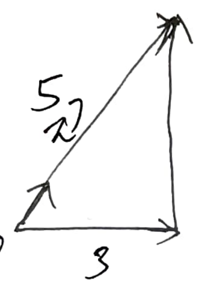
Linear Algebra

What is the unit vector of $\vec{a}$?



Let the unit vector of $\vec{a}$ be $\hat{u}$

unit vector on horizontal axis $\hat{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

vertical axis $\hat{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

So the vector $\vec{a}$ with magnitude 5 can be represented as $\vec{y} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

So how can you represent a unit vector $\hat{u}$ of $\vec{a}$

Basically $\hat{u}$'s magnitude should be 1

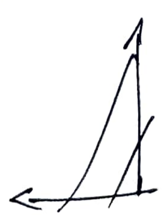
$\vec{a} = \sqrt{3^2+4^2} = \sqrt{9+16} = \sqrt{25} = 5$

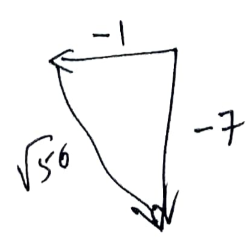
What would it take for $\vec{a}$ to be $\hat{u}$

$\Rightarrow \sqrt{(\frac{3}{5})^2+(\frac{4}{5})^2} = \sqrt{\left(\frac{9}{25}+\frac{16}{25}\right)}$

$= \sqrt{\frac{25}{25}} = \sqrt{1} = 1$

$\sqrt{4+49} = \sqrt{53}$

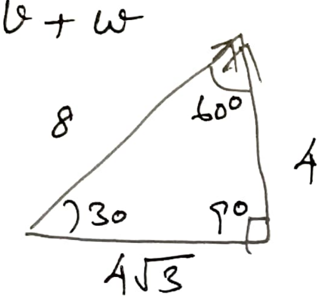




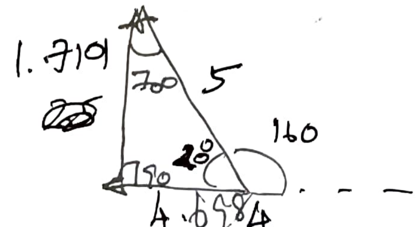
$\sqrt{(-1)^2+(-7)^2} = \sqrt{50}$

The magnitude of $\vec{v}$ is 8 and its direction angle is 30°
The magnitude of $\vec{w}$ is 5 and its direction angle is 160
find $v + w$

$v =$



$w =$



$\sin 30 = \frac{Opp}{Hyp} = \frac{opp}{8}$

$Opp = 4$

$\cos 30 = \frac{Adj}{Hyp} = \frac{Adj}{8}$

$8 \times \frac{\sqrt{3}}{2} = Adj$

$\sin 20 = \frac{Opp}{5}$

.342

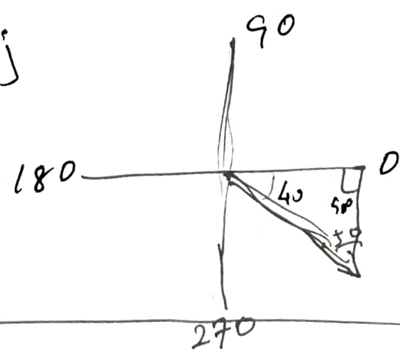
$Opp = 1.7101$

$\cos 20 = \frac{Adj}{5}$

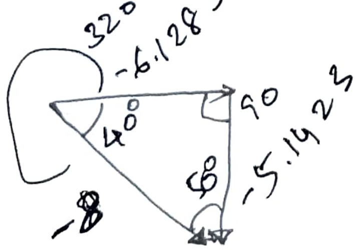
$4\sqrt{3}i + 4j$ $\quad -4.6984i + 1.7101j$

$6.9282i + 4j$

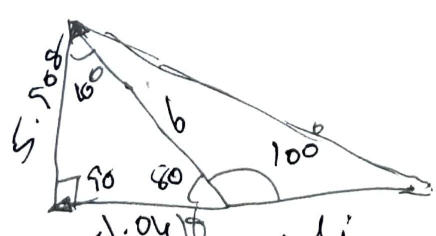
$2.2298i + 5.7101j$



$6.1283i - 5.1423j$



$\cos 40 = \frac{Adj}{-8}$



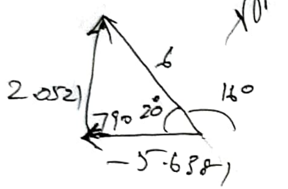
$\cos 80 = \frac{Adj}{6}$

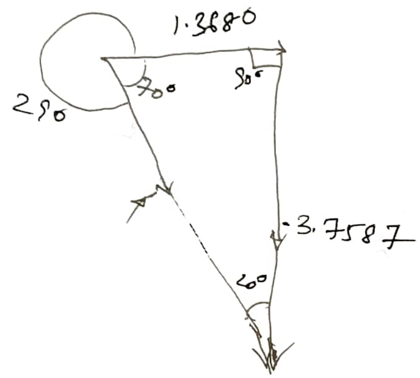
5.086

$-6.1283i - 5.1423j$

$-1.0418i + 5.908j$

$-7.1701 + 0.7657$





-4.27
-1.70

-4.9240 i + 0.8682 j

- 1.0418 i 5.9088 j

7.8784 i + 1.3891 j
5.7340 i 4.0150 j

1.3891
4.0150
5.4041

7.8784
5.7340
13.6124

7.123
11.8954

7.87

$8 . \sin 80 = 7.8784$

$8 . \cos 80 = 1.3891 i + 7.8784 j$

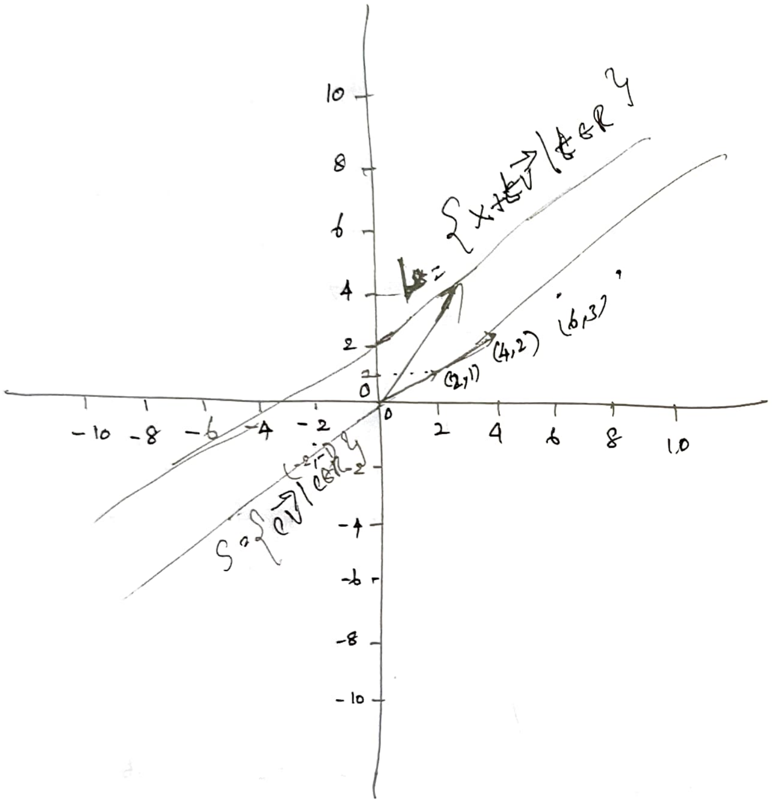
$5.7340 i - 4.0150 j$

7.1231 ~~11.9~~

3.86

$-5 \times \cos 10$	$5 \times \sin 10$	$-4.9240 i$	$+0.8682 j$
$8 \times \cos 60$	$-8 \times \sin 60$	4	-6.9282
		-0.924	-6.06

$4 \times \cos 50$	$-4 \sin 50$	$2.57 i$	$- 3.06 j$
$7 \times \cos 5$	$7 \sin 50$	$4.4995 i$	$+5.3623 j$
		7.0695	2.3023
$-7 \times \cos 80$	$7 \times \sin 80$	$-1.2155 i$	$+ 6.8936 j$
$5 \times \cos 70$	$-5 \times \sin 70$	$1.7101 i$	$- 4.6984 j$
		0.4945	2.1952



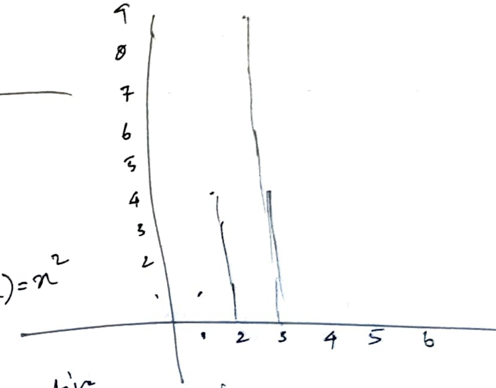
$\vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$S = \{c\vec{v} \mid c \in \mathbb{R}\}$ S is a set of all vectors out of $\vec{v}$ multiplied by a scalar c.

If $c = 2$

$2\vec{v} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$ When we multiply the $\vec{v}$ with a scalar the resulting vector is Colinear.

$y = x^2$

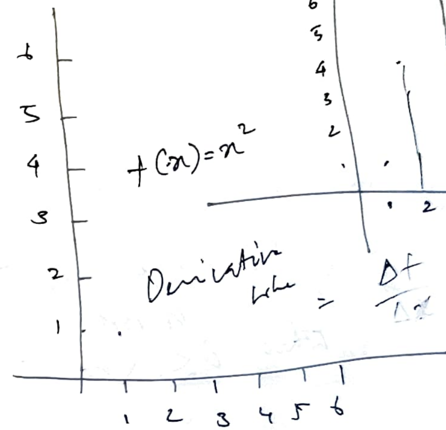


When $x = 1$ $y = 1$

$\Delta y = 1$

$\Delta x = 1$

Derivative $= 1$



Derivative when $= \frac{\Delta y}{\Delta x}$

$x = 2$

$y = 4$

$\Delta y = 3$

$\Delta x = 1$

Derivative is 3

$x = 3$

$y = 9$

$\Delta y = 6$

$\Delta x = 1$

derivat $= 6$

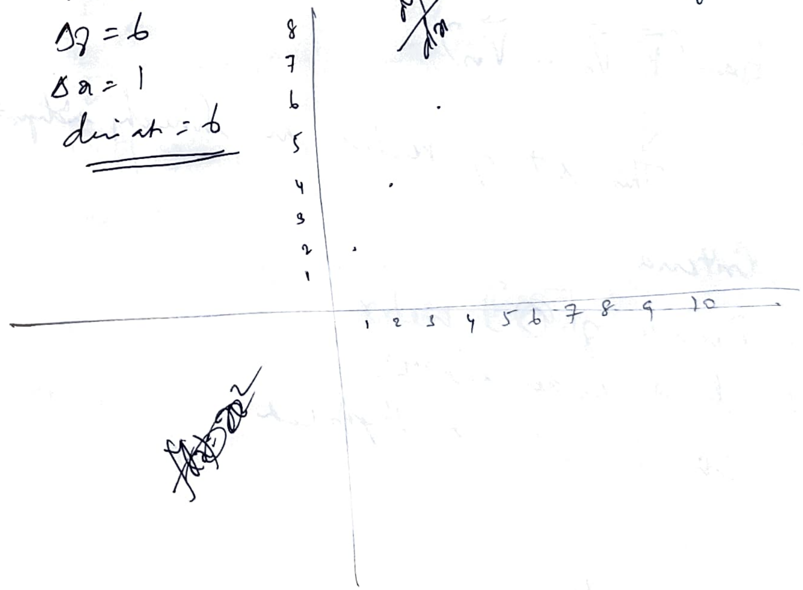
$y = x^2$

$\frac{dy}{dx} = 2x$

$y = mx$

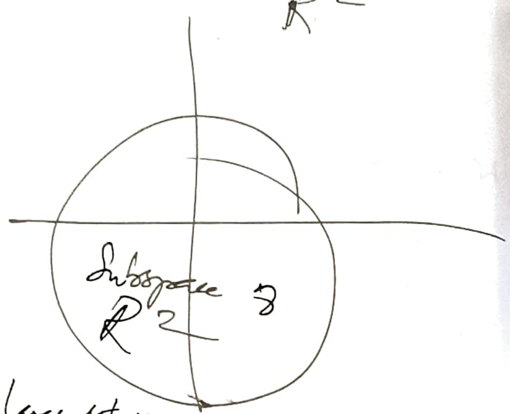
$y = 2 \cdot 1$

$y = 4$



Linear Subspaces ::

Subspace of R^n
$V \leftarrow$ Subset of R^n



$R^n = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right.$ ... large set of vectors $\}$

when $i \le 1 \le n$

Span of a vector ::

$V = \text{Span}(\vec{V_1}, \vec{V_2} \dots \vec{V_n})$

$\downarrow$ Subspace

$\downarrow$ This set of vectors are linearly independent

Subspace :: Criteria
i) Presence of $[0]$ vector
ii) closed under addition
iii) closed under multiplication

$\log_2 8 = 3$

$\log_y x = a \quad y^a = x$

which is $2^3 = 8$
negative not allow
$\log_{-2} 8 = x$
$-2^x = 8$

$\log_{-2} x = 1/2$
$\Rightarrow (-2)^{1/2} = x$
$= \sqrt{-2} \Rightarrow$ Square root of negative number is not possible

What is exponential "e"
It is the maximum possible growth continuously compounding a 100% growth over 1 time period

e

1 year: $1 \rightarrow 1 \} 2$
100% growth

Semi Annual: $1 \overset{0.5}{\_\_\_} 1 \overset{.75}{\rightarrow} 1.5$, 2.25

Quarter: $1 \overset{0.25}{-} 1 \overset{0.3125}{\rightarrow} 1.25 \rightarrow$

1.25

$1.25 \times \frac{25}{100}$

$1.25 + .3125 = 1.5625$

What is Cauchy Schwartz inequality?

$\vec{V} \cdot \vec{w} = \vec{w} \cdot \vec{V}$ commutation

$\vec{v} = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} \quad \vec{w} = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{bmatrix}$

$|\vec{x} \cdot \vec{y}| \le ||x|| \, ||y||$

Cauchy Swartz inequality ::
The dot product of two vectors is less than or equal to product of length of the vectors

$||x|| \, ||y|| \ge |\vec{x} \cdot \vec{y}|$

$\vec{x} = c\vec{y} \iff |\vec{x} \cdot \vec{y}| = ||\vec{x}|| \, ||\vec{y}||$

$$\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

Length or Magnitude

$$\|v\| = \sqrt{1^2 + 2^2 + 3^2 + 4^2}$$

$$= \sqrt{1 + 4 + 9 + 16}$$

$$= \sqrt{30}$$

Length of $\|v\| = \sqrt{30}$

$$\|v\|^2 = \vec{v} \cdot \vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 9 \\ 16 \end{bmatrix}$$

This is dot product

Triangle inequality

$$|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| . \|\vec{y}\|$$

$$\vec{x} = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix} \quad y = \begin{bmatrix} -1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$$=$$

$$\left\| \begin{bmatrix} -1 \\ 4 \\ -9 \\ 16 \end{bmatrix} \right\| = |\vec{x} \cdot \vec{y}| \quad \begin{bmatrix} 1 \\ 4 \\ 9 \\ 16 \end{bmatrix}$$ when the length is

$$\|\vec{x}\| = \sqrt{1^2 + 2^2 + (-3)^2 + 4^2} = \sqrt{1+4+9+16} = \sqrt{30}$$

$$\|\vec{y}\| = \sqrt{(-1)^2 + 2^2 + 3^2 + 4^2} = \sqrt{30}$$

$$\|\vec{x}\| . \|\vec{y}\| = 30 \qquad \sqrt{30} \times \sqrt{30} = 30$$

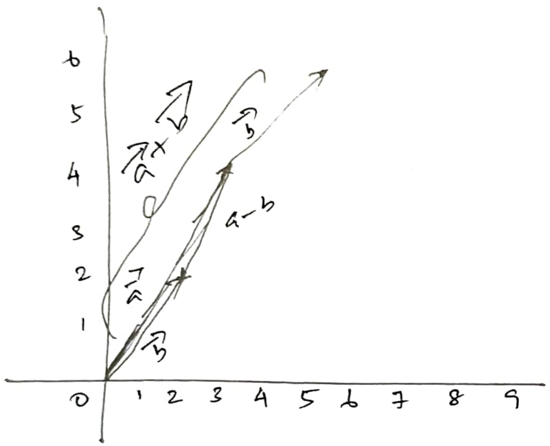
Defining the angle between vectors

$||\vec{a}||$ = length

$\mathbb{R}^2$

$\vec{a} \to$ Vector 1 = (2, 2)

$\vec{b}$ Vector 2 = (3, 4)

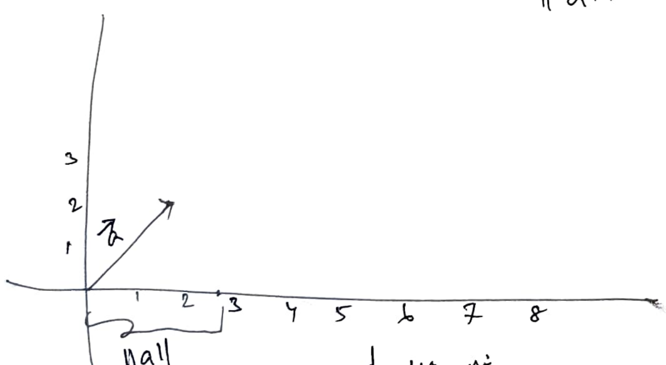


$\vec{a}+\vec{b} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$

$\vec{a}-\vec{b} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$

$\vec{a}\cdot\vec{b} = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$

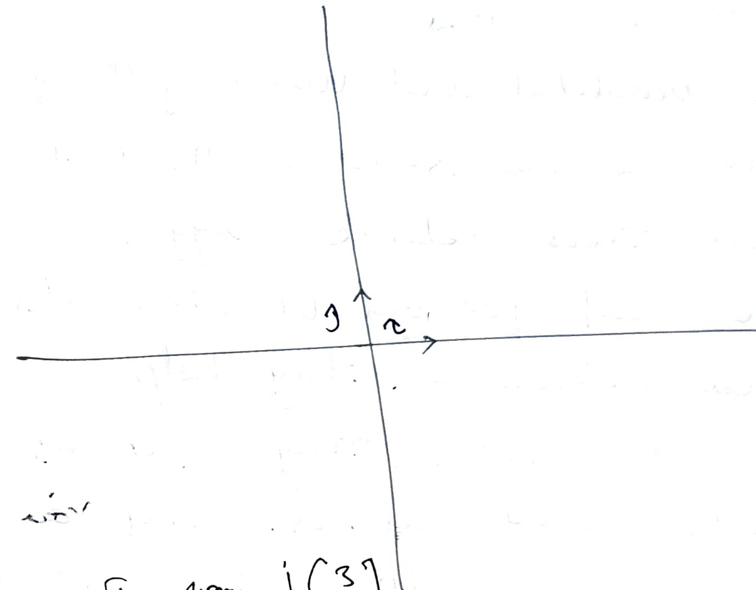
$||a|| = \sqrt{2^2+2^2}$
$= \sqrt{8}$
$= \sqrt{4\times2}$
$= 2\sqrt{2}$
$= 2.8284$



When $\vec{a}$ is straightened up in 1 dimension it shows the length of $\vec{a} = ||a||$

3 BlueBrown :: Essence of linear algebra

Linear Transformations



Transform j $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$

i $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$

fy 16 Q2

fy 17

fy 14 Q1
Q2
Q3
Q4
17 Q1
Q2
Q3
Q4

1
2
3
4
5
6
7

find the
{ hand 14   how am }
{ Im   velocity m a }
{ in }
{ }

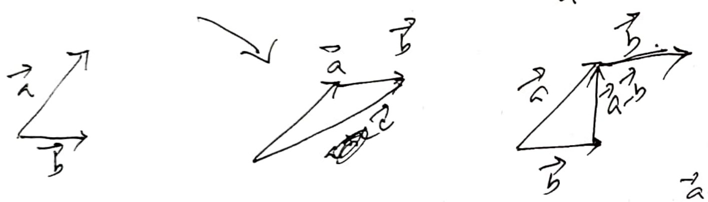
Theres a kind angel [illegible] to help just because you worked hard

$\vec{x} = \begin{bmatrix}1\\2\\3\end{bmatrix}$ $\quad \vec{x}\cdot\vec{x} = \begin{bmatrix}1\\2\\3\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{cases}1+4+9\\ =14\end{cases}$

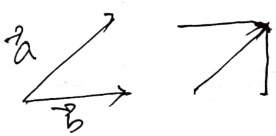
$\|\vec{x}\|^2 = \left(\sqrt{1^2+2^2+3^2}\right)^2 = \left(\sqrt{1+4+9}\right)^2 = 14$

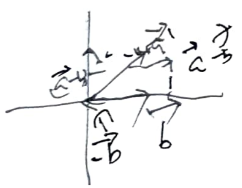
$\Rightarrow \vec{x}\cdot\vec{x} = \|x\|^2$

$\vec{a}+\vec{b}$ $\qquad\qquad\qquad\qquad \vec{a}-\vec{b}$



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





$\|\vec{a}\|^2 + \|\vec{b}\|^2 + 2\|\vec{a}\,\vec{b}\cos\theta\|$

$(\vec{a}\cdot\vec{a})$

$\vec{a}\cdot\vec{b} = \|a\|\|b\|\cos\theta$

$(\vec{a}+\vec{b})^2$ $\qquad = \|a\|^2 + \|b\|^2 + 2\|\vec{a}\,\vec{b}\|\cos$

$\cancel{\vec{a}^2}$ $(\vec{a}+\vec{b})(\vec{a}+\vec{b})$

$\|a\|^2 + \|b\|^2 + 2\vec{a}\vec{b}$

$\Rightarrow \vec{a}\cdot\vec{b} = \|\vec{a}\,b\|\cos\theta$

Goals

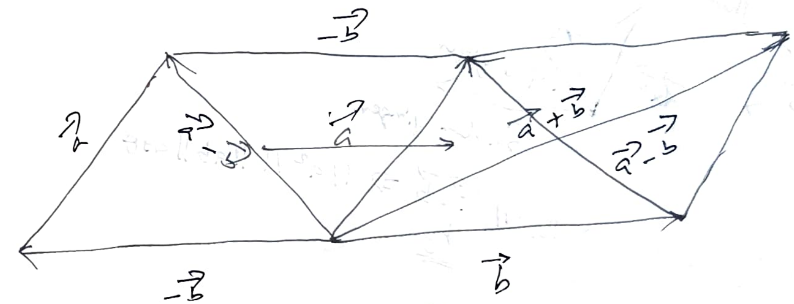
Linear Algebra

Calculus $\qquad$ Machine Learning

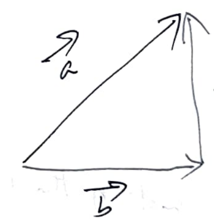
Statistics

Equation of a plane in $\mathbb{R}^3$:

Parallelogram Law :



$\Rightarrow$

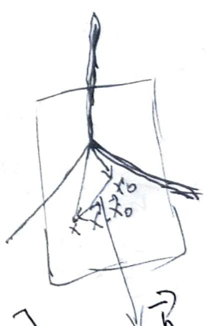


$\vec{a}-\vec{b}$ from the

Equation of a plane :

$Ax+By+Cz = D$

$\vec{n}\cdot(\vec{x}-\vec{x}_0) = 0$



$\vec{n} = \begin{bmatrix}n_1\\n_2\\n_3\end{bmatrix}$

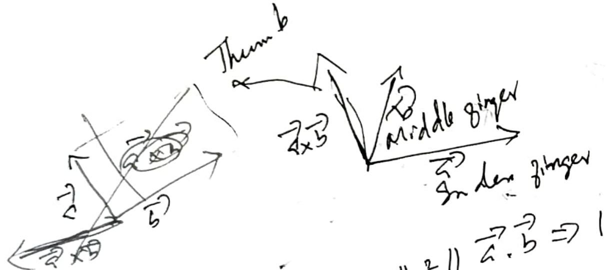
$\vec{x}_0 = \begin{bmatrix}x_0\\y_0\\z_0\end{bmatrix}$

$\vec{x} = \begin{bmatrix}x\\y\\z\end{bmatrix}$

Cross Product::

Cross product is limited to $\mathbb{R}^3$ whereas
dot product is applicable to $\mathbb{R}^n$

$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ $\vec{a} \times \vec{b} = \begin{bmatrix} (a_2 b_3 - a_3 b_2) \\ (a_3 b_1 - a_1 b_3) \\ (a_1 b_2 - a_2 b_1) \end{bmatrix}$



$(\vec{a} \times \vec{b}) \cdot \vec{a} = 0 \Rightarrow \|a^2\| \, \vec{a} \cdot \vec{b} \Rightarrow \|a^2\| \, \|ab\| \cos\theta$ $\downarrow 0$

$(\vec{a} \times \vec{b}) \cdot \vec{b} = 0$

$\vec{a} \cdot \vec{b} = 0$

Proof:: Cross product is the product of the vectors and an ax

$\Rightarrow \vec{a} \times \vec{b} = \|\vec{a}\| \|\vec{b}\| \sin\theta$

$\|\vec{a} \times \vec{b}\|^2 = (\vec{a} \times \vec{b}) \cdot (\vec{a} \times \vec{b}) = (a_2 b_3 - a_3 b_2)^2 + (a_3 b_1 - a_1 b_3)^2 + (a_1 b_2 - a_2 b_1)^2 \ldots$

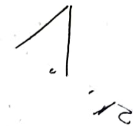
Dot and Cross

$\vec{a} \cdot \vec{b} = \|\vec{a}\| \|$

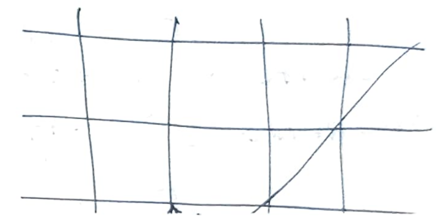
$\cos\theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|}$

$\vec{a} \cdot \vec{b} = \|\vec{a}\| \|\vec{b}\| \cos\theta$

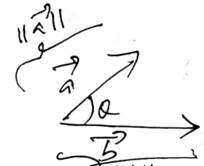
$\|\vec{a}\| \|\vec{b}\|$



$\theta = \cos^{-1} \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|}$



$\cos\theta = \frac{\text{adjacent}}{\text{hyp}}$

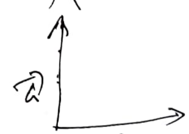
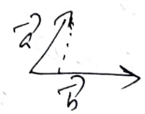


$\|a\| \|b\| \cos\theta$

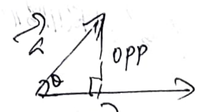
$\|b\| \|a\| \cos\theta \to \frac{\|b\| \|a\| \|b\|}{\|a\|}$

$\to \|b\|$ adjacent

You are multiplying vector $\vec{b}$ with the amount of $\vec{a}$ going alongside $\vec{b}$.

$\vec{a} \cdot \vec{b} = 0$



$\sin\theta = \frac{\text{opp}}{\|\vec{a}\|}$

$\|\vec{a}\| \sin\theta = \text{opp}$

$\|\vec{a} \times \vec{b}\| = \|\vec{a}\| \|\vec{b}\| \sin\theta$
$= \|\vec{b}\| \|\vec{a}\| \sin\theta$

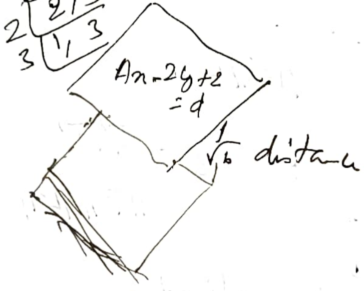
Cross Product:

Cross product is limited to $\mathbb{R}^3$

dot product is applicable

$1\ \ 4,5$
$2\ \ 2,3$
$3\ \ 1,3$

$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ $\vec{a},$



3152
4589
2536
10,277

493
65
73
74
176

3) 7596 (2532
75
9
9

11,158

$\frac{x-1}{2} = \frac{y-2}{3} = \frac{z-3}{4} = 6$

$6x-6-4y-8-3z-9 = 12$

$6x-4y-3z = 35$

$\frac{x-2}{3} = \frac{y-3}{4} = \frac{z-4}{5}$

~~20x-40~~

$60(x-2)$

(x,y,z) (x_0,y_0,z_0)

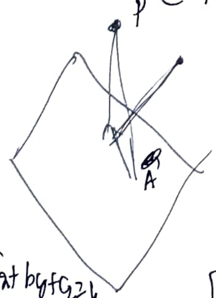
$x-x_0, y-y_0, z-z_0$ gives a line

$\frac{x-1}{2} = \frac{y-2}{3} = \frac{z-3}{4}$

$\frac{x-2}{3} = \frac{y-3}{4} = \frac{z-4}{5}$

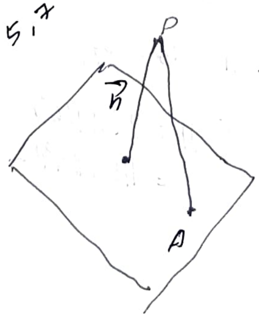
$\vec{a}\cdot\vec{b} = \|\vec{a}\|\|\vec{b}\|\cos$

$P(x_p, y_p, z_p)$



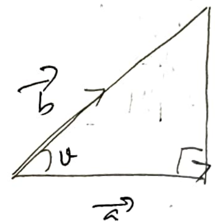
$\vec{PA}\cdot\vec{n}$

3,5,7



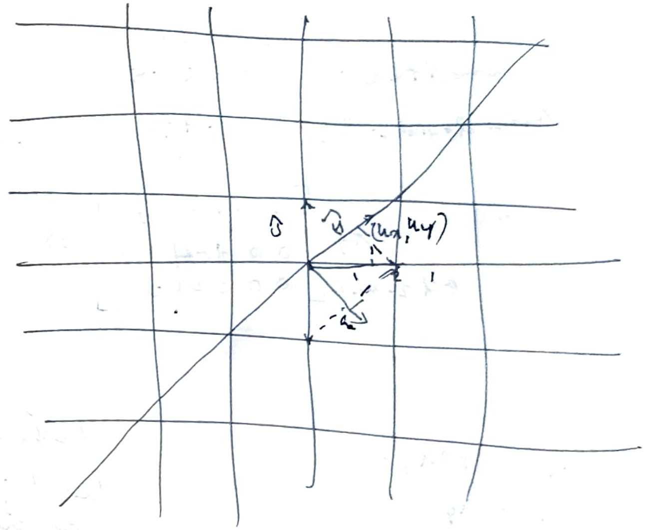
$ax+by+cz=k$

$n = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$



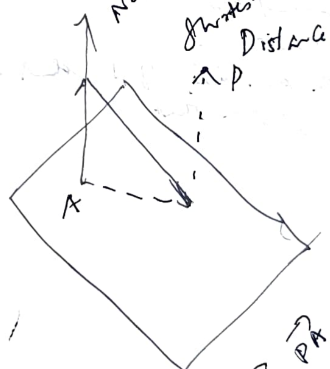
$\vec{a}\cdot\vec{b} = \|\vec{a}\|\|\vec{b}\|\cos\theta$

$\|\vec{a}\|\|\vec{b}\|$



Normal vector

Shortest Distance between a point and a plane



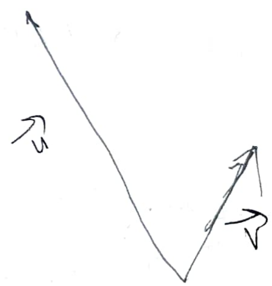
Equation of the plane

$ax+by+cz=d$

$n = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$

$\vec{n}\cdot\vec{PA}$

$= \|\vec{n}\|\|\vec{PA}\|\cos\theta$



$x_1 + 2x_2 + x_3 + x_4 = 7$
$x_1 + 2x_2 + 2x_3 - x_4 = 12$
$2x_1 + 4x_2 \quad + 6x_4 = 4$

Augmented Matrix

$\Rightarrow A = \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 7 \\ 1 & 2 & 2 & -1 & 12 \\ 2 & 4 & 0 & 6 & 4 \end{array}\right]$

Reduced row echelon form

minus first row — minus 2 × first row

$\left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 7 \\ 0 & 0 & -1 & 2 & -5 \\ 0 & 0 & -2 & 4 & -10 \end{array}\right]$

First row − second row

$= \begin{array}{l} \\ \times -1 \\ \text{Add } 2 \times \text{second} \end{array} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 7 \\ 0 & 0 & 1 & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \left[\begin{array}{cccc|c} 1 & 2 & 0 & 3 & 2 \\ 0 & 0 & 1 & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

Pivot ent — Pivot ent

Reduced row echelon form
leading entry on each row are 1.
zeroed out on the the last row.
The positioning of 1 should be with that

$x_1 + 2x_2 + 3x_4 = 2$
$x_3 - 2x_4 = 5$

Pivot variable $\Rightarrow x_3 = 5 + 2x_4$

are the variable related to pivot entry

Reduced Row Echelon form:

REF (or) RREF

Gauss Elimination — Gauss Jordan Elimination

$\begin{bmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

Everything above and below a 1 is 0

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & * & 0 & 0 & * & 0 & 0 & * \\ 0 & 0 & 1 & 0 & * & 0 & 0 & * \\ 0 & 0 & 0 & 1 & * & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \end{bmatrix}$

These 1s are called pivots or Basics.

Row operations:

i) Replace (You can only "add" and not "subtract", because subtraction is not commutative)

$a + b = b + a$
$a - b \neq b - a$

ii) Interchange (switch two rows)
iii) Scale (multiply by a scalar)

** You work column by column. Do not move to another column until you have gotten all the necessary zeros, then get your pivots (ones)

Hint: If you are working in column 1, use row 1 to get zeros, if you are in column 2 use your row 2 to get zeros etc.

$8x + 6y = 2$
$5x + 4y = -1$

$16x + 12y = 4$
$15x + 12y = -3$

$x = 7 \Rightarrow y = \frac{-36}{4} = -9$

$(7, -9)$.

$\left[\begin{array}{cc|c} 8 & 6 & 2 \\ 5 & 4 & -1 \end{array}\right] \Rightarrow \begin{array}{c} -5\times \\ 8\times \end{array}\left[\begin{array}{cc|c} 8 & 6 & 2 \\ 5 & 4 & -1 \end{array}\right] \sim \left[\begin{array}{cc|c} 8 & 6 & 2 \\ 0 & 2 & -18 \end{array}\right] \sim \left[\begin{array}{cc|c} 16 & 0 & 112 \\ 0 & 2 & -18 \end{array}\right]$

$\Rightarrow -5R_1 + 8R_2 \to R_2$

$-6R_2 + 2R_1 \to R_1$

$\Rightarrow$ Get the zeros and the ones (pivots)

$\sim \left[\begin{array}{cc|c} 1 & 0 & 7 \\ 0 & 1 & -9 \end{array}\right]$

$\frac{1}{16} R_1 \to R_1$
$\frac{1}{2} R_2 \to R_2$

Use REF to solve the system

$x-2x_2 = -1$

$-x_1+3x_2 = 3$

We need the form $\left[\begin{array}{cc|c} 1 & * & -1 \\ 0 & 1 & 3 \end{array}\right]$

$x_1 \quad x_2 =$

$$\left[\begin{array}{cc|c} 1 & -2 & -1 \\ -1 & 3 & 3 \end{array}\right] \sim \left[\begin{array}{cc|c} 0 & 1 & 2 \\ -1 & 3 & 3 \end{array}\right] \left[\begin{array}{cc|c} 1 & -2 & -1 \\ 0 & 1 & 2 \end{array}\right] \Rightarrow \text{general form}$$

$1R_1 + 1R_2 \to R_2$

$$\begin{cases} x_1 - 2x_2 = -1 \\ x_2 = 2 \end{cases}$$

now back substitute

$x_1 - 4 = -1$

$x_1 = 3$

REF needs back substitution

Subspace: S

$\vec{0} \in S \quad \vec{V}_1, \vec{V}_2 \in S \Rightarrow \vec{V}_1 + \vec{V}_2 \in S$ Subspaces are closed under addition

$c \in R \quad \vec{V}_1 \in S \quad c\vec{V}_1 \in S$

$A\vec{x} = 0 \quad N = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}$

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\vec{x}} = 0$$

RREF $\rightarrow$ form - 2 m

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 3 & 4 \\ 1 & 1 & 4 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \\ 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Column Space of a Matrix:

$$\underset{m \times n}{A} = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} \qquad \vec{V}_1, \vec{V}_2 \cdots \vec{V}_n \in \mathbb{R}^m$$

Column space is defined as all of the linear combinations of these vectors.

$C(A) = \text{Span of } (\vec{V}_1, \vec{V}_2 \cdots \vec{V}_n)$ if $\vec{a} \in C(A)$ then

$\vec{a} = c_1\vec{v}_1 + c_2\vec{v}_2 \cdots c_n\vec{v}_n$

$s\vec{a}$ is this in span?

$s\vec{a} = sc_1\vec{v}_1 + sc_2\vec{v}_2 + \ldots + sc_n\vec{v}_n$

$s\vec{a} \in C(A)$

Also $\vec{b} \in C(A)$

$\vec{b} = k_1\vec{v}_1 + k_2\vec{v}_2 + \cdots k_n\vec{v}_n$

$\vec{a} + \vec{b} \in C(A)$

$\Rightarrow (c_1+k_1)\vec{v}_1 + (c_2+k_2)v_2 + \cdots + (c_n+k_n)v_n \in C(A)$

$\{A\vec{x} \mid \vec{x} \in \mathbb{R}^n\} \quad A = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

$A\vec{x} = x_1\vec{v}_1 + x_2\vec{v}_2 + \ldots x_n\vec{v}_n \Rightarrow \{x_1\vec{v}_1 + x_2\vec{v}_2 + \ldots x_n\vec{v}_n \mid x_1, x_2 \ldots x_n \in \mathbb{R}^n\}$

How to find the null space of a matrix:

$N(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}$

Nullspace of A is all of vectors in $\mathbb{R}^n$ when the product of the matrix A and the vector is a zero vector.

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$x_1 + x_2 + x_3 + x_4 = 0$

$x_1 + 2x_2 + 3x_3 + 4x_4 = 0$

$4x_1 + 3x_2 + 2x_3 + x_4 = 0$

$\Rightarrow$ Augmented matrix

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 4 & 0 \\ 4 & 3 & 2 & 1 & 0 \end{array}\right]$$

Use Reduced Row Echelon form:

Consider this matrix:

$$\begin{bmatrix} 1 & 4 & 7 \\ 8 & 2 & 5 \\ 9 & 4 & 6 \end{bmatrix} \Longrightarrow$$ is comprised of the vectors $\begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}, \begin{bmatrix} 8 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 9 \\ 4 \\ 6 \end{bmatrix}$

These vectors can be expressed as

$$\left.\begin{aligned} \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} &= \hat{i}+4\hat{j}+7\hat{k} \\ \begin{bmatrix} 8 \\ 2 \\ 5 \end{bmatrix} &= 8\hat{i}+2\hat{j}+5\hat{k} \\ \begin{bmatrix} 9 \\ 4 \\ 6 \end{bmatrix} &= 9\hat{i}+4\hat{j}+6\hat{k} \end{aligned}\right\}$$ where i, j, k are the basis/unit vectors

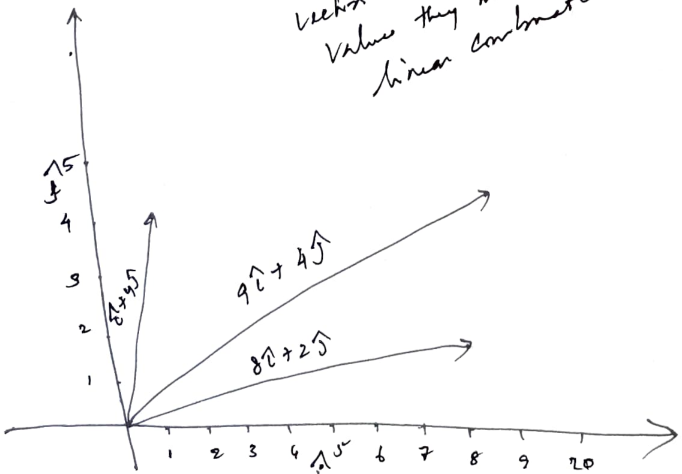
These 3 vectors can be represented in a matrix $\Big\}$ $\begin{bmatrix} 1 & 4 & 7 \\ 8 & 2 & 5 \\ 9 & 4 & 6 \end{bmatrix}$

Let us represent these vectors in a graph
Since it would be difficult to represent a 3D vector
let us consider the following 2D vectors

$\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ $\begin{bmatrix} 8 \\ 2 \end{bmatrix}$ $\begin{bmatrix} 9 \\ 4 \end{bmatrix}$

$\hat{i}+4\hat{j}$ $8\hat{i}+2\hat{j}$ $9\hat{i}+4\hat{j}$

The span of these vectors is all the possible values they make as linear combinations



~~The span of the linear combination~~

The resultant area of linear combination of two vectors is the span of the vectors. $\begin{bmatrix} 1 \\ 4 \end{bmatrix} \begin{bmatrix} 8 \\ 2 \end{bmatrix} \begin{bmatrix} 9 \\ 4 \end{bmatrix}$

If the span of the vectors lie in $\vec{A} = \begin{bmatrix} 1 & 4 \\ 8 & 2 \\ 9 & 4 \end{bmatrix}$ the 2D space then the column space of the matrix $\vec{A} \Rightarrow C(A) = 2$.

Similarly whatever ~~the span~~ dimension the span of given vectors cover becomes the column space of the corresponding matrix.

Linearly dependent vectors:

Assume there are three vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \end{bmatrix}$

If one vector can be expressed as a linear combination of other vectors then they are linearly dependent

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} 3 \\ 6 \end{bmatrix} \times \frac{1}{3} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \begin{bmatrix} 2 \\ 4 \end{bmatrix} \times 1.5 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

Linearly independent vectors:

The vectors that do not result as a linear combination of others

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 7 \\ 11 \end{bmatrix} \begin{bmatrix} 13 \\ 4 \end{bmatrix}$$

~~Thus~~ I think these are linearly independent vectors

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 7 \\ 11 \end{bmatrix} \neq \begin{bmatrix} 13 \\ 24 \end{bmatrix} \text{ or } 2 \times \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 7 \\ 11 \end{bmatrix} = \begin{bmatrix} 13 \\ 4 \end{bmatrix}$$

~~See the~~ Visit next page for consolidated result

Null space of a Matrix

When you get $\vec{0}$ when your matrix is multiplied by a vector or a set of vectors i.e The product of a matrix and the set of all vectors that give $\vec{0} \Rightarrow \{A\vec{x} = \vec{0} \mid \vec{x} \in R^n + A \in m \times n\}$

Consolidated points

linearly dependent vectors

$\Rightarrow$ let $\{\vec{v_1}, \vec{v_2} \ldots \vec{v_n}\}$

If they are linearly dependent

$$\vec{v_1} = \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} \quad \vec{v_2} = \begin{bmatrix} 8 \\ 2 \\ 5 \end{bmatrix} \quad \vec{v_3} = \begin{bmatrix} 9 \\ 4 \\ 6 \end{bmatrix}$$

$$\vec{A} = \begin{bmatrix} 1 & 4 & 7 \\ 8 & 2 & 5 \\ 9 & 4 & 6 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = 0 \qquad \vec{C}$$

$$c_1 \vec{v_1} + c_2 \vec{v_2} + \ldots + c_n \vec{v_n} = 0$$

$$v_1 = -\frac{1}{c_1}(c_2 \vec{v_2} + \ldots + c_n \vec{v_n})$$

If the vectors are linearly dependent then atleast one of the scalar values among $\{c_1, c_2 \ldots c_n\}$ should be non-zero.

In other words you can express the linearly dependent vectors in such a way that their sum or linear combination is zero.

On the contrary linearly independent vectors cannot be expressed by manipulating a scalar alongside it. You can never get a $\vec{0}$. You can only get a $\vec{0}$ when all of $\{c_1, c_2 \ldots c_n\}$ are 0.

So effectively ~~the~~ if you look at the matrix $\vec{A}$ and the vector $\vec{c}$

The null space of a matrix for linearly independent vectors can only be the $\vec{0}$ i.e if only $\vec{c} = \vec{0}$ when $\{\vec{c_1}, c_2 \ldots c_n = 0\}$ occurs.

Column space of a vector is only the basis vectors. i.e in a set of vectors $\{v_1, v_2 \ldots v_n\}$ only the number of linearly independent vectors equals to the column space

$$\vec{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 4 & 3 \\ 3 & 4 & 1 & 2 \end{bmatrix} \qquad C(A) = \text{Span of the column vectors}$$

$$\Rightarrow \begin{array}{l} x + y + z + k \\ 2x + y + 4z + 3k \\ 3x + 4y + z + 2k \end{array}$$

What would be the span of column vectors

$$x\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} + z\begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix} + k\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

It means if the basis vectors x, y, z, k are linearly independent i.e suppose x and z are in the same dimension and can be expressed like $\vec{x} = c\vec{z}$ or something like that then z ~~is~~ does not constitute a column space alone. It belongs to the column space of x or vice versa.

How to determine column space and null space of a matrix

① – $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

To find the column space of A, we should find the basis vectors of A.

To find the basis vectors, we have to find the linearly independent vectors of A

We can figure out the linearly independent vectors by finding the NULL SPACE of A.

If A contains all linearly independent vectors then $N(A) = \vec{0}$ that is the nullspace of A is only the $\vec{0}$.

If the vectors are linearly dependent then the nullspace of A has more than $\vec{0}$.

① find the nullspace of A

$A\vec{x} = \vec{0}$

Determine the value of components of $\vec{x}$

find the redundant (or) linearly dependent vectors

Then determine the column space of A

Inference:

The basis vectors → Mutually independent vectors
↓
Their span is
↓
Column space

A basis is a linearly independent spanning set.

$A = \{a_1, a_2 \dots a_n\}$ basis V

Any other spanning set should have atleast n elements

Claim: $B = \{b_1, b_2 \dots b_m\}$ $m < n$ spans V (falsify the set)

Let me call a new set $B_1' = \{a_1, b_1, b_2 \dots b_m\}$ If $a_1 \in V$ and $b_1, b_2 \dots b_m$ spans V then ~~they~~ a_1 can be linear combination of other vectors. i.e linear combination of B_1' can be used to arrive at any member.

If you keep removing b_x and replace with a_x, at one point we will end up with $\{a_1, a_2 \dots a_m\}$ but $\{a_1, a_2 \dots a_m, \dots, a_n\}$ are linearly independent and forms the basis of the span.

So $\{a_1, a_2 \dots a_m\}$ will not be able to span whatever $\{a_1, a_2 \dots a_m \dots a_n\}$ spans i.e a_n cannot be a linear combination of one or more of $\{a_1, a_2 \dots a_m\}$

Dimension of the nullspace is the number of vectors that one basis for the span (or) the number of linearly independent vectors.

Nullity is the number of free variables or the non pivot variables

Dimension of the column space or Rank:

Dimension of the column vectors

C(A) : Span of each of the column vectors

Span is some sense the more fundamental concept.

If you have any collection of vectors you can take all the possible linear combinations of them. The resulting set of vectors is a vector space called the span of the original collection.

The column space of a matrix is all possible linear combination of the column vectors ~~that~~ make up the matrix

$$\begin{bmatrix} x_1^1 & x_2^1 & x_3^1 \\ x_1^2 & x_2^2 & x_3^2 \\ x_1^3 & x_2^3 & x_3^3 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} + b$$

$$= w_1 \begin{bmatrix} x_1^1 \\ x_1^2 \\ x_1^3 \end{bmatrix} + w_2 \begin{bmatrix} x_2^1 \\ x_2^2 \\ x_2^3 \end{bmatrix} + w_3 \begin{bmatrix} x_3^1 \\ x_3^2 \\ x_3^3 \end{bmatrix} + b$$

$$= \begin{bmatrix} w_1 x_1^1 \\ w_1 x_1^2 \\ w_1 x_1^3 \end{bmatrix} + \begin{bmatrix} w_2 x_2^1 \\ w_2 x_2^2 \\ w_2 x_2^3 \end{bmatrix} + \begin{bmatrix} w_3 x_3^1 \\ w_3 x_3^2 \\ w_3 x_3^3 \end{bmatrix} + b$$

$$\hat{y} = \begin{bmatrix} \hat{y}^1 \\ \hat{y}^2 \\ \hat{y}^3 \end{bmatrix} = \begin{bmatrix} w_1 x_1^1 + w_2 x_2^1 + w_3 x_3^1 + b \\ w_1 x_1^2 + w_2 x_2^2 + w_3 x_3^2 + b \\ w_1 x_1^3 + w_2 x_2^3 + w_3 x_3^3 + b \end{bmatrix}$$

$$J(\theta_0,\theta_1) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

$$h_\theta(x^{(i)}) - y^{(i)}$$

$$\theta_0 + \theta_1 x - y^{(i)}$$

$$h_\theta(x) = \theta_0 + \theta_1 x$$

$$\theta_0 = 0$$

$$\theta_1 = 1$$

$$\frac{1}{8}\left((5-4)^2 + (3-4)^2 + (0-1)^2 + (4-3)^2\right)$$

$$\frac{1}{8}\left(1^2 + 1 + 1 + 1\right)$$

$$\frac{4}{8}$$

$$h_\theta(x) = \theta_0 + \theta_1 x$$

$$-1 + 2\times 6$$

$$u = \begin{bmatrix}4\\-4\\-3\end{bmatrix} \quad v = \begin{bmatrix}4\\2\\4\end{bmatrix}$$

$$u^T = \begin{bmatrix}4 & -4 & -3\end{bmatrix}$$

u_{21} should be equal u^T_{12}

$$\begin{bmatrix}4 & -4 & -3\end{bmatrix}\begin{bmatrix}4\\2\\4\end{bmatrix}$$

1×3 $\quad$ 3×1

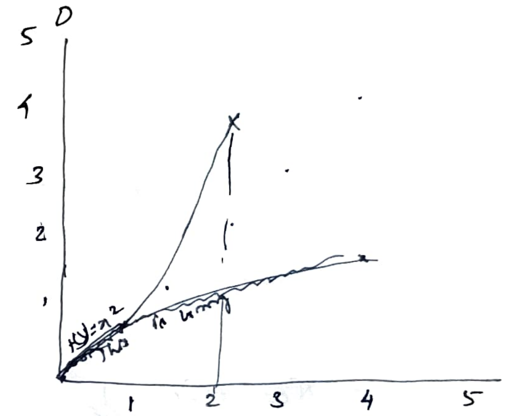
$16 - 8 - 12$

$\cancel{16} + 8 - 12$

-4

$$\frac{dy}{dx} = \frac{\Delta y}{\Delta x}$$

$f(x) = y$



$f(x) = x^2$

$$\frac{dy}{dx} = 2x$$

$f(x) = y^2$

$$\frac{\Delta y}{\Delta x} = 2x$$

let $x = 1$ $y = 1$

$x = 2$ $y = 4$

$\theta_0 + (\theta_1 x^{(i)})^2 + 2\theta_0\theta_1 x^{(i)}$

$h_\theta(x^{(i)})^2 + y(i)^2$

$- 2h_\theta(x^{(i)})\,y(i)$

$$\frac{\partial}{\partial\theta_j}J(\theta_0,\theta_1) = \frac{\partial}{\partial\theta_j}\cdot\frac{1}{2m}\sum_{i=1}^{M} h_\theta\left(x^{(i)}) - y^{(i)}\right)^2$$

$$= \frac{\partial}{\partial\theta_j}\cdot\frac{1}{2m}\sum_{i=1}^{M}\left(\theta_0 + \theta_1 x^{(i)} - y(i)\right)^2$$

$\hookrightarrow (\theta_0 + \theta_1 x^{(i)} + y(i))$

$(\theta_0 + \theta_1 x^{(i)} - y(i))$

θ_0 $j = 0$: $\frac{\partial}{\partial\theta_0}J(\theta_0,\theta_1)$

$$= \frac{\partial}{\partial\theta_0}\frac{1}{2m}\sum_{i=1}^{M}\left(\theta_0 + \theta_1 x^{(i)} - y(i)\right)^2$$

$$= \cancel{\partial}\ 2\times\frac{1}{2m}\sum_{i=1}^{M} 1 + \theta_1 x^{(i)} - y(i) = \frac{1}{m}\sum_{i=1}^{m} 1 + \theta_1 x$$

θ_1 $j = 1$ $= \frac{\partial}{\partial\theta_1}\frac{1}{2m}\sum_{i=1}^{m}\left(\theta_0 + \theta_1 x^{(i)} - y(i)\right)^2$

Logistic Regression

Email: Spam / Not Spam?

Online Transactions: fraud (Yes / No)

Tumor: Malignant / Benign?

$y \in \{0, 1\}$ 0: Negative class
↳ benign tumor
1: Positive class
↳ Malignant tumor

Usually Negative class conveys the absence of something
Positive class conveys the presence of something

$h_\theta(x) = g(\theta^T x)$ → $g(z) \geq 0.5$

$P(y=1 | x; \theta)$

$h_\theta(x) = g(\theta^T x) \geq 0.5$ Zoom

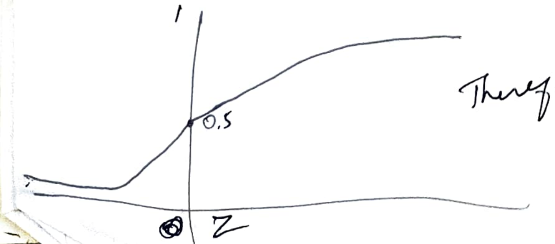
$g(\theta_0 + \theta_1 x_1 + \theta_2 x_2) \geq 0.5$

For $g(z) \geq 0.5$

$z \geq 0$

Therefore $z = \theta_0 + \theta_1 x_1 + \theta_2 x_2$

$\theta_0 + \theta_1 x_1 + \theta_2 x_2 \geq 0$



Suppose $\theta_0 = 5$ $\theta_1 = -1$ $\theta_2 = 0$ so that $h_\theta(x) = g(5 - x_1)$

i.e. $z = 5 - x$ $\frac{1}{1 + e^{-z}}$

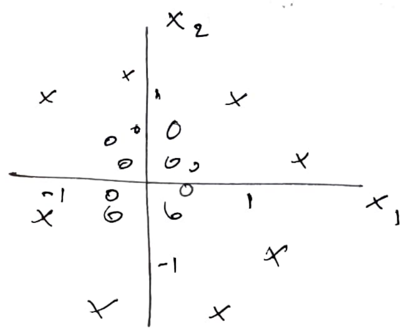
$\Rightarrow 5 + (-1 \times x_1) + (0 \times x_2) \geq 0$

$5 - x_1 \geq 0$

$x_1 \leq 5$ ~~$-x_1 \geq -5$~~ ~~$x_1 \geq 5$~~

$x_1 \leq 5$

$g(5 - x_1) \geq 0.5$
for that
$5 - x_1 \geq 0$



$x_1 \geq 5$

$\frac{1}{1 + e^{-z}}$

$z \Rightarrow 0 = \frac{1}{2}$

$z \Rightarrow -\alpha \Rightarrow e^{\alpha} = \alpha \Rightarrow 1$

$z = \alpha \Rightarrow e^{-\alpha} = 1 \to 0$

Logistic Regression Cost function

$$J(\theta) = \frac{1}{m} \sum_{i=1}^{m} Cost(h_\theta(x^{(i)}), y^{(i)})$$

$$Cost(h_\theta(x^{(i)}), y) = \begin{cases} -\log(h_\theta(x)) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

Note: $y = 0$ or 1 always
This equation can be written in one line

$$Cost(h_\theta(x), y) = -y \log h_\theta(x) - (1-y) \log (1 - h_\theta(x))$$

$$\Rightarrow J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)} \log h_\theta(x^{(i)}) + (1 - y^{(i)}) \log (1 - h_\theta(x^{(i)}))\right]$$

Principle of Maximum likelihood estimation

is used to arrive at the cost function

Study when you are free

Also the cost function that we came up is shown,

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)} \log h_\theta(x^{(i)}) + (1-y^{(i)}) \log(1-h_\theta(x^{(i)}))\right]$$

$$\frac{\partial}{\partial \theta_j} J(\theta) = \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})\, x_j^{(i)}$$

Gradient Descent formula:

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

$$\Rightarrow \theta_j - \alpha \times \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})\, x_j^{(i)}$$

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix}$$

-X: Even though the update rules of linear regression and logistic regression look identical

$$\theta_j := \theta_j - \alpha \times \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})\, x_j^{(i)}$$

$h_\theta(x)$ differs

(PTO)

Linear Regression $\rightarrow h_\theta(x) \rightarrow \theta^T x \Rightarrow \theta_0 x_0 + \theta_1 x_1 \ldots + \theta_n x_n$

Logistic Regression $\rightarrow h_\theta(x) \rightarrow \frac{1}{1+e^{-\theta^T x}}$

So they are not the same

$$g(X\theta) = \begin{vmatrix} x_0 & x_1 \\ x_0 & x_2 \\ x_0 & x_3 \end{vmatrix} \begin{vmatrix} \theta_0 \\ \theta_1 \end{vmatrix} = 3\times 1 \text{ matrix} = \begin{vmatrix} x_0\theta_0 + x_1\theta_1 \\ x_0\theta_0 + x_2\theta_2 \\ x_0\theta_0 + x_3\theta_3 \end{vmatrix}$$

$3\times 2 \quad 2\times 1$

$$\theta := \theta - \frac{\alpha}{m} X^T (g(X\theta) - \vec{y})$$

$$\begin{vmatrix} x_0 & x_0 & x_0 \\ x_1 & x_2 & x_3 \end{vmatrix} \begin{vmatrix} x_0\theta_0 + x_1\theta_1 \\ x_0\theta_0 + x_2\theta_1 \\ x_0\theta_0 + x_3\theta_1 \end{vmatrix} = \begin{vmatrix} 2\times 1 \end{vmatrix}$$

$2\times 3 \quad 3\times 1$

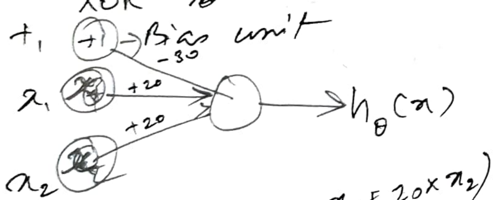
XOR NEURAL NETWORKS

Basic Rules	Symbol	True when
Binary operator, Two input	$\oplus$	inputs are different

Let 0 = False

p	q	XOR $p \oplus q$	AND $p \wedge Q$	OR $p \mid Q$	XNOR
0	0	False 0	False	False	TRUE
1	0	True 1	False	True	FALSE
0	1	True 1	False	True	FALSE
1	1	False 0	True	True	TRUE

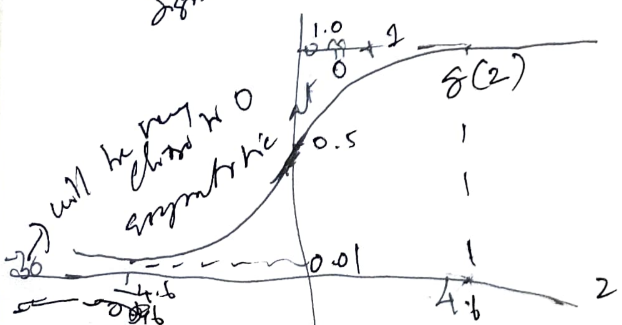
XOR $\Rightarrow$ NOT (XOR) = XNOR



$$h_\theta(x) = g(-30 + 20 \times x_1 + 20 \times x_2)$$

θ^1_{10}, θ^1_{11}, θ^1_{12}

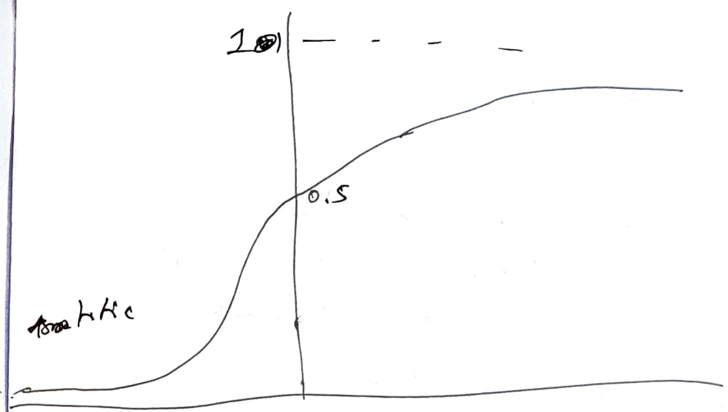
Sigmoid function



x_1	x_2	$h_\theta(x)$
0	0	$g(-30) \approx 0$
0	1	$g(-30+20) = g(-10) \approx 0$
1	0	$g(-10) \approx 0$
1	1	$g(10) \approx 1$

From the previous table 0 0 0 1

It is obvious $h_\theta(x) = 1$ if and only if $x_1 = 1 = x_2$



x_1	x_2	h_θ
0	0	$g(-10) \approx 0$
1	0	$g(10) \approx 1$
0	1	$g(10) \approx 1$
1	1	$g(30) \approx 1$

Represent logical OR function

$g(-10 + 20x_1 + 20x_2)$

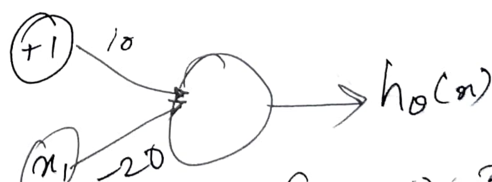
$x_1, x_2 = (0,0) \Rightarrow g(-10)$

$x_1, x_2 = (1,0)$ $g(-10+20+0) \Rightarrow g(10)$

$x(0,1)$ $g(-10+0+20) \Rightarrow g(10)$

$x(1,1)$ $g(-10+20+20)$ $g(30)$

NEGATION : NOT X1



$$h_\theta = g(10 - 20x_1)$$

x_1	$h_\theta(x)$
0	$g(10) \approx 1$
1	$g(-10) \approx 0$

These values are essentially NOT X1 function

General idea of negation is to put a large negative value before it

For ~~tagegs~~ egs

$-20x_1$

NOT x_1 AND NOT x_2 $\Rightarrow$ x_1 x_2 20

Qn

Suppose you have a multi-class classification problem with 10 classes. Your neural network has 3 layers and the hidden layer (layer 2) has 5 units. Using the one vs all method described here, how many elements does $\Theta^{(2)}$ have

a_0^2 θ_0^2

a_1^2 θ_1^2

a_2^2 θ_2^2

a_3^2 θ_3^2

a_4^2 θ_4^2

a_5 θ

θ dim is

$S_{(j+1)} \times S_j + 1$

$\Rightarrow$ 10×6

$10 \times (5+1)$

$= 60$

$\delta(-20+30x_1+30x_2)$

x_1	x_2	h_θ
0	0	$\delta(-20) \approx 0$
1	0	$\delta(10) \approx 1$
0	1	$\delta(10) \approx 1$
1	1	$\delta(40) \approx 1$

x_0	θ_0
x_1	θ_1
x_2	θ_2

x_0

1	0.5	1.9	x
1	1.2	2.7	y
			z

$x + 0.5y + 1.9z$

$x + 1.2y + 2.7z$

$a21 = a21 + x(1) \cdot \theta(1,1)$

$\Rightarrow a21 + x(2) \cdot \theta(1,2)$

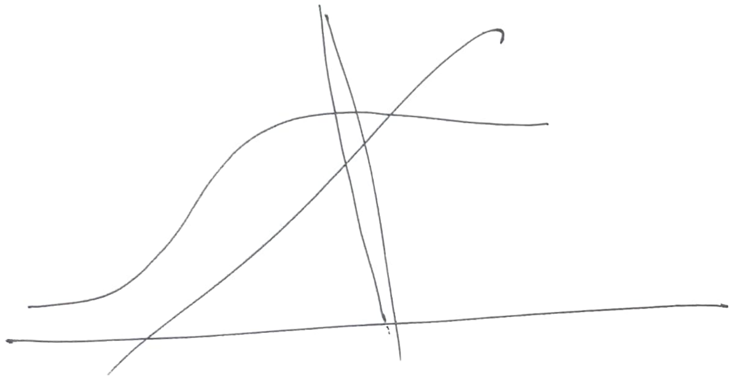
$\Rightarrow a21 + x(3) \cdot \theta(1,3)$

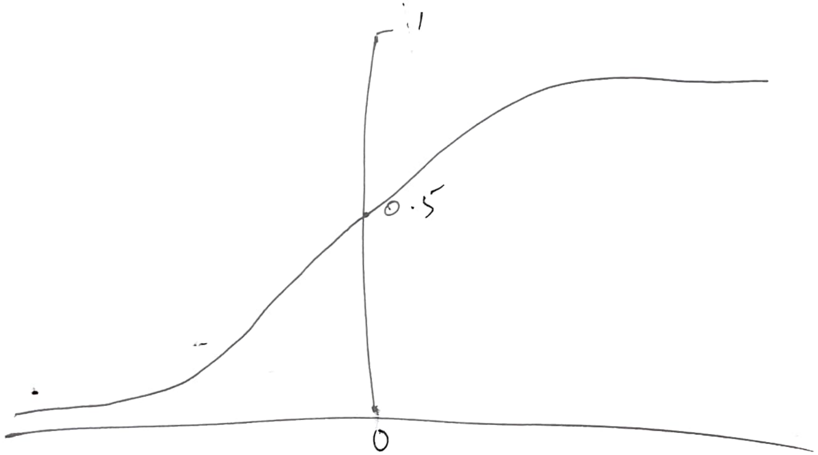
$a22 \Rightarrow a22 + x(1) \cdot \theta(2,1)$

$\Rightarrow a22 + x(2)$

$\Rightarrow a22 +$

$\approx a_{23}$





1 -1.7 -0.2

$g(\cancel{-0.5})$

$g(\cancel{1.2})$

x_1	x_2	h_θ
1	1	$g(-0.5) \simeq 0$
0	1	$g(1.2) \simeq 1$
1	0	$g(-0.7) \simeq 0$
0	0	$g(1) \simeq 1$

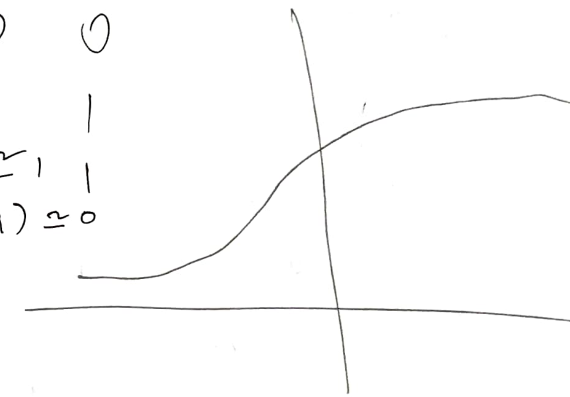
~~1 -0.2 -0.7~~

30 20 -20

x_1	x_2	h_θ
0	0	$g(30) \simeq 1$
1	0	$g(10) \simeq 1$
0	1	$g(10) \simeq 1$
1	1	$g(-20) \simeq 0$

1 -0.2 -1.7

x_1	x_2	$h(\theta)$	
0	1	$g(-0.7)$	0
1	0	$g(0.8)$	1
0	0	$g(1) \simeq 1$	1
1	1	$g(-0.9) \simeq 0$	



Swap 1 -1.7 -0.2

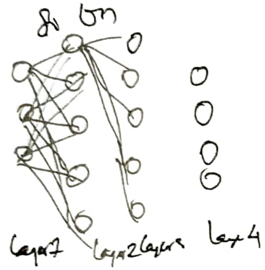
x_1	x_2					
0	1	$g(0.8)$	1	0	1	1
1	0	$g(-0.7)$	0	0	1	1
0	0	$g(1)$	1	0	~~0~~	0
1	1	$g(-0.9)$	0	1	~~0~~	0

AND

-30 20 20

x_1	x_2			
0	0	$g(-30) \simeq 0$	0	1
1	0	$g(-10) \simeq 0$	1	0
0	1	$g(-10) \simeq 0$	1	1
1	1	$g(10) \simeq 1$	0	0

Neural Network:



Training $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots (x^{(m)}, y^{(m)})\}$

L = total no of layers in network

S_l = no of units (not counting bias unit) in layer l

$s_1 = 3, \; s_2 = 5$

$h_\theta(x) \in \mathbb{R}$

k output units

Multi-class Classification

$y \in \mathbb{R}^k$ Eg $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$

Cost function: Logistic Regression

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)} \log h_\theta(x^{(i)}) + (1-y^{(i)}) \log(1-h_\theta(x^{(i)}))\right] + \frac{\lambda}{2m}\sum_{j=1}^{n}\theta_j^2$$

Understanding differentiation

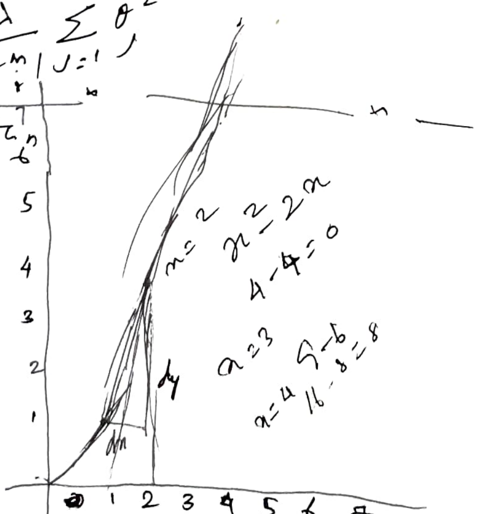
$y = f(x) = x^2$

$\frac{dy}{dx} = 2x$

when $x=1$ $\frac{dy}{dx} = 2$

$x=2$ $\frac{dy}{dx} = 4$

$x=8$ $\frac{dy}{dx} = 16$



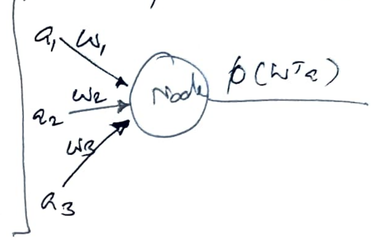
Brian Dolhansky

ϕ: Sigmoid function

$w = \theta$

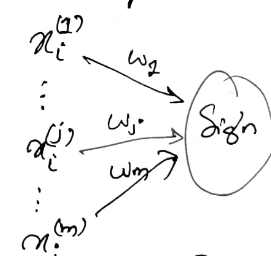
a = activation units

$\phi(\sum w_i a_i) = \phi(w^T a)$ → Visually this looks like



Our goal is to train a network using labelled data so that we can then feed it a set of inputs and it produces the appropriate outputs for unlabelled data. We can do this because we have both the input x_i and the desired target output y_i in the form of data pairs. Training in this case involves learning the correct edge weights to produce the target output given the input

Single layer Perceptron:



$\to h(x_i, w) = \text{sign}(w^T x_i) = y_i$

$$\text{sign}(w^T x_i) = \begin{cases} 1 & \text{if } w^T x_i > 0 \\ 0 & \text{if } w^T x_i = 0 \\ & \text{if } w^T x_i < 0 \end{cases}$$

The decision boundary is ~~$w^T x_i = 0$~~ the line where it is equally probable that a point on the line belongs to either class

$h(x_i, w) = \text{sign}$

Back propagation algorithm:

Layer L

$\delta^{(L)} = y - a^{(L)}$ The last layer is the difference between the vector value of actual results and correct results

For other layers

$\delta^{(l)} = ((\Theta^{l})^T \delta^{(l+1)}) . g'(z^{(l)})$

$g'(z^{(l)}) = a^{l} . (1 - a^{l})$

$$\frac{\partial J(\Theta)}{\partial \Theta_{ij}^{(l)}} = \frac{1}{m} \sum_{t=1}^{m} a_j^{(t)(l)} \delta^{(t)(l+1)}$$

$\delta^{(l+1)}$ and $a^{(l+1)}$ are vectors with s_{l+1} elements

Similarly $a^{(l)}$ is a vector of s_l elements

Back propagation algorithm

Given training set $\{(x^{(1)}, y^{(1)}) \ldots (x^{(m)}, y^{(m)})\}$

Set $\Delta_{ij}^{(l)} := 0$ for all (l, i, j)

For training example $t = 1$ to m:

Set $a^{(1)} := x^{(t)}$

Perform forward propagation to compute $a^{(l)}$ for $l = 2, 3 \ldots L$

using $y^{(t)}$ compute $\delta^{(L)} = \cancel{\ } a^{(L)} - y^{(t)}$

Compute $\delta^{(L-1)}, \delta^{(L-2)} \ldots \delta^{(2)}$ using $\delta^{(l)} = ((\Theta^{l})^T \cancel{\ } \delta^{(l+1)}) . * a^{(l)} . * (1 - a^{(l)})$

$\Delta_{ij}^{(l)} := \Delta_{ij}^{(l)} + a_j^{(l)} \delta_i^{(l+1)}$ or with vectorization $\Delta^{l} := \Delta^{l} + \delta^{(l+1)} (a^{(l)})^T$

$D_{ij}^{(l)} := \frac{1}{m} (\Delta_{ij}^{l} + \lambda \Theta_{ij}^{l})$ ...

without regularization

$$\frac{\partial}{\partial \Theta} J(\Theta) \approx \frac{J(\Theta + \epsilon) - J(\Theta - \epsilon)}{2\epsilon}$$

With multiple theta matrices, we can approximate the derivative with respect to Θ_j as follows

$$\frac{\partial}{\partial \Theta} J(\Theta) \approx \frac{J(\Theta_1, \ldots \Theta_j + \epsilon, \ldots \Theta_n) - J(\Theta_1, \ldots \Theta_j - \epsilon \ldots \Theta_n)}{2\epsilon}$$

A small value for ϵ (epsilon) such $\epsilon = 10^{-4}$

$$\frac{J(\Theta + \epsilon) - J(\Theta - \epsilon)}{2\epsilon}$$

$$\frac{J(1.01) - J(.99)}{.02}$$

Bias / variance :

Bias is essentially underfitting

High variance is essentially overfitting

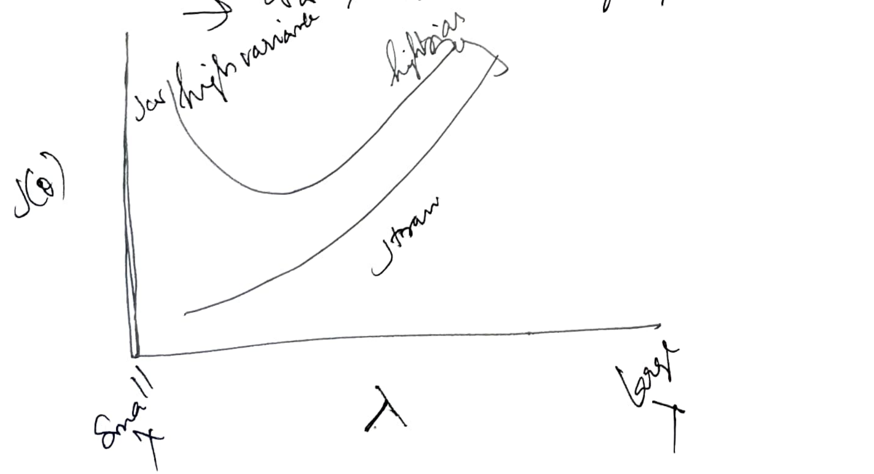
Bias { Training, Crossvalidation } → Errors → high

Variance { Training → low errors; Crossvalidation → high errors }

Regularization → Bias / Variance :

λ → high λ causes underfitting or bias

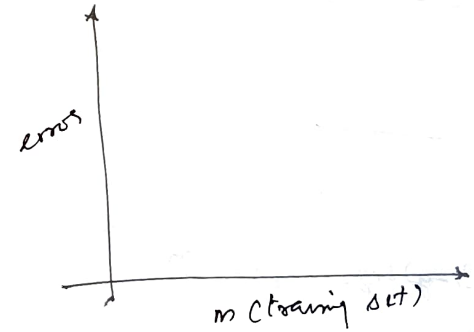
→ low λ causes overfitting or variance



Learning curves :

$$J_{train}(\theta) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$$

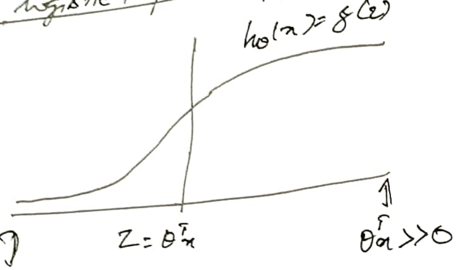
$$J_{cv}(\theta) = \frac{1}{2m_{cv}} \sum_{i=1}^{m_{cv}} (h_\theta(x_{cv}^{(i)}) - y_{cv}^{(i)})^2$$



Support Vector Machines:- (SVM)

Alternative view of logistic regression:

$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$



Cost of example $= -y \log h_\theta(x) + (1-y)\log(1-h_\theta(x)))$

$$= -y \log \frac{1}{1+e^{-\theta^T x}} + (1-y)\log\left(1-\frac{1}{1+e^{-\theta^T x}}\right)$$

When z or $\theta^T x$ is very large then the function $\frac{1}{1+e^{-\theta^T x}}$ contributes ~~very little~~ less to the cost function.

Support Vector Machines ⇒ Large Margin Classifier

Maths behind support vector machines

Vector inner product.

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$\theta_0 + \theta_1 x_1 + \theta_2 x_2 \geq 0$

3,1,0

$3 + 1 \times x_1 + 0 \times x_2$

$= 3 + x_1$

$x_1 \geq -3$

-3,1,0

$-3 + x_1 + 0 \times x_2 \geq 0$

$x_1 \geq 3$

3,0,1

$3 + x_2 \geq 0$

$x_2 \geq -3$

-3,0,1

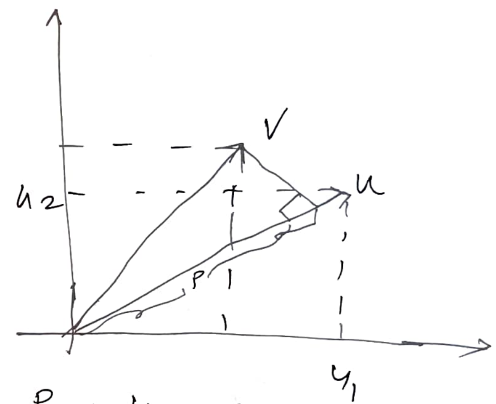
$-3 + x_2 \geq 0$

Vector Inner Product ⇒ $u^T v$

[is also called as]

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$u^T v = ?$



(norm of u)

$\|u\|$ = length of vector u

By Pythagoras theorem

$u = \sqrt{u_1^2 + u_2^2} \in \mathbb{R}$

Similarly $v = \sqrt{v_1^2 + v_2^2}$

p is the length of the projection of the vector v onto vector u

$\in \mathbb{R}$

$u^T v = p \times \|u\|$

$= \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

$= u_1 \times v_1 + u_2 v_2$

$u^T v = v^T u$

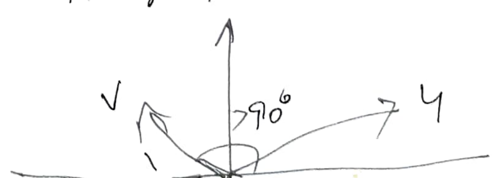
$= \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$

$= [v_1 u_1 + v_2 u_2]$

which means

You can also project vector u on vector v.

P = signed



$u^T v =$

If angle is > 90° then p would be negative

SVM Decision Boundary :

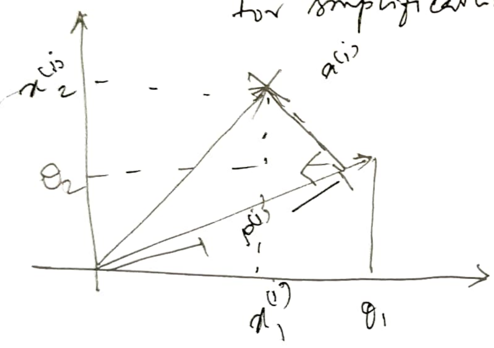
$\min_\theta \frac{1}{2}\sum_{j=1}^{n} \theta_j^2$

s.t. $\theta^T x^{(i)} \geq 1$ if $y^{(i)} = 1$

$\theta^T x^{(i)} \leq -1$ if $y^{(i)} = 0$

for simplification ignore θ_0 or set $\theta_0 = 0$

features $n = 2$



When $n = 2$

$$\min_\theta \frac{1}{2}\sum_{j=1}^{n} \theta_j^2 = \frac{1}{2}(\theta_1^2 + \theta_2^2) = \frac{1}{2}\left(\sqrt{\theta_1^2 + \theta_2^2}\right)^2$$

$$= \frac{1}{2} \times ||\theta||^2$$

$\theta^T x^{(i)} = ?$

$u^T v$

$\theta_0 = 0 \quad \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$

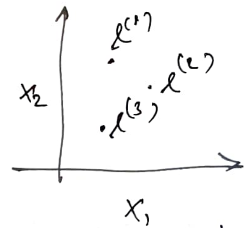
$\theta^T x^{(i)} = p^{(i)} \cdot ||\theta||$

$= \theta_1 x_1^{(i)} + \theta_2 x_2^{(i)}$

$p^{(i)} \cdot ||\theta||$

Kernels 1 :: Gaussian kernel

kernel 1



landmarks $l^{(1)}, l^{(2)}, l^{(3)}$

Given x: $f_1 = \text{similarity}(x, l^{(1)}) = \exp\left(-\frac{||x - l^{(1)}||^2}{2\sigma^2}\right)$ sigma

$f_2 = \text{similarity}(x, l^{(2)}) = \exp\left(-\frac{||x - l^{(2)}||^2}{2\sigma^2}\right)$

$f_3 =$ ...

The similarity function is called kernel

This is Gaussian kernel

Can also be written as $k(x, l^{(i)})$

$f_1 \approx \exp\left(-\frac{0^2}{2\sigma^2}\right) \approx 1$

When x is close to the landmark

$f_1 = \exp\left(\frac{(\text{large number})^2}{2\sigma^2}\right) \approx 0$

kernels 2: SVM

Where and how do you get landmarks

Q.

$$\downarrow x^1$$

$f_1^{(i)} = \mathrm{Sim}(x^{(i)}, l^{(1)})$

So basically it will run through all

$x^{(i)} \rightarrow$ (training examples)

So the above will be like

$f_1^{(0)}, f_1^{(1)}, f_1^{(2)}, f_1^{(3)}$ till $f_1^{(m)}$

Unsupervised algorithms:

Clustering: $k \to$ clusters
μ - cluster centroids

K-Means algorithm:

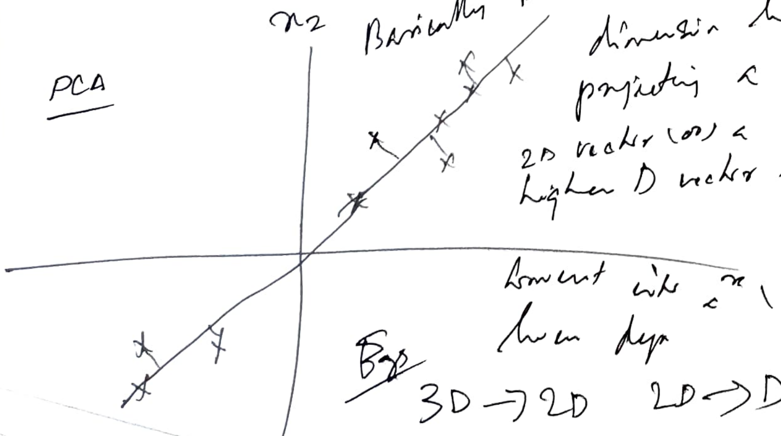
Randomly initialize k cluster centroids
$\mu_1, \mu_2 \ldots \mu_k \in \mathbb{R}^n$

Repeat {
for $i = 1$ to m
$c^{(i)}$:= index (from 1 to k) of cluster centroid closest to $x^{(i)}$
for $k = 1$ to k
μ_k := average (mean) of points assigned to cluster k
}

Random initialization:

Should have $k < m$ — No of clusters and cluster centroids less than m

PCA



Basically PCA reduces the dimension by projecting a 2D vector (or) a higher D vector to lower dim

Eg: 3D → 2D   2D → D

Dimensionality reduction:

VENKATRAMANAN. R

This is basically reducing multiple dimensions of a feature into 2D (or) 3D

Principal component Analysis:

The most popular dimensionality reduction algo.

Reconstruction from compressed representation:

1000 D to 100 D (or) feature vector ~~with~~

Ability to uncompress

$Z = U^T_{reduce} x$

$X_{approx} = U_{reduce} \times Z^{(i)}$ → This product of $U_{reduce} \cdot Z^{(i)}$ uncompresses the compressed data to an acceptable approximation

Choosing k (number of principal components)

Average squared projection error: $\frac{1}{m}\sum_{i=1}^{m} \| x^{(i)} - x^{(i)}_{approx} \|^2$ → Difference between original x and the reconstructed x

Total variation in the data
$= \frac{1}{m}\sum_{i=1}^{m} \| x^{(i)} \|^2$

Typically, choose k to be the smallest value so that

$$\frac{\frac{1}{m}\sum_{i=1}^{m} \| x^{(i)} - x^{(i)}_{approx} \|^2}{\frac{1}{m}\sum_{i=1}^{m} \| x^{(i)} \|^2} \le 0.01 \quad (1\%)$$

99% of variance is retained

90°

$\begin{bmatrix}0&-1\\1&0\end{bmatrix}$

Original î ĵ

$\begin{bmatrix}1&0\\0&1\end{bmatrix}$

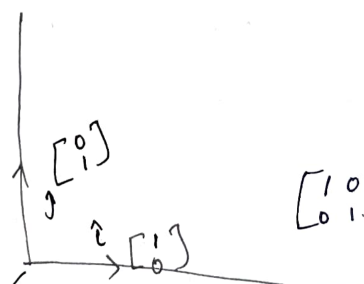
Suppose the space is rotated by 90

$\begin{bmatrix}0&-1\\1&0\end{bmatrix}$ To rotate back or make the transformation to $\begin{bmatrix}1&0\\0&1\end{bmatrix}$ would be

A^{-1} this would change by î.

$\begin{bmatrix}0&1\\-1&0\end{bmatrix}$

$$\begin{array}{cc} i & j \end{array}$$
$\begin{bmatrix}0&-1\\1&0\end{bmatrix}\begin{bmatrix}0&1\\-1&0\end{bmatrix} =$



$\begin{bmatrix}1&0\\0&1\end{bmatrix}$

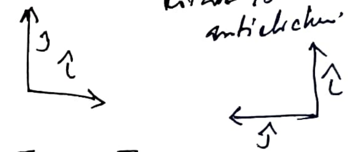
Rotate 90° anticlock

$\begin{bmatrix}0&-1\\1&0\end{bmatrix}$

ĵ we derive

Inverse of a Matrix :

Rotate 90° anticlockwise



$\begin{bmatrix}1&0\\0&1\end{bmatrix}$ $\begin{bmatrix}0&-1\\1&0\end{bmatrix}$

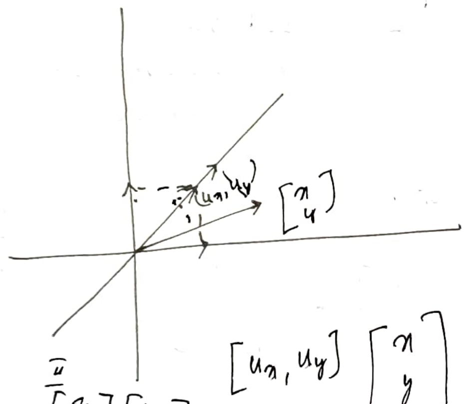
What is the inverse of this matrix

Identity Matrix

$\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}3&3\\4&4\end{bmatrix}$

$= \begin{bmatrix}3&3\\4&4\end{bmatrix}$

Vector dot product ::



$[u_x, u_y] \begin{bmatrix} x \\ y \end{bmatrix}$

I

$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

$\vec{a} \times \vec{b} = \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{bmatrix}$

II

Vector Cross product

$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

$\vec{a} \times \vec{b} = \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ \cancel{a_1b_3 - a_3b_1} \\ a_1b_2 - a_2b_1 \end{bmatrix}$

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

$\begin{bmatrix} 1 \\ -7 \\ 1 \end{bmatrix} \times \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -28-2 \\ 5-4 \\ 2+35 \end{bmatrix} = \begin{bmatrix} -30 \\ 1 \\ 37 \end{bmatrix}$

This vector is orthogonal to both these vectors

$\vec{a}\cdot\vec{c}$ or $\vec{a}\cdot\vec{b}$ should be 0

2) $\begin{bmatrix} 1 \\ -7 \\ 1 \end{bmatrix} \begin{bmatrix} -30 \\ 1 \\ 37 \end{bmatrix} = -30 - 7 + 37 = 0$

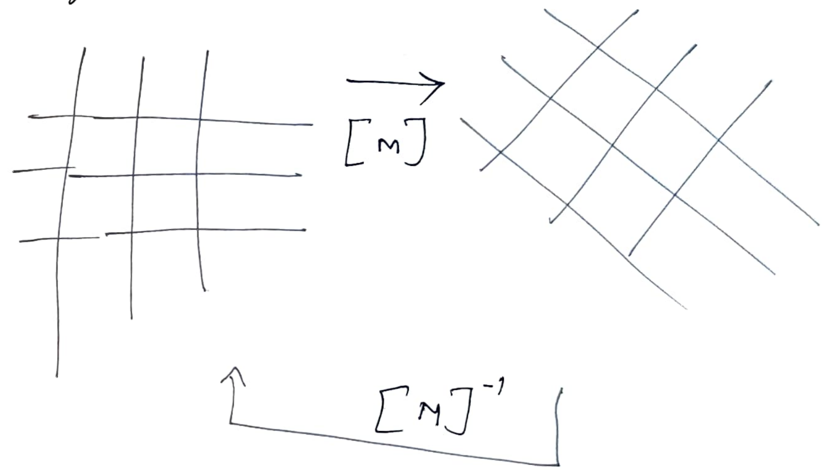
$\begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix} \begin{bmatrix} -30 \\ 1 \\ 37 \end{bmatrix} = -150 + 2 + 148 = 0$

Hence proved

Notes ::

Inverse of a Matrix ::

The transformation that is needed to reset the first transformation

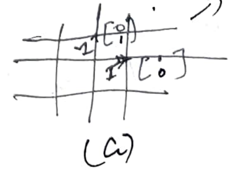


If M is a 2×2 matrix then

$MM^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ the unit vectors or identity matrix

Change of basis vectors (or) change of co-ordinate system

The orientation and the size of the grids are different



If there's a vector in b, what would be that vector in a's co-ordinates.

So Move the basis vectors of b to a's

Find the position of these vectors

b's $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ would transform to a's.

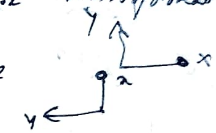
Eigen vectors and Eigen values: 1

Eigen vectors' ~~do not~~ span does not change when the space on which they exist undergo a matrix ~~information~~ transformation. The factor by which the eigen vector scales is its eigen value.

Abstract vector spaces:

You are dealing with a space that exists independently from the co-ordinates you give it. The co-ordinates are arbitrary that depends on the choice of the basis vectors.

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ makes a 90° counter clockwise transformation



the the matrix transformation would be

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

The inverse of this transformation is

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

90° Rotation counter clockwise

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

90° Rotation clockwise

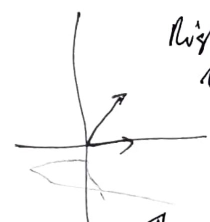
$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

Clockwise 90°

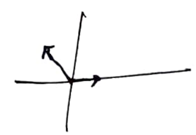
x becomes $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$ y becomes $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$



Rightward shear

$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

The inverse is leftward shear



$\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$

Cauchy Schwartz inequality

$\vec{x}, \vec{y} \in \mathbb{R}^n \quad |\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \, \|\vec{y}\|$

$\vec{V} = \begin{vmatrix} V_1 \\ V_2 \\ V_3 \end{vmatrix}$

$\|\vec{V}\| = \sqrt{v_1^2 + v_2^2 + v_3^2} \geqslant 0$

$\|\vec{V}\| = \vec{V} \cdot \vec{V}$

$\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$

$$\begin{array}{r} 29 \\ 14 \\ \hline 116 \\ 29 \\ \hline 406 \end{array} \quad 3$$

$|\vec{x} \cdot \vec{y}| = 2 + 6 + 12 = 20$

$\|\vec{x}\| \|\vec{y}\| = \sqrt{1^2+2^2+3^2} \times \sqrt{2^2+3^2+4^2}$

$= \sqrt{1+4+9} \times \sqrt{4+9+16}$

$= \sqrt{14} \times \sqrt{29}.$

$= \sqrt{406} > 20$

$\Downarrow$

20.149

Dot Products of two vectors is always less than or equal the absolute value (or) height of the products of the same two vectors.

Revision :

key indexed counting :

0 - 255

or let us consider

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z

Lets take an array like ①

①

d	3
f	5
a	0
a	0
c	2
e	4
e	4
e	4
f	5
d	3
d	3
b	1
c	2
b	1
a	0

Now create an array that represents the alphabets as its index and add up the count in each index. We will limit to a, b, c, d, e, f. Let the 0th index be unpopulated

	0
a	3
b	2
c	2
d	2
e	3
f	2

After adding the counts now cumulate it

	0
a	3
b	5
c	7
d	9
e	12
f	14

Now there are 3 chars lesser than b

5 < c 2 > a

0
d
f

```
int[] a = new int[array.length]
int[] count = new int[R+1]
for (int i=0; i<a.length; i++)
    count[a[i]+1]++
for (int j=0; j<count.length; j++)
    count[j+1] = count[j]+count[j+1])

for (int k=0; k<aux.length; k++)
    aux[count[a[k]]++] = a[k]
                          3
    aux[7] =
```