Sigmoid function:-

⇒ It is specifically, is used as gating function for 3 gates (in, out, forget in LSTM.

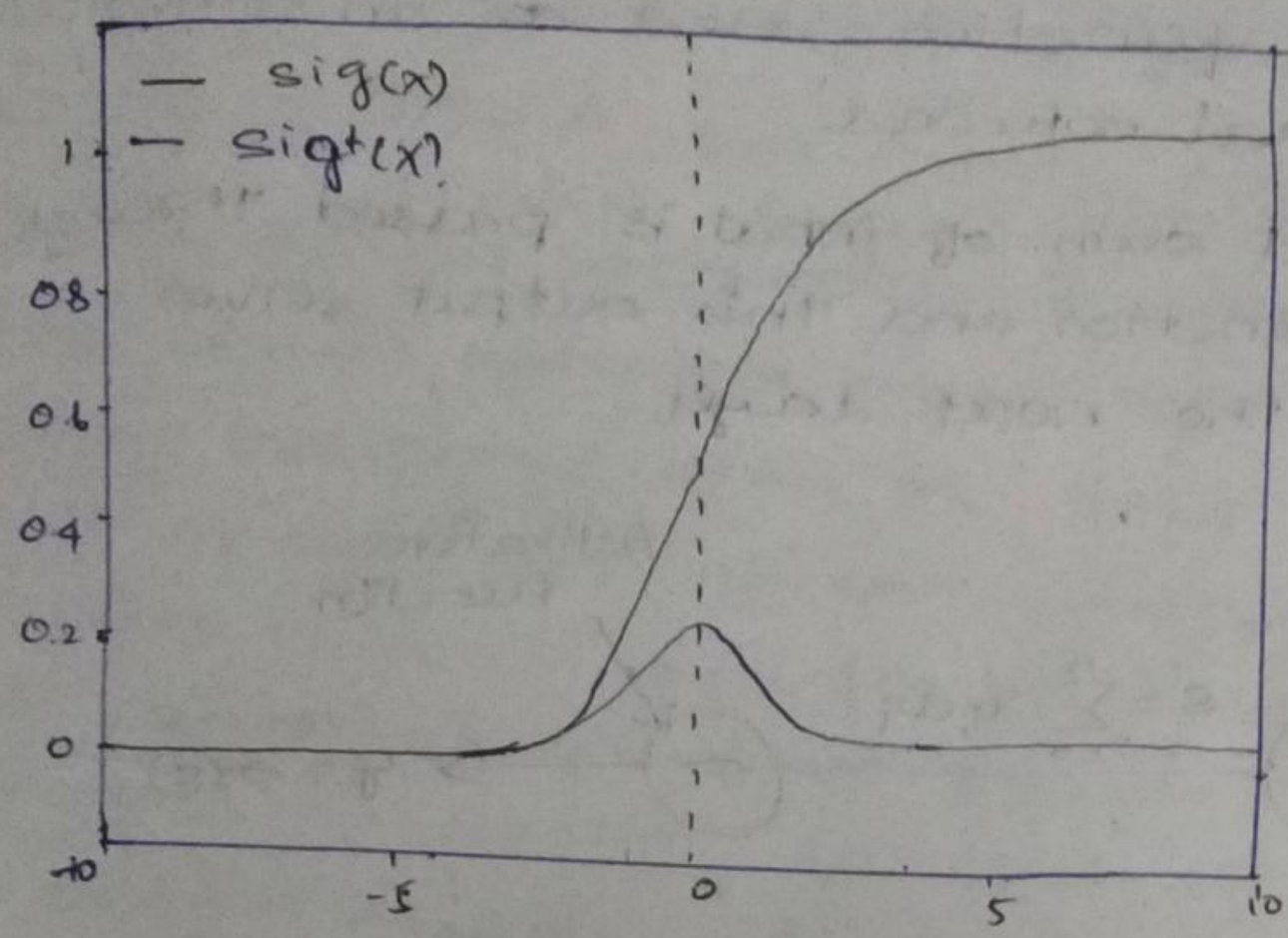
⇒ It outputs a value b/w 0 and 1, and it can either let no flow or complete flow of information throughout the gates

⇒ sigmoid function is a special form of logistic function and is usually denoted by $\sigma(x)$ or $sig(x)$. It is given by

$\sigma(x) =$

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

Properties & Identities of $\sigma(x)$ function:



Plot of sigmoid fun and its derivative.

$$\sigma(x) = 1 - \sigma(-x)$$

$$\sigma = \frac{1}{1+e^{-x}} = \frac{e^x}{e^x+1}$$

$$\sigma'(x) = \sigma(x)(1-\sigma(x))$$

① Thei properties:

* Domain $(-\infty, +\infty)$, Range $(0, +1)$, $\sigma(0) = 0.5$
* The function is monotonically increasing.
* The function is continuous everywhere.
* The function is differentiable everywhere in its domain.
* For values < -10 → the fens value is almost 0
For values > 10, the functions value are almost 1.

The sigmoid as squashing Function:

⇒ sigmoid function is also called a squashing function and its domain is the set of all real nums $(-\infty, +\infty)$, and its range is $(0,1)$

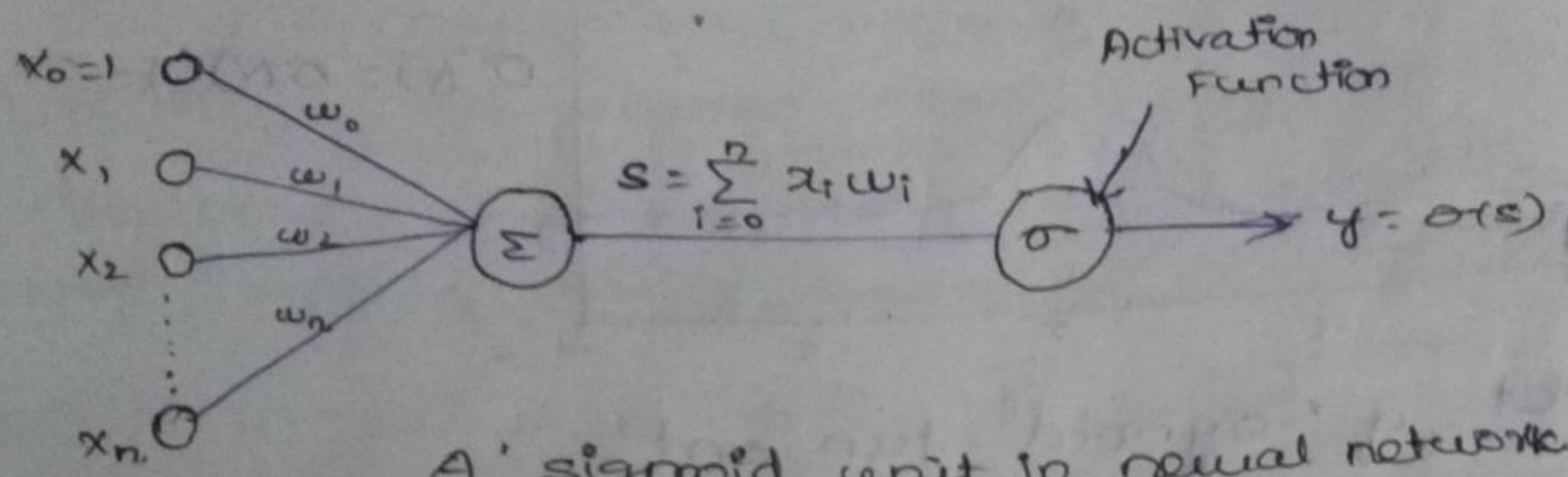
⇒ if the input to the functn is either a very large -ve num or very large +ve

number. The output is always between 0 and 1.

Sigmoid As an Activation Function in Neural Network:

⇒ Sigmoid function used as an activation function in neural networks.

⇒ A weighted sum of input is passed through an activation function and this output serves as an input to the next layer.



A sigmoid unit in neural network

⇒ When the activation function for neuron is a sigmoid function is guaranteed then the output value is between 0 and 1.

⇒ Sigmoid is a non-linear function, so the output also would be non-linear function of inputs.

⇒ If a neuron employs a sigmoid function as an activation function it is termed as a sigmoid unit.

Why sigmoid Function is Important in neural Networks?

⇒ If we use a linear activation function in neural network, then the model can only learn linearly separable problems.

⇒ However, with addition of just 1 hidden layer and a sigmoid activation function in the hidden layer, the neural network can easily

learn a non-linearly seperable problem.

⇒ Using a non-linear function produces non-linear boundaries and hence, the sigmoid function can be used in neural networks for learning complex decision functions.

~~Advantage~~: Why sigmoid function is not used in hidden layers:

Sigmoid activation is computationally slow & the neural network may not converge fast during training. when the input values are too small or too high, it can cause the neural network to stop learning, this issue is known as the vanishing gradient problem. This is why the sigmoid activation function should not be used in hidden layers.

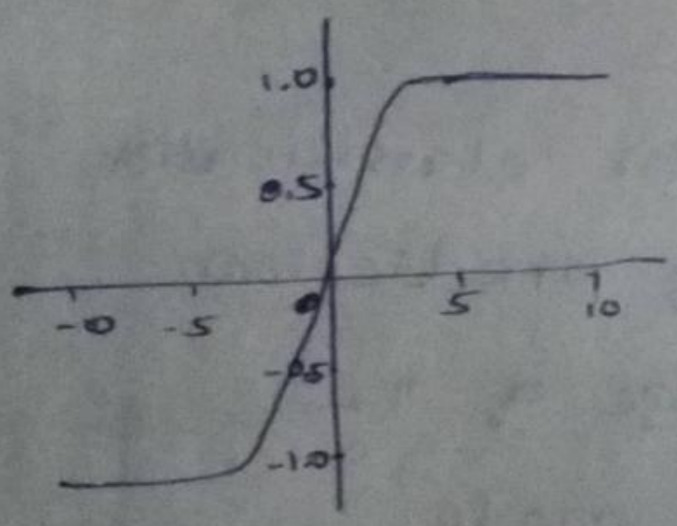
Tanh Activation:-

⇒ Tanh Activation is an activation function used for neural networks:

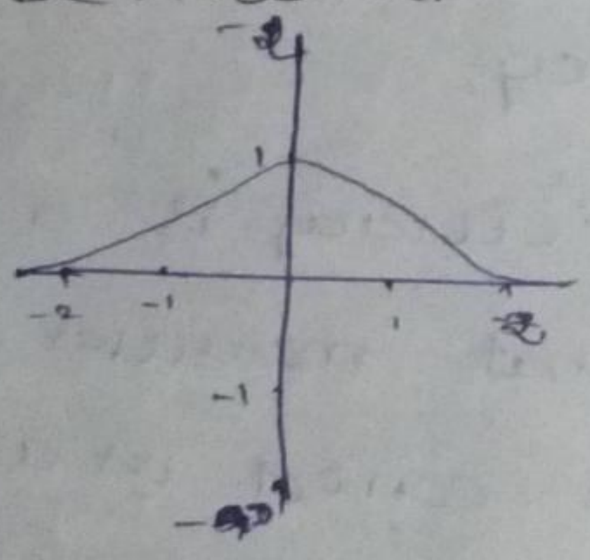
$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

⇒ Tanh squashes a real-valued number to the range [-1, 1]. It's non-linear too

Tanh Function:-



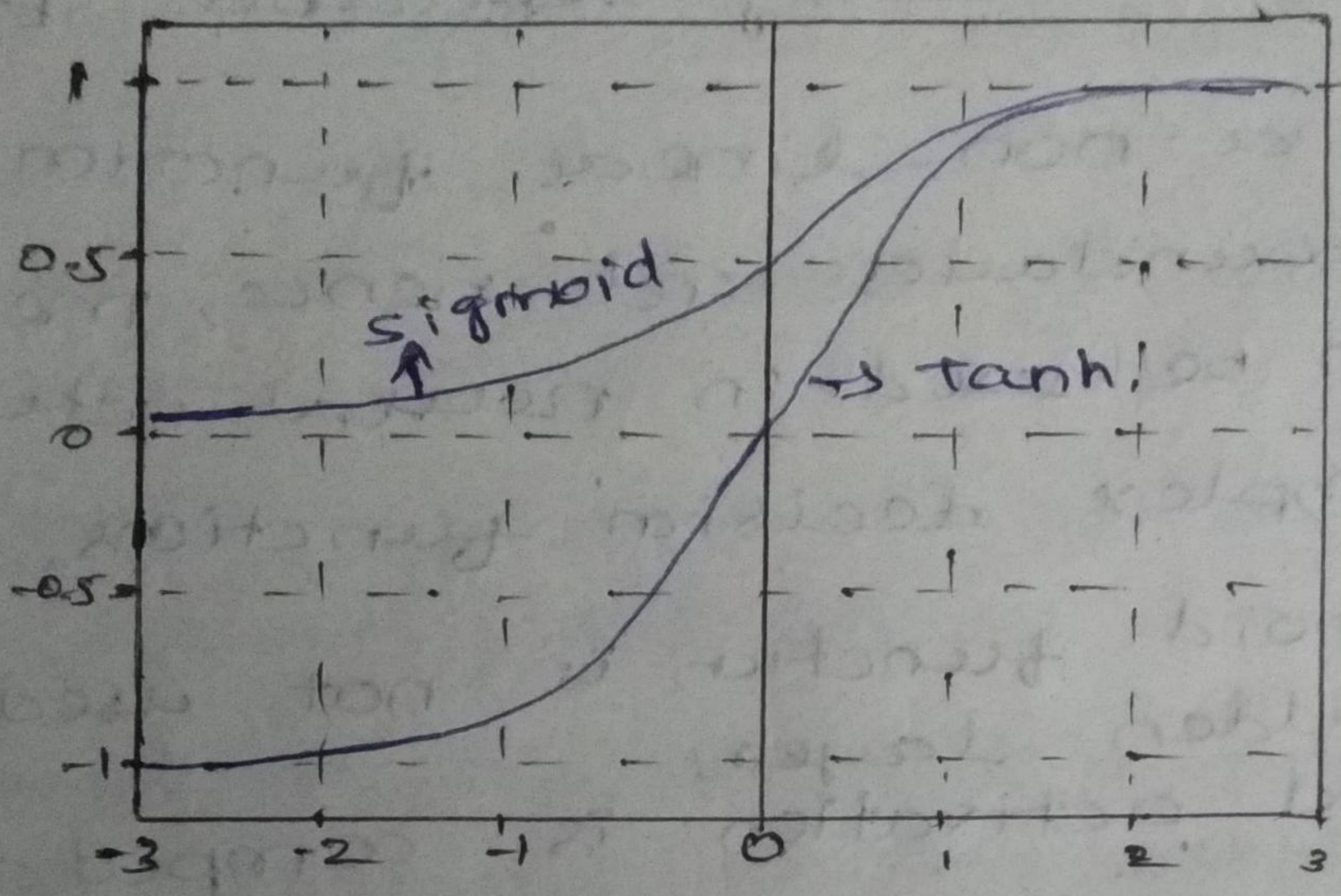
Derivative:



$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

$$\tanh'(z) = 1 - \tanh(z)^2$$

Derivative function give us almost same a sigmoids derivative function.



tanh Vs sigmoid fenction.