + w_0 = 0$

→ An interesting fact about straight line-





→ Coming back to our fish-sorting problem –

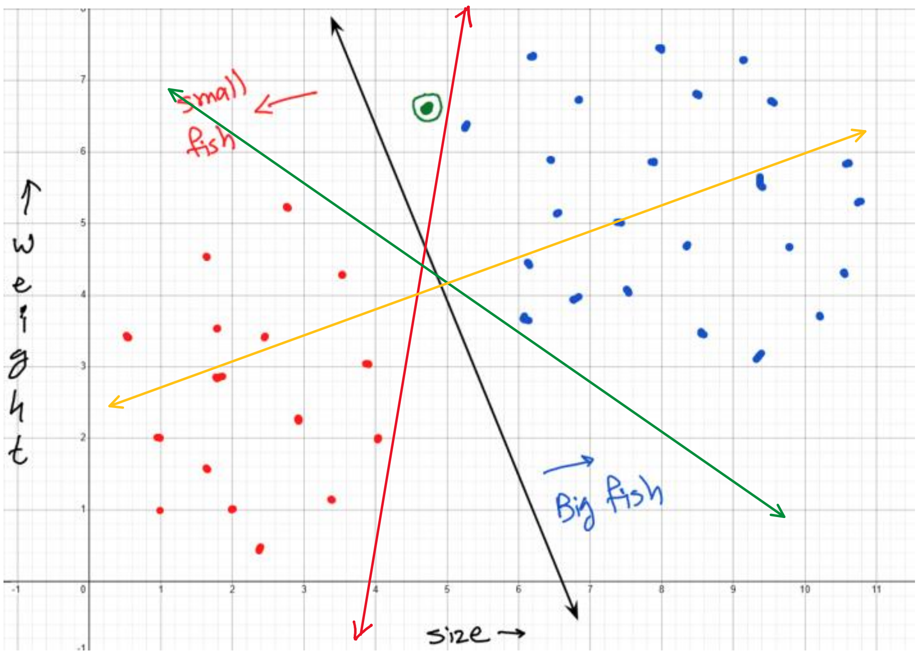


→ If we find out this classifier's equation ($w_1 x_1 + w_2 x_2 + w_0 = 0$ in the form of: $2x_1 + 3x_2 + 1 = 0$) just by putting values of x_1 & x_2 (in our case - size & weight) in that equation, we can determine the type of the fish (if $lhs > 0$, big fish & if $lhs < 0$, small fish or otherwise)

→ Therefore, the entire ML is focussed on finding these weights $w_1, w_2, \ldots, w_n$ & w_0

☆ Which of the followings is a better classifier?

✯ Which of the followings is a better classifier?



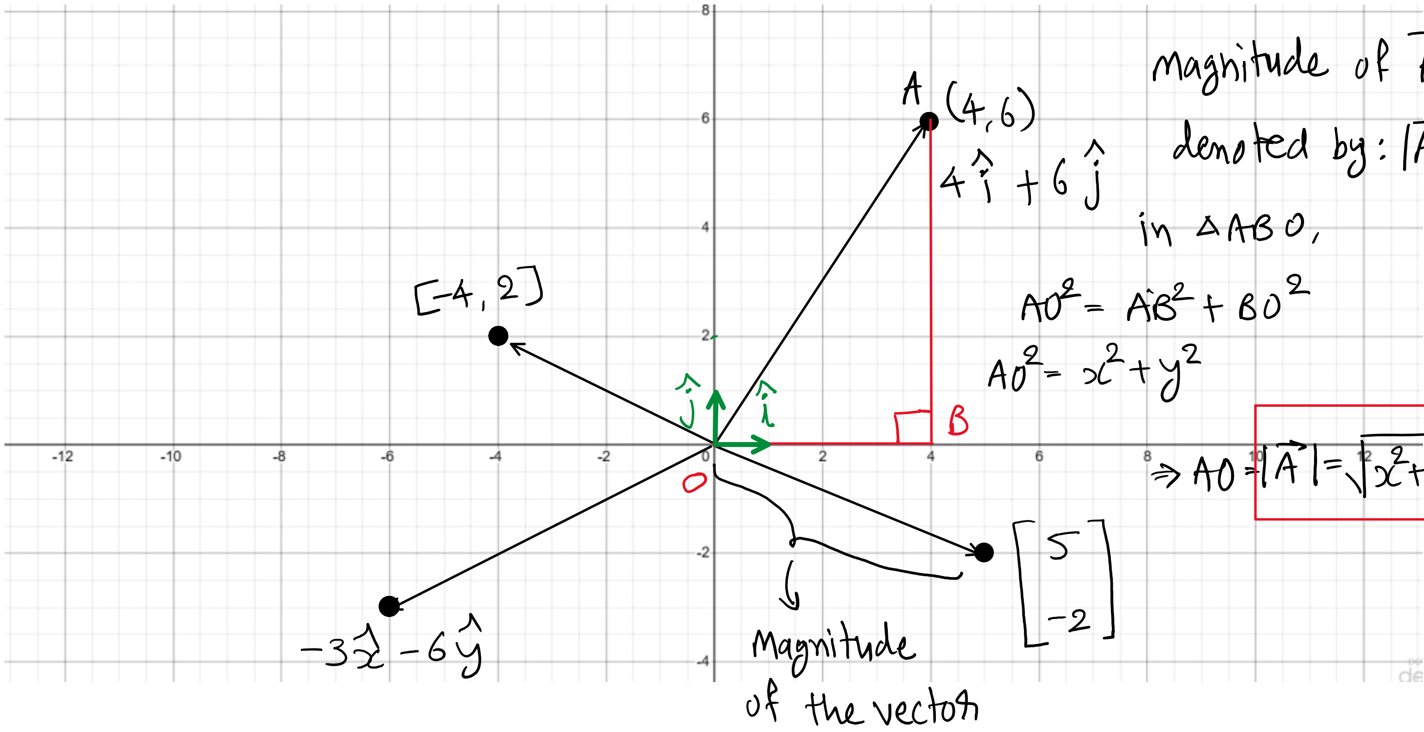
Yellow is not a good classifier as it has already misclassified a lot of points.

Green, Red & Black lines have not misclassified any points yet. But because Red & Green lines are very close to some points, it is possible that they will misclassify new datapoints in the future. The probability of such misclassification for the black line is less because it is more distant from both Red & Blue points.

∴ The key is to maximize the distance of points from our classifier.

∴ We need to learn how to find distance of a point from the given line. We also represent points as "vectors" hence, we need to learn about vectors first.

✯ Vectors - Any point can be considered as a vector. (In ML context)



magnitude of $\vec{A}$ denoted by: $|\vec{A}|$

in $\triangle ABO$,

$AO^2 = AB^2 + BO^2$

$AO^2 = x^2 + y^2$

$\Rightarrow AO = \boxed{|\vec{A}| = \sqrt{x^2+y^2}}$

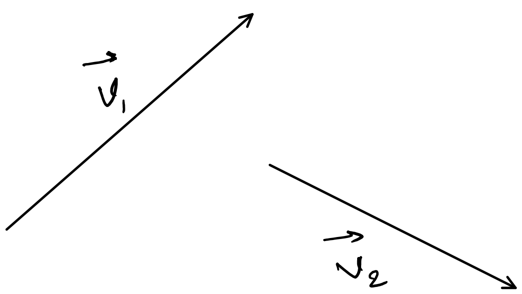
$\hat{i}$ & $\hat{j}$ are called **unit vectors** in the directions of x & y axis respectively.

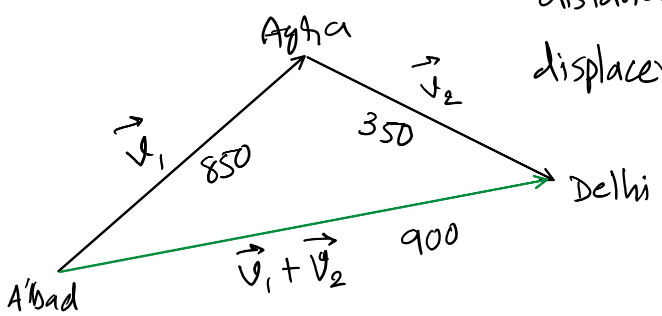
Unit vector is any vector that has magnitude of 1 unit.

→ Addition of vectors :

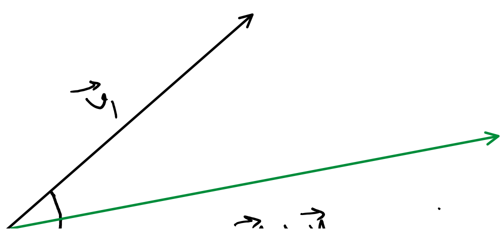
distance = 1200

displacement = 900

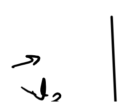


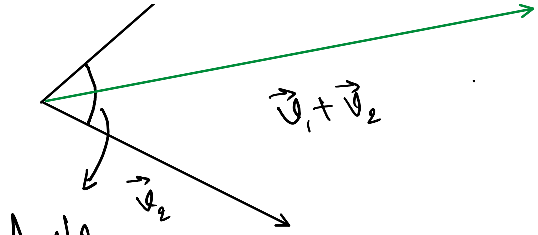


→ An alternate way of vector addition:

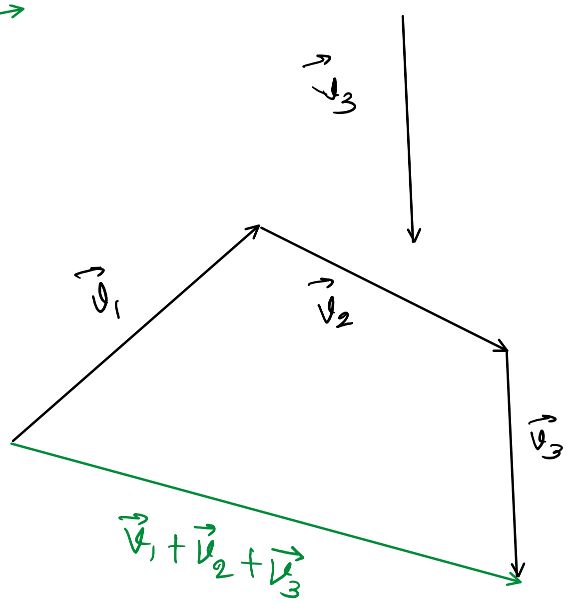


→ Adding multiple vectors:

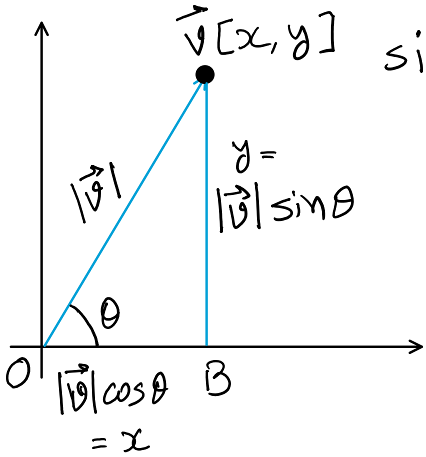




Angle between two vectors



→ Vectors in the perspective of trigonometry



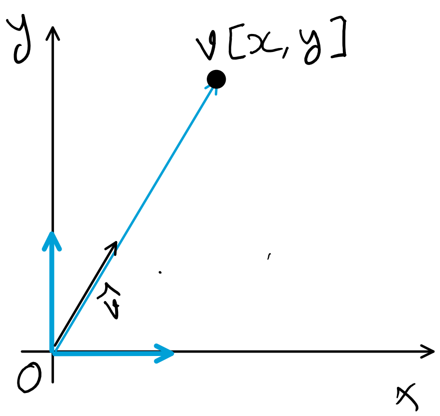
$\sin\theta = \dfrac{y}{|\vec{v}|}$ $\qquad \cos\theta = \dfrac{x}{|\vec{v}|}$ $\qquad \tan\theta = \dfrac{y}{x}$

$\boxed{y = |\vec{v}| \sin\theta}$ $\qquad \boxed{x = |\vec{v}| \cos\theta}$ $\qquad \boxed{\therefore y = x\tan\theta}$

$|\vec{v}| = \sqrt{x^2 + y^2} \Rightarrow |\vec{v}| = \sqrt{x^2 + x^2\tan^2\theta} = \sqrt{x^2(1+\tan^2\theta)}$

$$\boxed{\therefore |\vec{v}| = x\sqrt{1+\tan^2\theta}}$$

→ A deeper dive into unit vectors: Suppose we want to find unit vector in the direction of $\vec{v}$ that is $\hat{v}$.



$\therefore |\hat{v}| = 1$ (As it is a unit vector)

Will it be [1, 1]? No.

Because P[1, 1] has magnitude ≠ 1:

$|\vec{P}| = \sqrt{x^2 + y^2} = \sqrt{1^2 + 1^2} = \sqrt{2} \neq 1$

$\therefore \vec{P}[1, 1]$ is NOT a unit vector.

... direction of any vector $\vec{v}$ is

Unit vector in the direction of any vector $\vec{v}$ is that vector divided by its magnitude.

$$\boxed{\hat{v} = \frac{\vec{v}}{|\vec{v}|}}$$

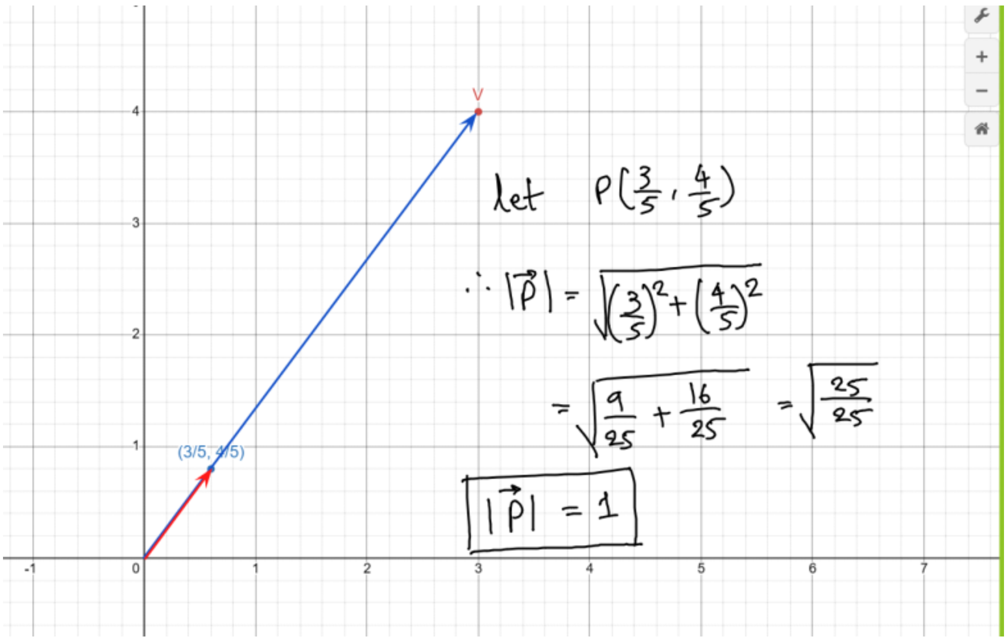
Suppose $\vec{v}\,[3, 4]$ is our vector

$$\therefore \hat{v} = \frac{\vec{v}}{|\vec{v}|} = \frac{[3,\ 4]}{\sqrt{3^2 + 4^2}} = \frac{[3,\ 4]}{\sqrt{9+16}} = \frac{[3,\ 4]}{5}$$

$$\hat{v} = \left[\frac{3}{5}, \frac{4}{5}\right]$$

If we plot this point $\left[\frac{3}{5}, \frac{4}{5}\right]$, it will be a vector in the direction of $\vec{v}\,[3, 4]$ & it will have magnitude $= 1$

$\therefore$ It is unit vector in the direction of $\vec{v}$.



let $P\left(\frac{3}{5}, \frac{4}{5}\right)$

$$\therefore |\vec{P}| = \sqrt{\left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2}$$

$$= \sqrt{\frac{9}{25} + \frac{16}{25}} = \sqrt{\frac{25}{25}}$$

$$\boxed{|\vec{P}| = 1}$$

In general, for $\vec{v}\,[x, y]$, $|\vec{v}| = \sqrt{x^2 + y^2}$ & $\vec{v} = x\hat{i} + y\hat{j}$

$\therefore \hat{v} = \frac{\vec{v}}{\ldots}$ will turn to:

$\therefore \quad \hat{v} = \dfrac{\vec{v}}{|\vec{v}|}$ will turn to:

$$\hat{v} = \frac{x\hat{i} + y\hat{j}}{\sqrt{x^2+y^2}} \Rightarrow \boxed{\hat{v} = \frac{x}{\sqrt{x^2+y^2}} \cdot \hat{i} + \frac{y}{\sqrt{x^2+y^2}} \cdot \hat{j}}$$

☆ 'Norm' of a vector :-

norm of a vector is nothing but its magnitude only, but there are many ways to calculate this magnitude. To understand that can we say that magnitude of a vector is essentially **distance** of that point from the origin. And, there are multiple ways to compute distance.

e.g. distance between a warehouse & dilivery location can be computed as the crow flies (for drone dilivery) or as the 'L-distance' (for traditional dilivery).

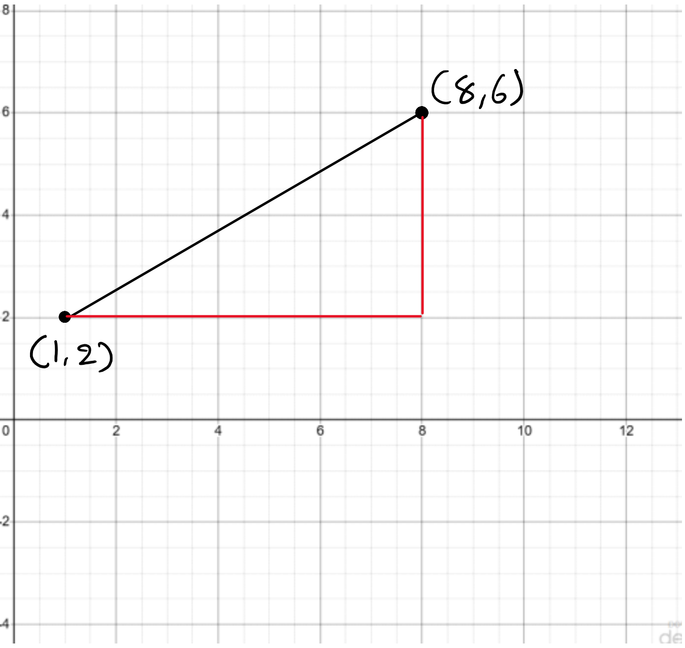
Straight line distance

= Euclidean Distance

$= \sqrt{(x_2-x_1)^2+(y_2-y_1)^2}$ OR

For magnitude of a vector, we write it as:

$$\sqrt[2]{x^2+y^2}$$



This is also known as **2-norm** and denoted by $\|\vec{v}\|$

→ 'L-distance' = Manhattan Distance

$$= (x_2-x_1)+(y_2-y_1) = \sqrt[1]{(x_2-x_1)^1+(y_2-y_1)^1}$$

For magnitude of a vector, we can re-write this as:

$\sqrt[1]{x^1+y^1}$ ← This magnitude is also called **1-norm** of the vector.

& denoted by $|\vec{v}|$.

Similarly, 3-norm $= \|\vec{v}\|_3 = \sqrt[3]{x^3+y^3}$

⋮

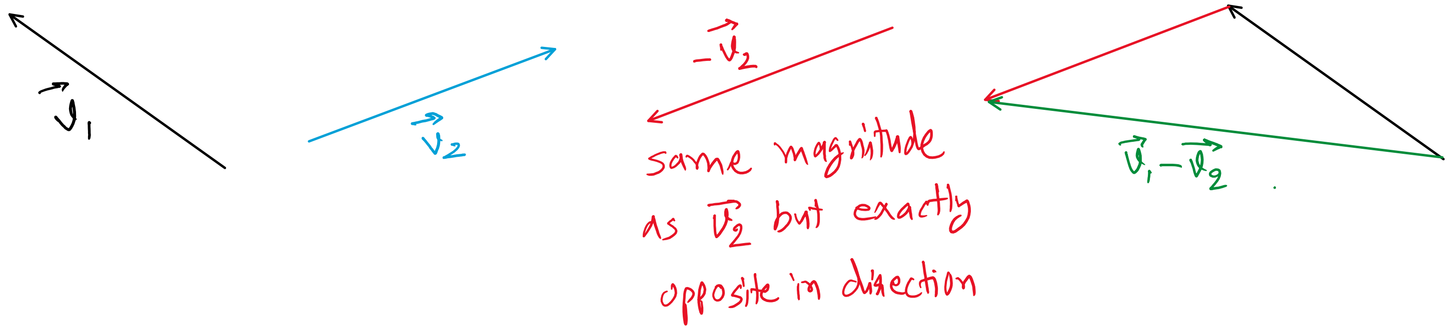
n-norm $= \|\vec{v}\|_n = \sqrt[n]{x^n+y^n}$

1-norm is also denoted as $\|\vec{v}\|_1$

☆ Subtraction of two vectors

$\vec{v}_1 - \vec{v}_2 = \vec{v}_1 + (-\vec{v}_2)$

$\vec{v}_1 + (-\vec{v}_2)$:



As per the formula: Let $\vec{v}_1[x_1, y_1]$ & $v_2[x_2, y_2]$

$$\therefore \vec{v}_1 = x_1 \hat{i} + y_1 \hat{j} \quad \& \quad \vec{v}_2 = x_2 \cdot \hat{i} + y_2 \hat{j}$$

$$\therefore \vec{v}_1 - \vec{v}_2 = x_1 \cdot \hat{i} + y_1 \hat{j} - (x_2 \cdot \hat{i} + y_2 \cdot \hat{j})$$

$$= \underline{x_1 \cdot \hat{i}} + \underline{\underline{y_1 \cdot \hat{j}}} - \underline{x_2 \cdot \hat{i}} - \underline{\underline{y_2 \cdot \hat{j}}}$$

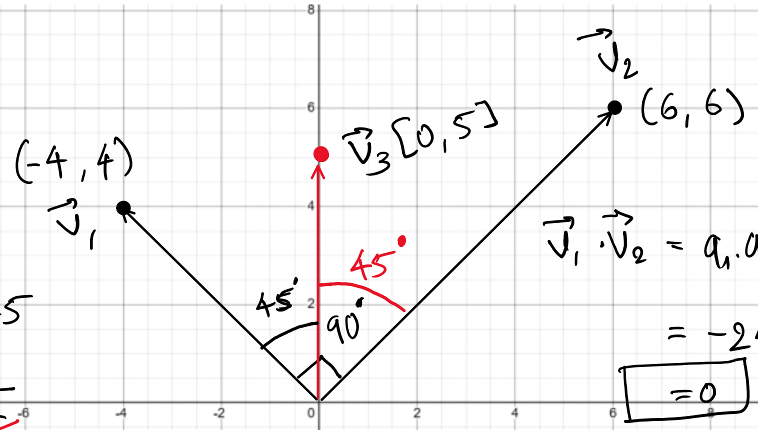
$$\boxed{\vec{v}_1 - \vec{v}_2 = (x_1 - x_2) \cdot \hat{i} + (y_1 - y_2) \cdot \hat{j}}$$

☆ Dot product of two vectors

– Multiplication of two vectors that **results in a scalar.**

Let $\vec{v}_1[a_1, b_1]$ & $\vec{v}_2[a_2, b_2]$ then first we write $\vec{v}_2$ as a column vector i.e. $\vec{v}_2 \begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$ and then we multiply the elements as this: $[\underline{a_1} \quad \underline{b_1}] \cdot \begin{bmatrix} \underline{a_2} \\ \underline{b_2} \end{bmatrix}$

$$\boxed{\therefore \vec{v}_1 \cdot \vec{v}_2 = a_1 \cdot a_2 + b_1 \cdot b_2} \quad \underline{\underline{\text{OR}}} \quad \boxed{\vec{v}_1 \cdot \vec{v}_2 = |\vec{v}_1| \cdot |\vec{v}_2| \cdot \cos\theta}$$



$\vec{v}_1 \cdot \vec{v}_3 = 0 + 20$

$\boxed{= 20}$

$|\vec{v}_1| \cdot |\vec{v}_3| \cdot \cos\theta$

$= 4\sqrt{2} \times 5 \times \cos 45$

$= 4\cancel{\sqrt{2}} \times 5 \times \frac{1}{\cancel{\sqrt{2}}}$

$\boxed{= 20}$

$\vec{v}_1 \cdot \vec{v}_2 = a_1 \cdot a_2 + b_1 \cdot b_2$

$= -24 + 24$

$\boxed{= 0}$

$|\vec{v}_1| \cdot |\vec{v}_2| \cdot \cos\theta = |\vec{v}_1||\vec{v}_2| \cos 90$

$\boxed{= 0}$

$\vec{v}_2 \cdot \vec{v}_3 = 0 + 30$

$\boxed{= 30}$

$= |\vec{v}_2||\vec{v}_3| \cos 45 = 6\cancel{\sqrt{2}} \times 5 \times \frac{1}{\cancel{\sqrt{2}}}$

$\boxed{= 30}$

$|\vec{v}_1| = \sqrt{(-4)^2 + 4^2} = \sqrt{32} = \sqrt{16 \times 2} = 4\sqrt{2}$

$|\vec{v}_2| = \sqrt{36 + 36} = \sqrt{72} = \sqrt{36 \times 2} = 6\sqrt{2}$

$|\vec{v}_3| = \sqrt{25} = 5$

→ Matrix Multiplication as **Collection of Dot Products**

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_1 \cdot x_1 + b_1 \cdot x_2 + c_1 \cdot x_3 \\ a_2 \cdot x_1 + b_2 \cdot x_2 + c_2 \cdot x_3 \\ a_3 \cdot x_1 + b_3 \cdot x_2 + c_3 \cdot x_3 \end{bmatrix}$$

$[a_1\ b_1\ c_1] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$[a_2\ b_2\ c_2] \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$[a_3\ b_3\ c_3] \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$