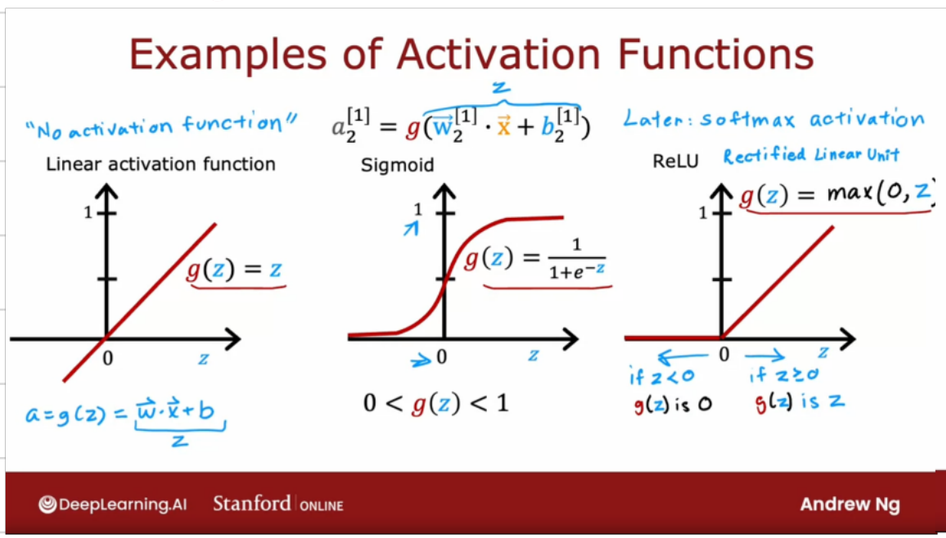
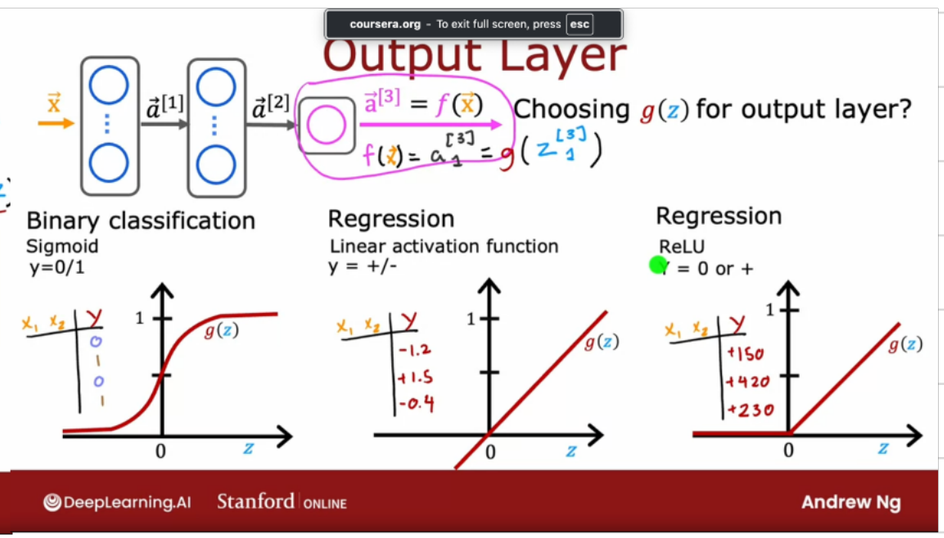
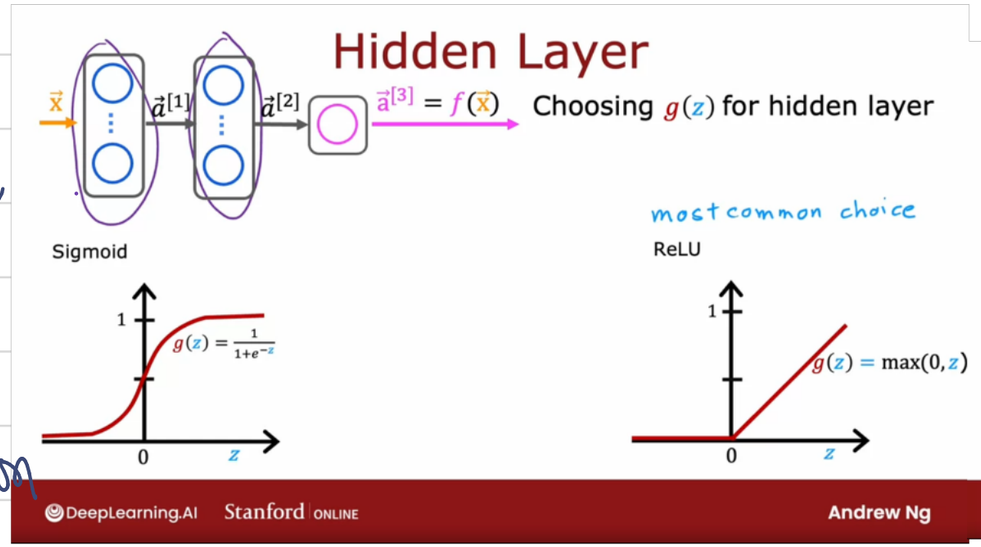
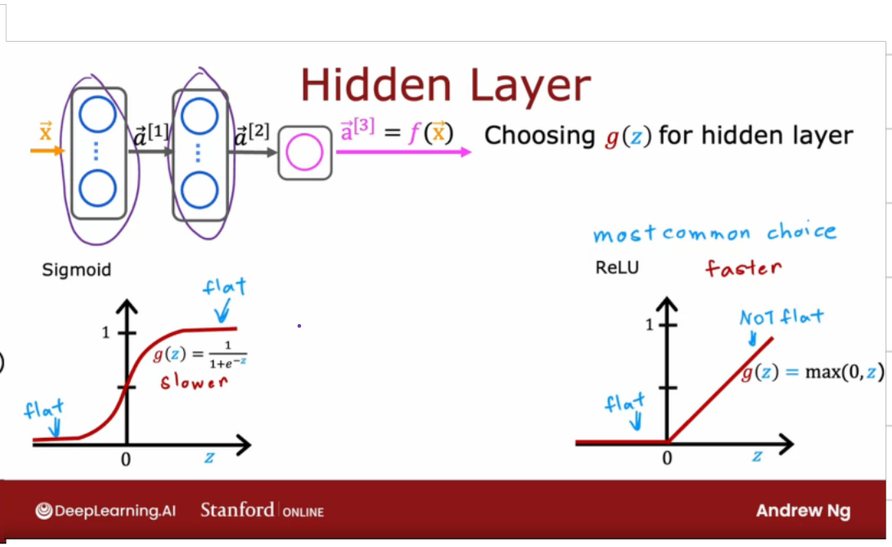
Day-28, oct-30, 2024.





See how $\vec{a}^{[3]} = f(\vec{x})$ applied Activation function





Sigmoid function is only used in binary classification or simpler operation since $[0, 1]$ i.e. $0 < g(z) < 1$

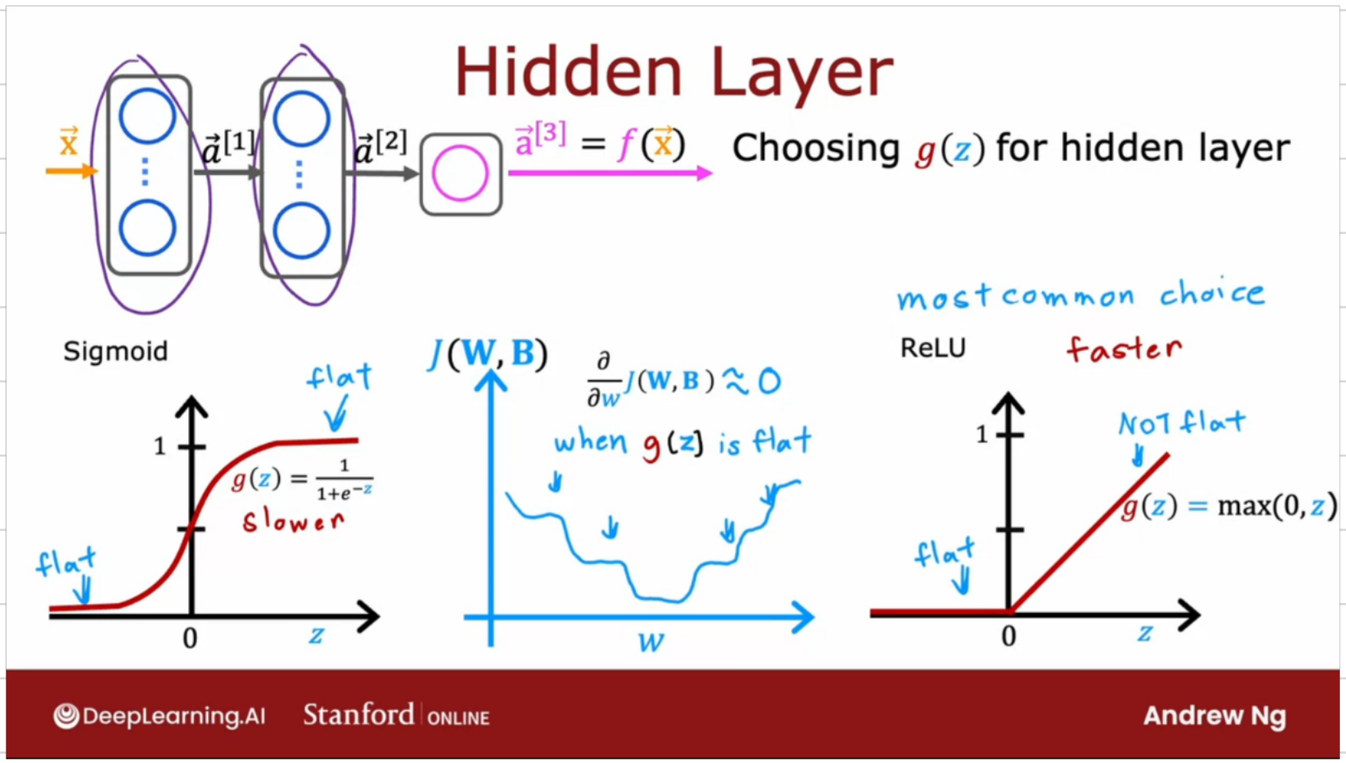
Choose Activation function

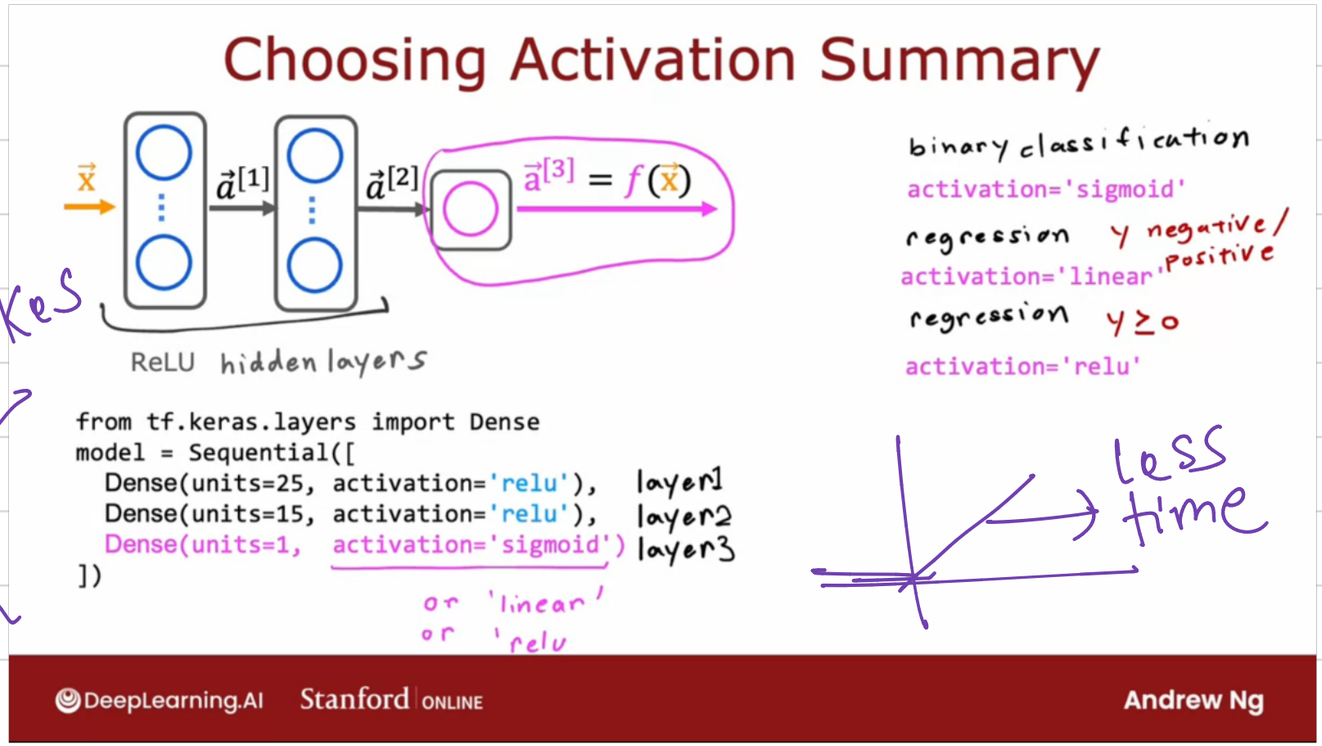
Most widely used is ReLu because it is Simple & faster $[0,\infty)$ than the Sigmoid function

ie. $\max(0,z)$

ie. $f(z) \geq 0$

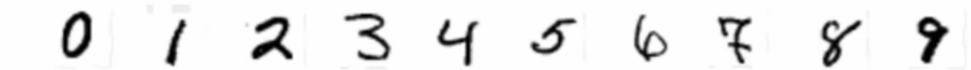
See Sigmoid has flat



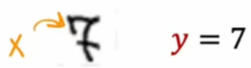


flat takes More Computation

MNIST example



$y = 0$ 1 2 3 4 5 6 7 8 9

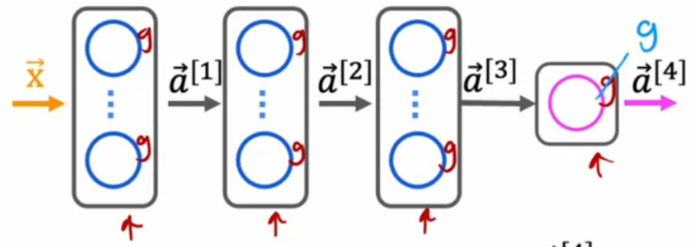


$y = 7$

DeepLearning.AI Stanford ONLINE Andrew Ng

Example

$g(z) = z$



$$\vec{a}^{[4]} = \vec{w}_1^{[4]} \cdot \vec{a}^{[3]} + b_1^{[4]}$$

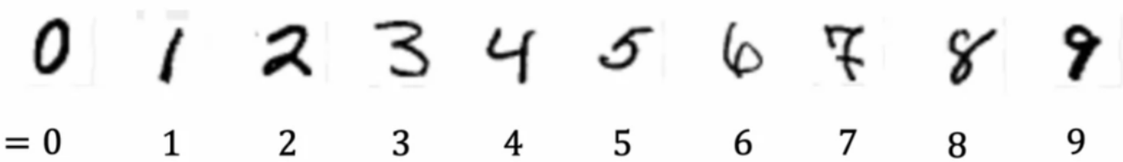
all linear (including output)
↳ equivalent to linear regression

$$\vec{a}^{[4]} = \frac{1}{1+e^{-(\vec{w}_1^{[4]} \cdot \vec{a}^{[3]} + b_1^{[4]})}}$$

output activation is sigmoid
(hidden layers still linear)
↳ equivalent to logistic regression

Multi-class what if I have to classify multiple y? then we use softmax for the multiple class output.

MNIST example



$y = 0$ 1 2 3 4 5 6 7 8 9

$y = 7$

multiclass classification problem:
target y can take on more than two possible values

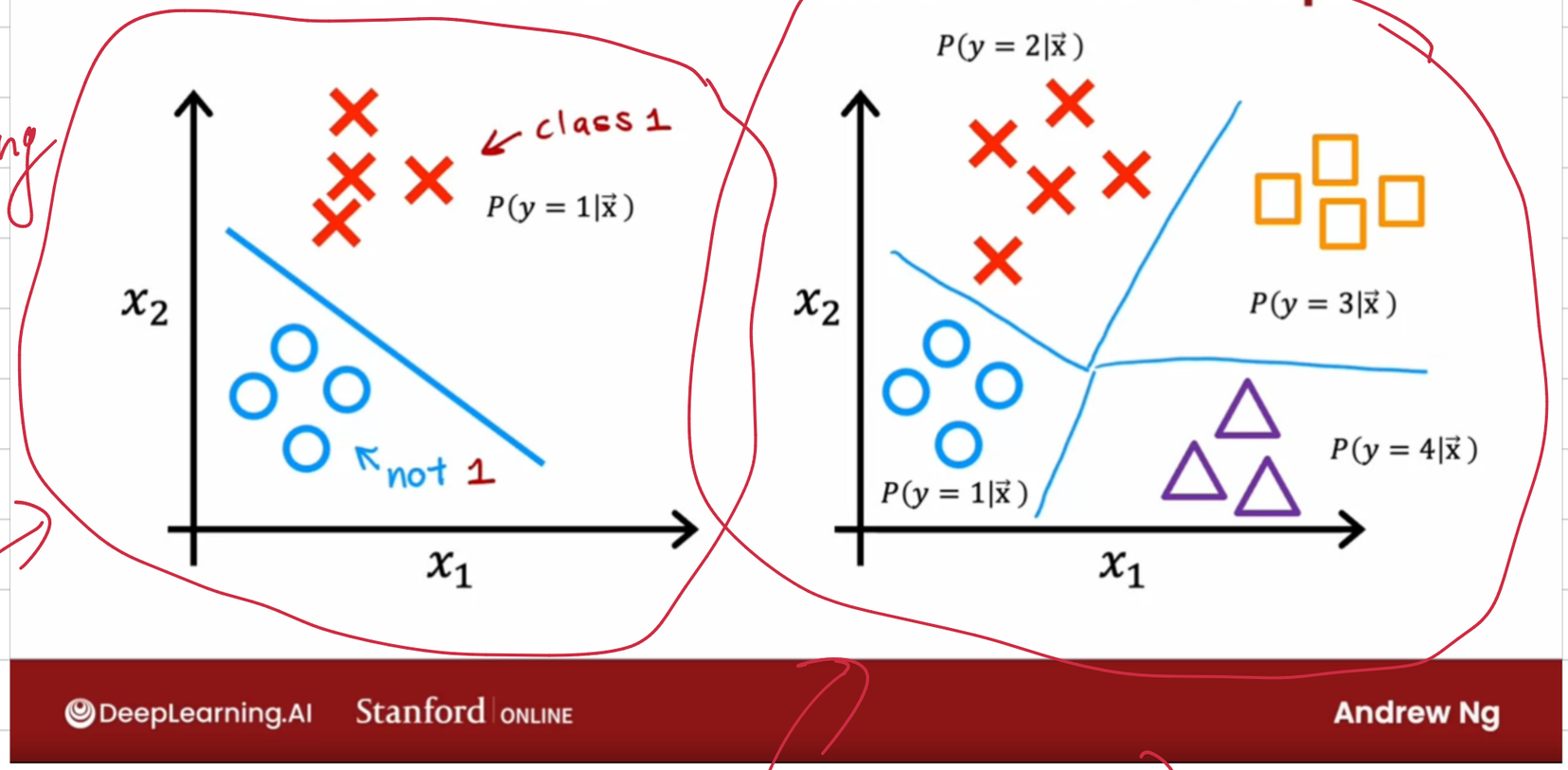
DeepLearning.AI Stanford ONLINE Andrew Ng

→ Activation function

Must be used According to ANN Architecture.

→ ANN architecture has layers of Neural layer

Activation functions are found in Hidden & output layer.



Sigmoid

Softmax (multi-class)

$P = (y = 2|\vec{x})$

$P = (y = 3|\vec{x})$

$P = (y = 2|\vec{x})$