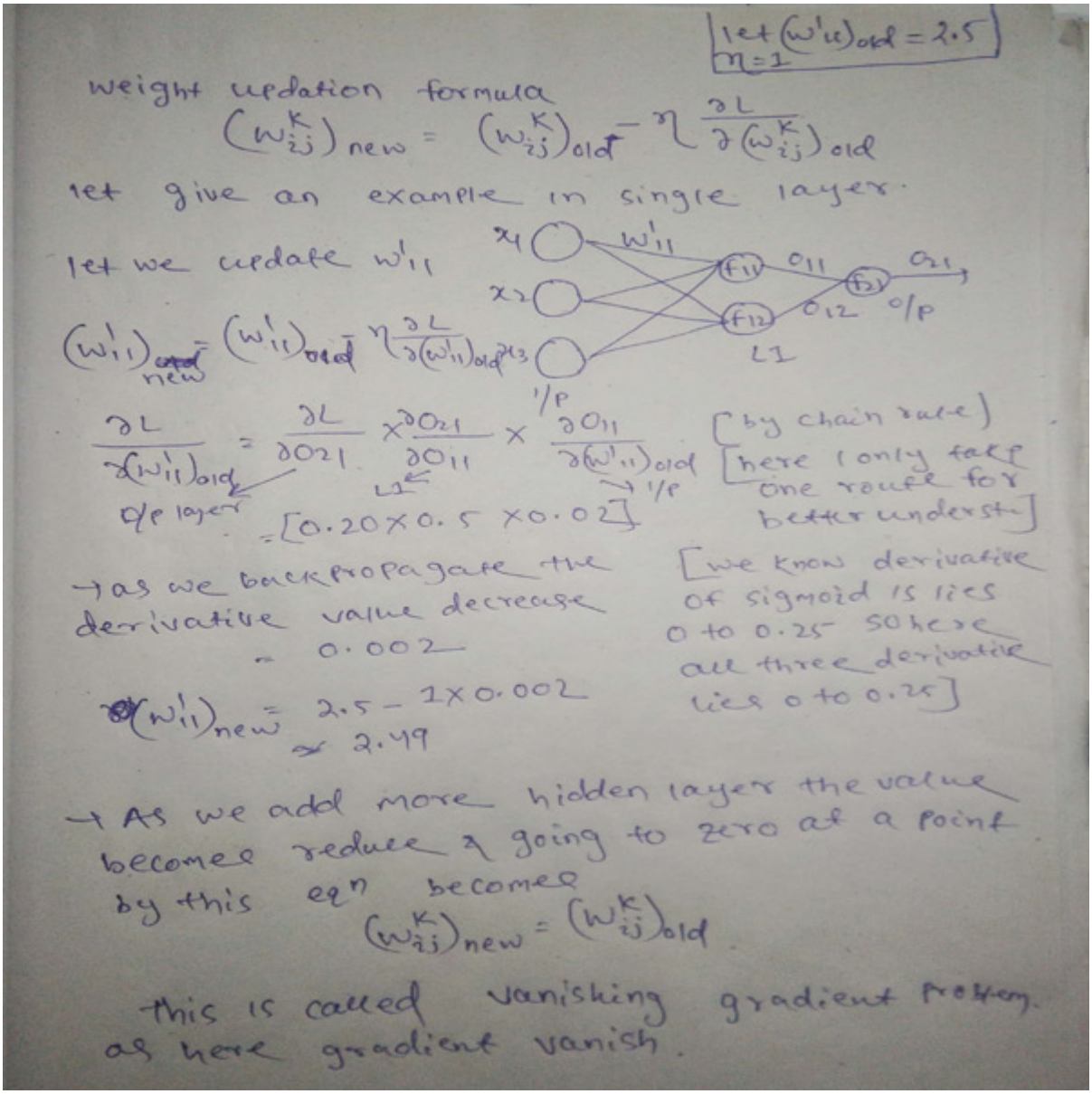
# Vanishing Gradient Problem

- Vanishing gradient problem is a common problem that we face while training deep neural networks.Gradients of neural networks are found during back propagation.
- Generally, adding more hidden layers will make the network able to learn more complex arbitrary functions, and thus do a better job in predicting future outcomes. This is where Deep Learning is making a big difference.

**Now during back-propagation i.e moving backward in the Network and calculating gradients, it tends to get smaller and smaller as we keep on moving backward in the Network . Below is Just a simple demonstration of Vanishing Gradient Problem in single layer.**




- This Happen because of we use sigmoid and tanh activation function in hidden layer. As sigmoid and tanh deriative 0.25,1 respectively. so by calculating number of hidden layer the derivative becomes 0 so avoid it we use RELU activation function in hidden layer.

## Exploding gradient Problem

- We have discussed about vanishing gradient problem.Now we will get in to exploding gradient problem.Earlier we discussed what happens when our gradient becomes very small.Now we will discuss what will happen if it gets large.
- In deep networks or recurrent neural networks, error gradients can accumulate during an update and result in very large gradients.
- These in turn result in large updates to the network weights, and in turn, an unstable network.The explosion occurs through exponential growth by repeatedly multiplying gradients through the network layers that have values larger than 1.0.This will ultimately led to an total unstable network.