Calculus

15 March 2025 09:49

✯ Why Calculus

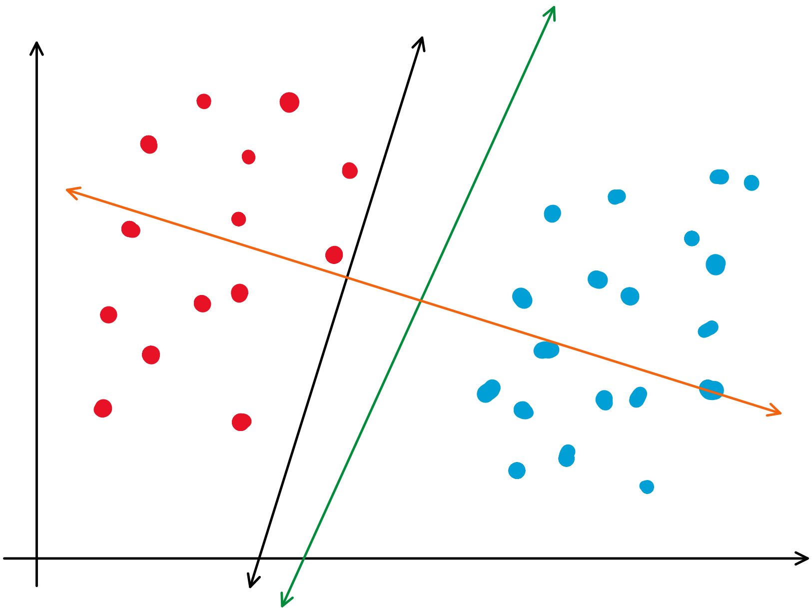
✯ Functions

✯ Limit

✯ Some Imp. Functions

✯ Why Calculus?

→ Which of the following lines is the best seperator?



→ Amongst orange & black -

Ans: Black

Why: Accuracy

$$\text{accuracy} = \frac{\text{no. of correctly classified point}}{\text{total points}} \times 100$$

$$acc_b = \frac{100}{\cancel{100}} \times \cancel{100}\ \% = 100\%$$

$$acc_o = \frac{(25+25)}{\cancel{100}} \times \cancel{100}\ \% = 50\%$$

∴ Accuracy of black seperator > Accuracy of Orange seperator

∴ Black classifier is better than the orange.

→ But which one of Black & Green is a better boundary? And why?

boundary? And why?

$acc_g = \frac{100}{\cancel{100}} \times \cancel{100}\% = 100\%$ ∴ Accuracy of black & green both is the same!

But the the value of loss function of green line (Error) is very high for orange line, less for black line & very low for green line and hence, green is the best choice out of these 3.

∴ We are interested in finding the line for which the value of the loss function is minimum.

If the eq^n of that line is $\vec{w} \cdot \vec{x} + w_0 = 0$

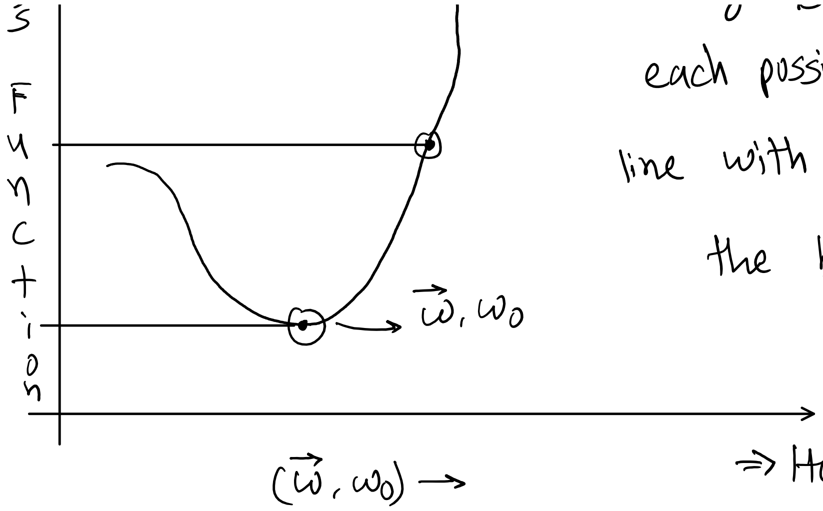
where $\vec{w} = [w_1, w_2, w_3, \ldots, w_n]$ & $\vec{x} = [x_1, x_2, x_3, \ldots x_n]$

(making the equation: $w_1x_1 + w_2x_2 + \ldots + w_nx_n + w_0 = 0$) then,

we are interested in finding $\vec{w}$ & w_0 for **which, the value of the loss function is minimum.**

Suppose, we draw a graph of Loss function $\rightarrow (\vec{w}, w_0)$ and it comes to be as this:

If we find $\vec{w}, w_0$ for minimum using Brut-force (Finding loss for each possible line & then picking the



each possible line & then picking the line with minimum loss), will it be the best approach?

<u>Ans</u>: No. Why?

⇒ How will we do this?

→ We will take some random values of $\omega_1, \omega_2, \ldots, \omega_n$ ($\vec{\omega}$) & ω_0 and then find value of Loss function using

$$L(\vec{\omega}, \omega_0, x, y) = -\sum_{i=1}^{n} \frac{\omega^T . x_i + \omega_0}{\|\vec{\omega}\|} . y_i$$

→ But there are infinite possible values of such $\vec{\omega}$ & ω_0. Even if we take values of $\vec{\omega}$ & ω_0 in the interval $[0, 20)$ with the step size of 0.1 and there are just two features x_1 & x_2 ($\therefore$ $\vec{\omega}$ has only two components ω_1 & ω_2 along with ω_0)

Variable	Possible Values
ω_1	200
ω_2	200
ω_0	200

This way, we will have $200 \times 200 \times 200 = 80,00,000$ lines to find loss function

→ Now imagine taking $[-200, 200)$ interval with the step size 0.01 & having 7 features!

... loss function!!

step size 0.01 & having 7 features!

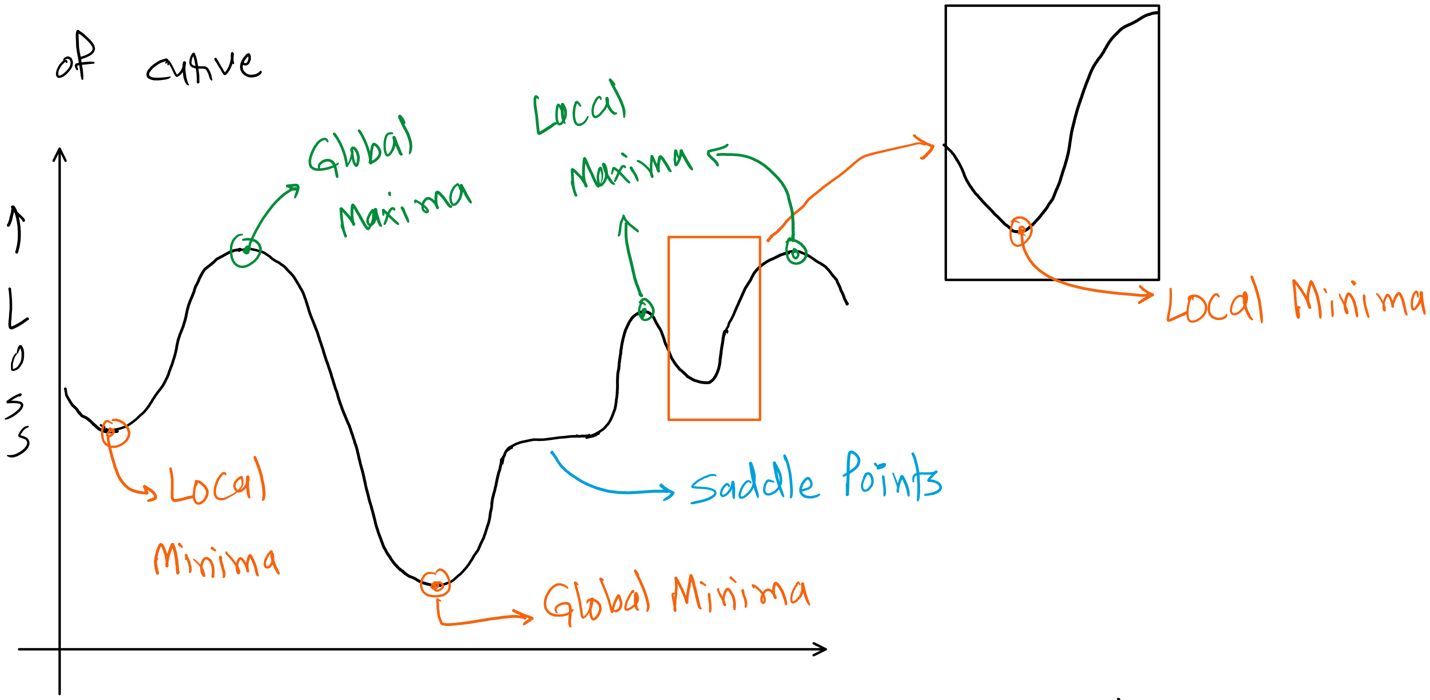
$20{,}000^8$ lines we will have to compute loss function!!

This is going to be computationally very expensive.

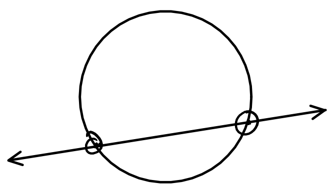
∴ We need a smarter way to do this & calculus can help us doing so.

★ How will calculus help us to solve this problem?

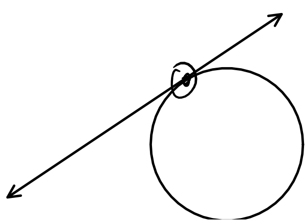
→ To understand that, let's first learn a few jargons of curve

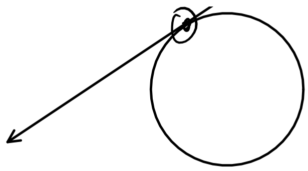


→ Chord - A line that **intersects** a curve/circle.

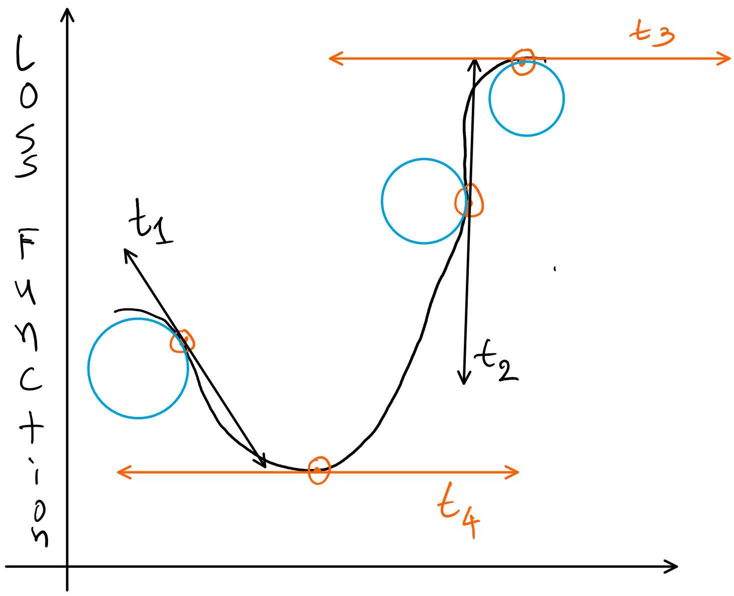


→ Tangent - A line that **touches** a curve





→ If we draw tangents to our curve at various points, they will look like this :-



An interesting observation:

The slope of the tangent at the minima or at the maxima (or even at the saddle points) is zero.

Hence, we can say that if we find the tangent whose slope is zero then the point is either minima or maxima or a saddle point.

And the slope of a curve at a particular point is called **differentiation** of that **function** at that point.

∴ We need to understand concepts of **functions & differentiation.**

✱ Functions

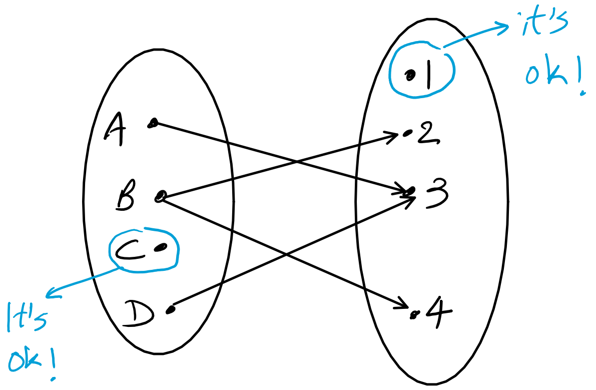
→ A Relation (that is not a function)

Relation : 'x is sister of y'

→ A Relation that is a Function

function : 'x has got grade y'

Relation : 'x is sister of y'

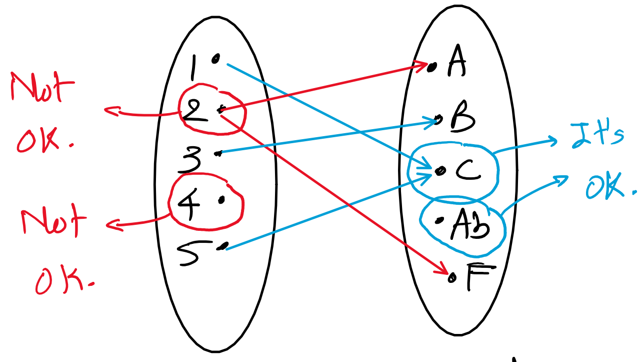


Girls → Domain

Boys → Codomain

Range = {2, 3, 4}

function : x has got grade y



students → Domain

grades → Codomain

Range = { A, B, C, F }

→ Examples of functions :

① $f(x) = x^2$

$$\underset{=3}{x} \longrightarrow \boxed{\begin{array}{c}\text{Function}\\ f(x)=x^2\end{array}} \longrightarrow \underset{=9}{y} = f(x) = x^2$$

$f(-2) = (-2)^2 = 4$

$f(2) = (2)^2 = 4$

$f(4) = (4)^2 = 16$

② $f(x, y) = 3x^2 + 2y$

$$\begin{array}{l} x \longrightarrow \\ y \longrightarrow \end{array} \boxed{f(x,y)} \longrightarrow z$$

$f(2, 3) = 3(2)^2 + 2 \times 3 = 18$

③ $f(x_1, x_2) = w_1 x_1 + w_2 x_2 + w_0$ where $w_1 = 2, w_2 = -3, w_0 = 4$

$\therefore f(x_1, x_2) = 2x_1 - 3x_2 + 4$

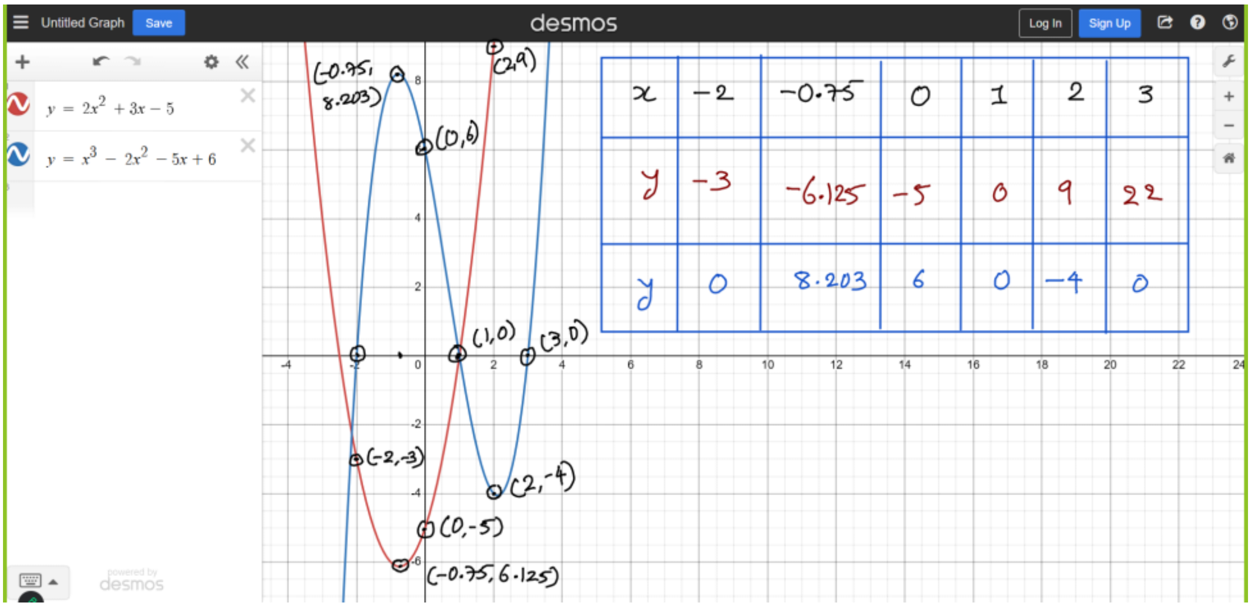
If we want to change the line, we need to take new values of w_1, w_2 & w_0. A more general form of a Hyperplane is :

④ $f(x_1, x_2, \dots, x_n) = w_1 x_1 + w_2 x_2 + \dots + w_n x_n + w_0$

④ $f(x_1, x_2, \ldots, x_n) = w_1x_1 + w_2x_2 + \ldots + w_nx_n + w_0$

where $\underbrace{w_1, w_2, \ldots, w_n}_{\vec{w}}$ & w_0 are some constants.

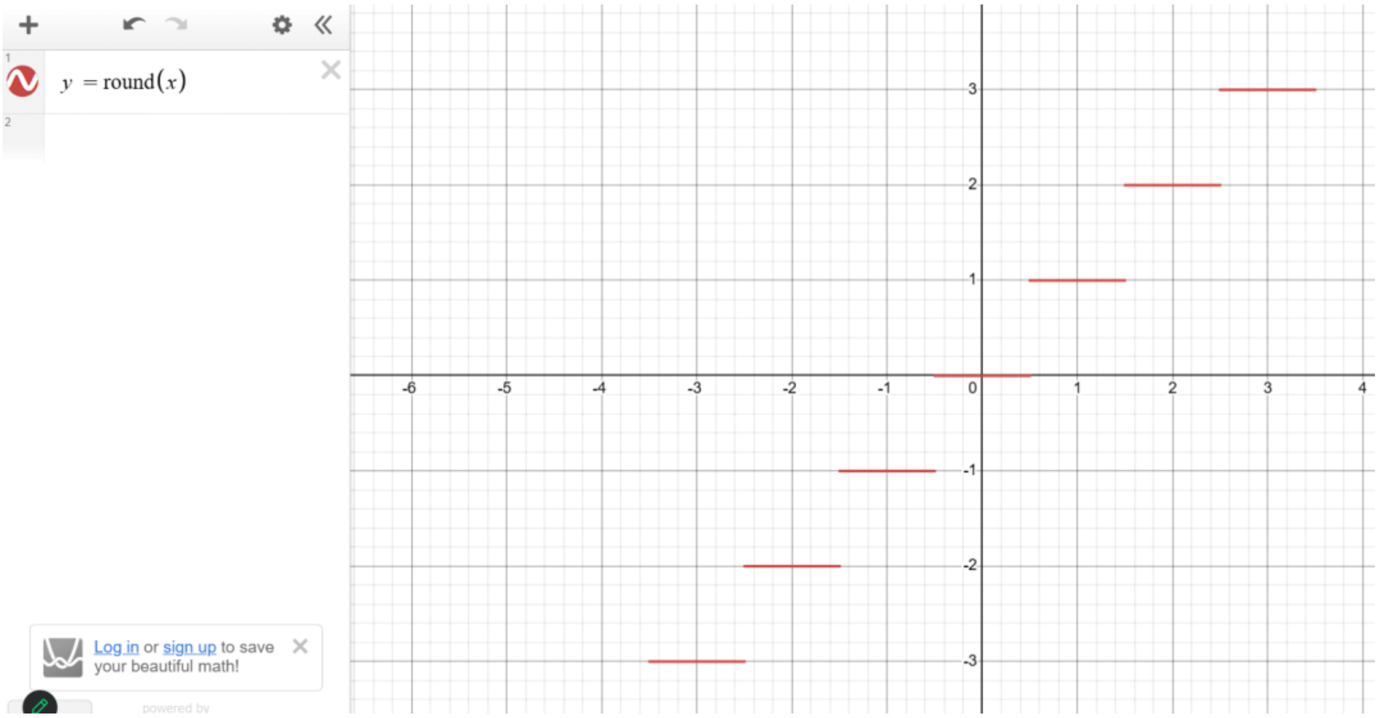
→ If we have a function, we can also draw its curve.

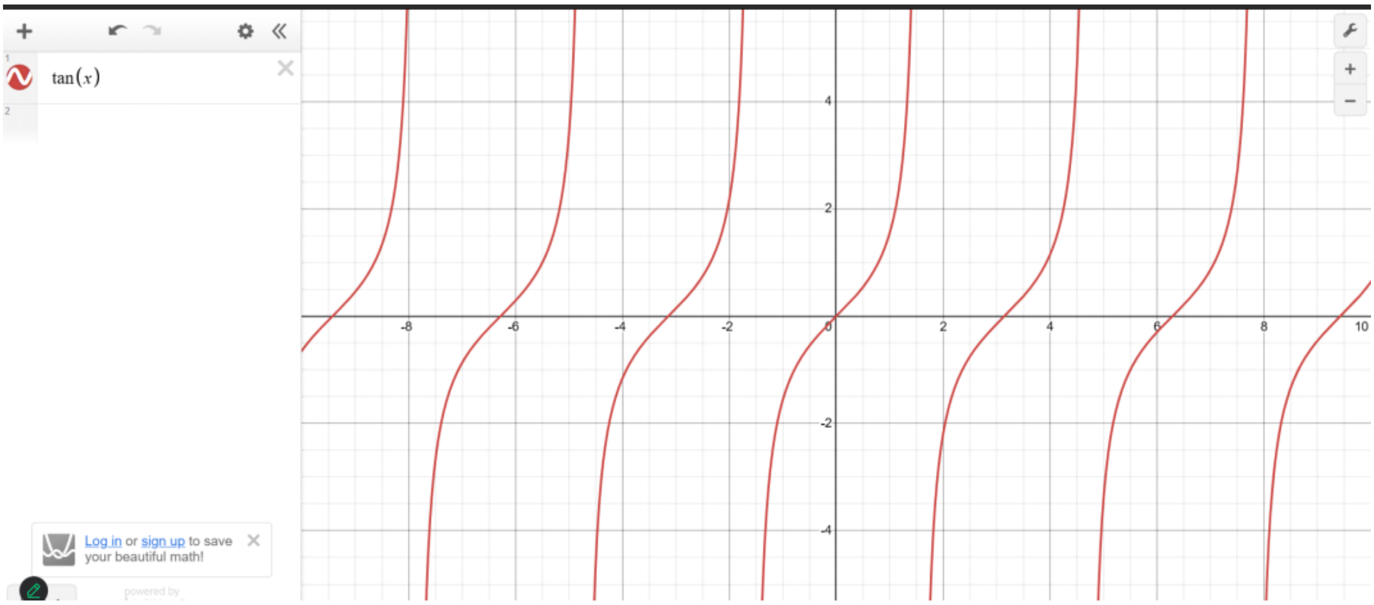


x	-2	-0.75	0	1	2	3
y	-3	-6.125	-5	0	9	22
y	0	8.203	6	0	-4	0

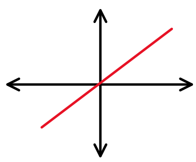
→ strategy – We will make a graph (curve) of our loss function & then we will find (somehow) minima of that curve. That point will give us line that has minimum loss (error) or the best classifier.

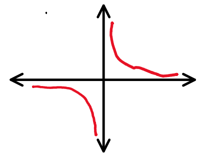
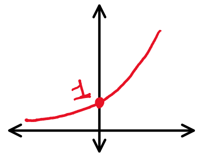
→ **Continuous Functions** – Functions whose curves can be drawn without lifting the pen are called continuous functions.





→ Some Important Functions

Function	Domain	Range	Continuous	Graph
① $y = x$	$[-\infty, +\infty]$	$[-\infty, +\infty]$	Y	

② $y = \frac{1}{x}$	$[-\infty, +\infty]$	$[-\infty, +\infty] - \{0\}$	N	
③ $y = e^x$	$[-\infty, +\infty]$	R^+	Y	

→ Understanding log :

$10^2 = 100 \rightarrow$ If $x^2 = 100$ then $x = ?$

$$\Rightarrow \sqrt{x^2} = \sqrt{100} \Rightarrow x = 10$$

∴ We take root both the sides.

But what if the question is this:

If $10^x = 2$ then $x = ?$ <u>OK</u> $10^x = 1000$ then $x = ?$

∴ We need an operator that helps us solving this type of problems. $\boxed{\text{Solution} = \log}$

power

base $10^{x} = 1000 \Rightarrow x = \log_{10} 1000$

Now we know that the answer of $10^x = 1000$ is $x = 3$

$$\boxed{\therefore \log_{10} 1000 = 3}$$

$10^x = 2 \Rightarrow x = \log_{10} 2$

From log table, $\log_{10} 2 = 0.3010$

∴ $10^{0.3010}$ must be equal to 2

In general, $\boxed{\log_{p} q = r \iff p^{r} = q}$

p: base, r: power

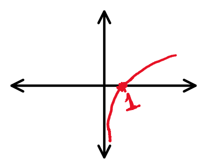
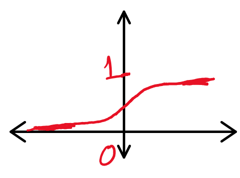
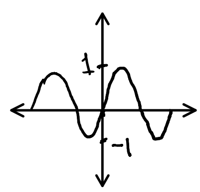
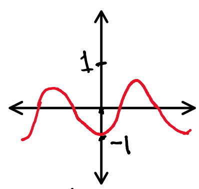
$\log_{10} 2 = ? \Rightarrow \log_{10} 2 = x \Rightarrow 10^{x} = 2 \Rightarrow x = 0.3010$

$\left.\begin{array}{l} \log_{10} 1 = ? \Rightarrow \log_{10} 1 = x \Rightarrow 10^{x} = 1 \Rightarrow x = 0 \\ \log_{e} 1 = ? \Rightarrow \log_{e} 1 = x \Rightarrow e^{x} = 1 \Rightarrow x = 0 \end{array}\right\} \therefore \log_{b} 1 = 0$

$\log_{10} -10 = ? \Rightarrow \log_{10} -10 = x \Rightarrow 10^{x} = -10 \qquad 10^{-5} = \frac{1}{10^{5}} = 0.00001 > 0$

$\therefore$ log of -ve numbers do not exist.

$\log_{10} 10 = \underline{1}$, $\log_{10} 100 = \underline{2}$, $\log_{10} 1000 = \underline{3}$, $\log_{10} 10{,}000 = \underline{4}$

Function	Domain	Range	Continuous?	graph
(4) $y = \log x = \log_{10} x$ $y = \ln x = \log_{e} x$	R^{+}	R	Y	
(5) $y = \frac{1}{1+e^{-x}}$ (sigmoid)	R	$[0, 1]$	Y	
(6) $y = \sin\theta$	R	$[-1, -1]$	Y	
(7) $y = \cos\theta$	R	$[-1, 1]$	Y	
(8) $y = \tan\theta$	R	$[-\infty, +\infty]$	N