## **Zero centered functions**

1. Why do logistic neurons saturate?
   1. Consider a pre activation function:
   2. And the activation function 
   3. In cases where the weights are initialised to very high or very low values, the weighted summation term (a) will become very large or very small (very negative).
   4. This could result in the neuron attaining saturation
   5. Remember to initialise the weights to small values
2. Another shortcoming with the logistic function is that it is not zero centered
   1. Zero centered: The function is spread out equidistant around the 0 point, i.e. it takes an equal number of positive and negative values.
   2. The logistic function ranges from 0 to 1
   3. The tanh function is a zero centered sigmoid function
3. Consider the simple neural network with logistic sigmoid neurons
   1. Consider the following gradients
   2. The bracketed terms are common
   3. Both h21 and h22 are outputs of the logistic function, so they are always positive (i.e. ranging from 0 to 1)
   4. Due to this, at all times, both and will always be of the same sign, either positive or negative. They cannot be different from each other since the bracketed part is common between them and the logistic function output is always positive
   5. The gradients w.r.t all the weights connected to the same neuron are either all +ve or all -ve
   6. Thus, this limits the directions in which the weights can be updated
4. Thus, we cannot arrive at the local minima as fast as possible by moving in all directions.
5. Also, logistic function is computationally expensive because of ex