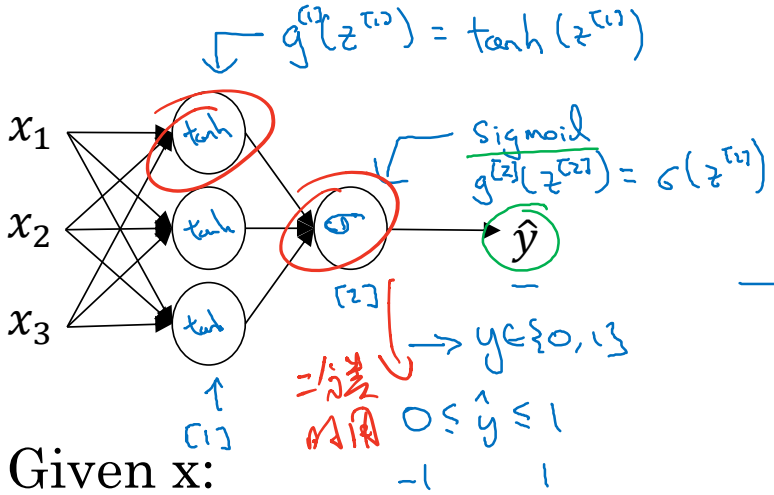
deeplearning.ai

One hidden layer Neural Network

Activation functions

Activation functions



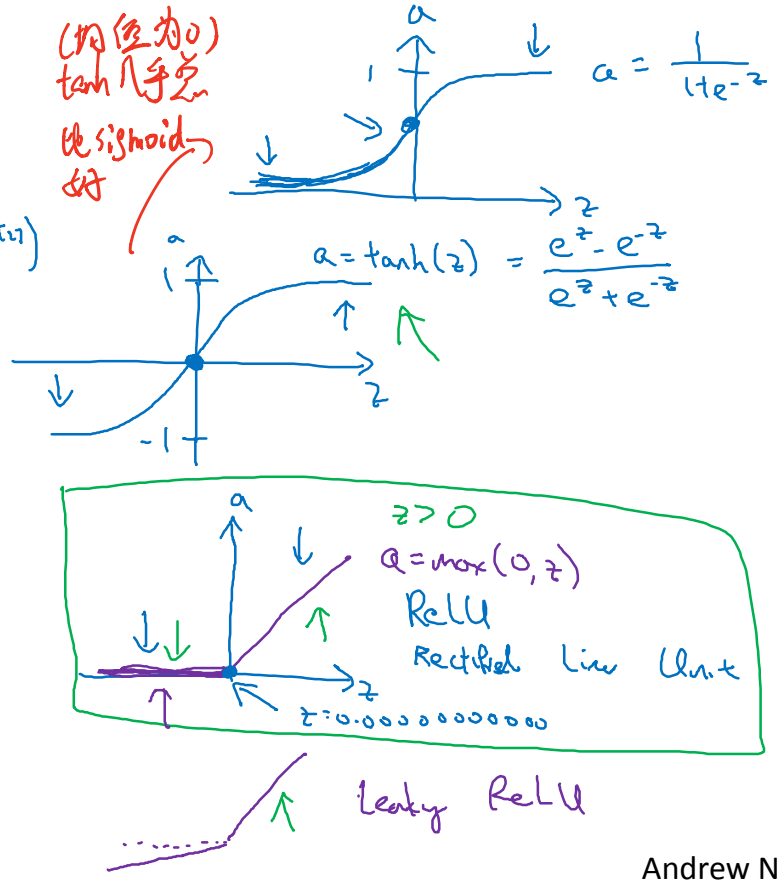
Given x:

$$z^{[1]} = W^{[1]}x + b^{[1]}$$

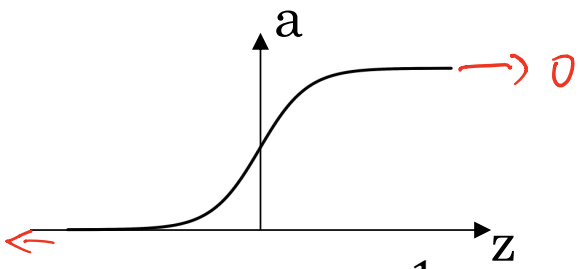
$$a^{[1]} = \sigma(z^{[1]}) \quad g^{[1]}(z^{[1]})$$

$$z^{[2]} = W^{[2]}a^{[1]} + b^{[2]}$$

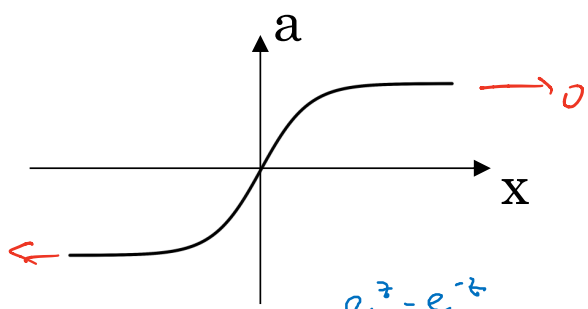
$$a^{[2]} = \sigma(z^{[2]}) \quad g^{[2]}(z^{[2]})$$



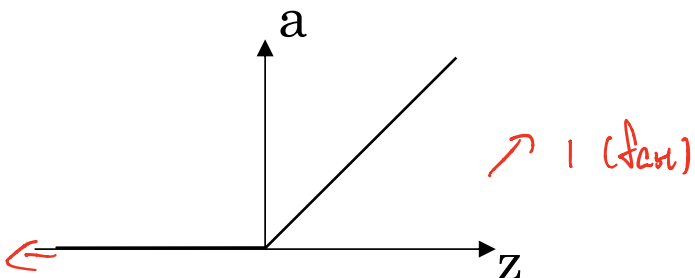
Pros and cons of activation functions



sigmoid: $a = \dfrac{1}{1+e^{-z}}$ 二分类输出层

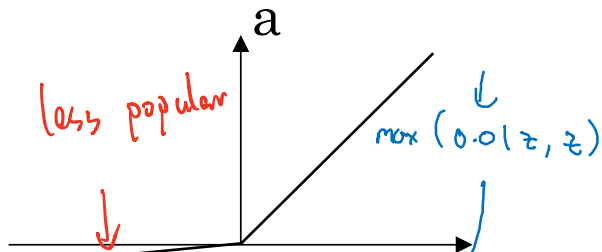


tanh: $a = \dfrac{e^{z}-e^{-z}}{e^{z}+e^{z}}$



1 (slope)

ReLU $a = \max(0, z)$ default



Leaky ReLU $a = \max(0.01z, z)$

Andrew Ng