



deeplearning.ai

Deep Neural Networks

Forward and backward
propagation

Forward propagation for layer l



$$z^{(l)} = W^{(l)} \cdot a^{(l-1)} + b^{(l)}$$

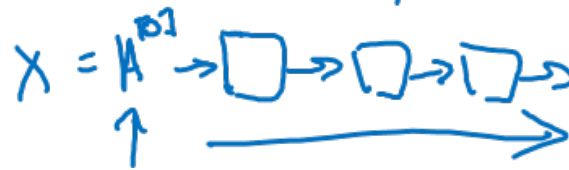
$$a^{(l)} = g^{(l)}(z^{(l)})$$

Vectorized:

$$Z^{(l)} = W^{(l)} \cdot A^{(l-1)} + b^{(l)}$$

$$A^{(l)} = g^{(l)}(Z^{(l)})$$


$a^{(0)}$
 $A^{(0)}$



Backward propagation for layer l

→

→



$$\underline{dz}^{[l]} = \underline{da}^{[l]} * g^{[l]'}(z^{[l]})$$

$$\underline{dw}^{[l]} = \underline{dz}^{[l]} \cdot \underline{a}^{[l-1]}$$

$$\underline{db}^{[l]} = \underline{dz}^{[l]}$$

$$\underline{da}^{[l-1]} = W^{[l]T} \cdot \underline{dz}^{[l]}$$

$$\underline{dz}^{[l-1]} = W^{[l+1]T} \underline{dz}^{[l]} * g^{[l]'}(z^{[l-1]})$$

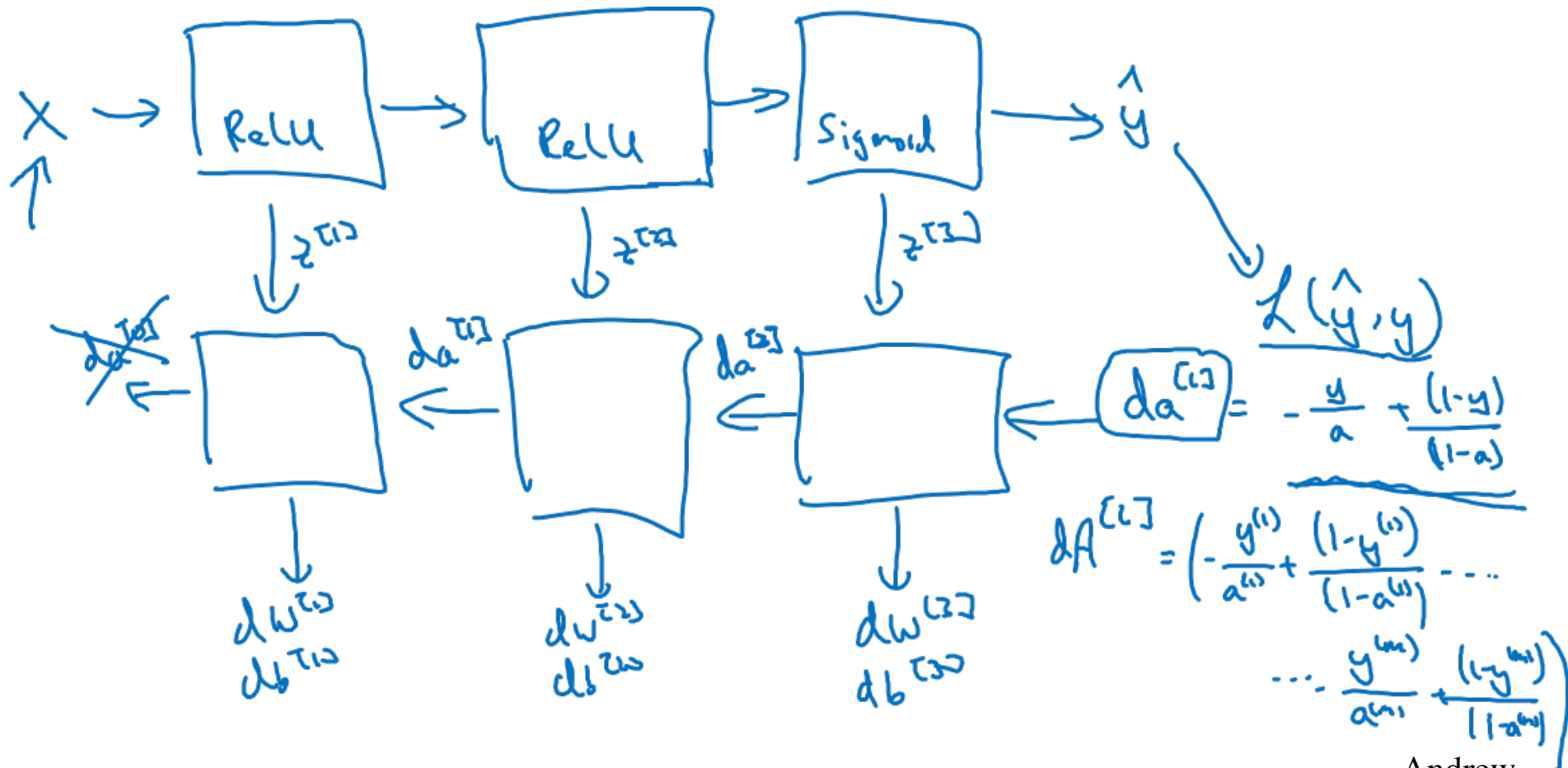
$$\underline{dz}^{[l]} = \underline{dA}^{[l]} * g^{[l]'}(z^{[l]})$$

$$\underline{dw}^{[l]} = \frac{1}{n} \underline{dz}^{[l]} \cdot A^{[l-1]T}$$

$$\underline{db}^{[l]} = \frac{1}{n} \text{np.sum}(\underline{dz}^{[l]}, \text{axis}=1, \text{keepdims}=\text{True})$$

$$\underline{dA}^{[l-1]} = W^{[l]T} \cdot \underline{dz}^{[l]}$$

Summary



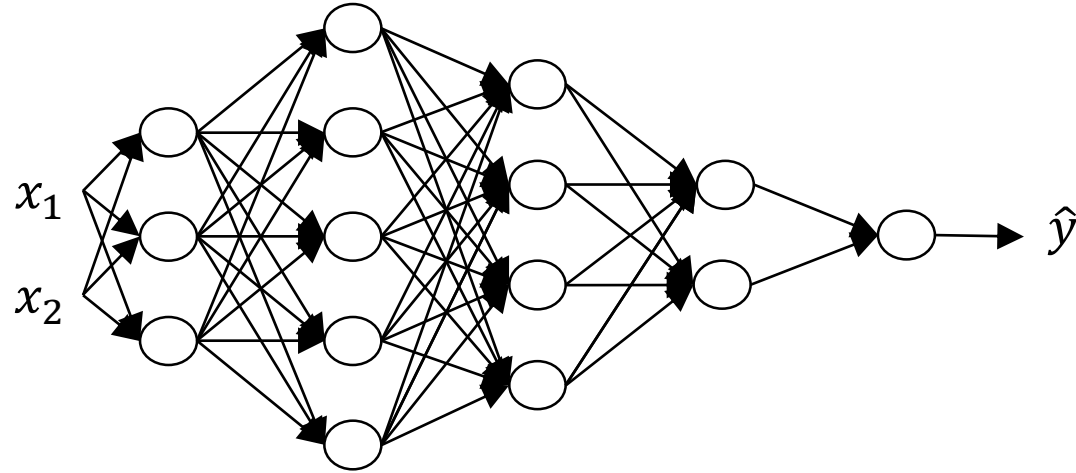


deeplearning.ai

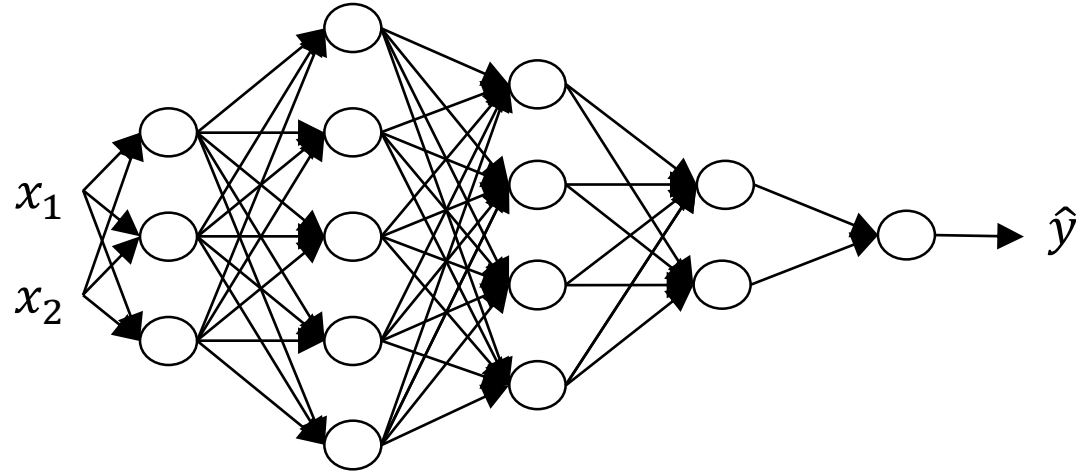
Deep Neural Networks

Getting your matrix
dimensions right

Parameters $W^{[l]}$ and $b^{[l]}$



Vectorized implementation





deeplearning.ai

Deep Neural Networks

Forward and backward
propagation

Forward propagation for layer l



$$z^{(l)} = W^{(l)} \cdot a^{(l-1)} + b^{(l)}$$

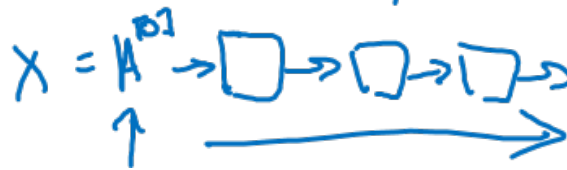
$$a^{(l)} = g^{(l)}(z^{(l)})$$

Vectorized:

$$Z^{(l)} = W^{(l)} \cdot A^{(l-1)} + b^{(l)}$$

$$A^{(l)} = g^{(l)}(Z^{(l)})$$


$a^{(0)}$
 $A^{(0)}$



Backward propagation for layer l

→

→



$$\underline{dz}^{[l]} = \underline{da}^{[l]} * g^{[l]'}(z^{[l]})$$

$$\underline{dw}^{[l]} = \underline{dz}^{[l]} \cdot \underline{a}^{[l-1]}$$

$$\underline{db}^{[l]} = \underline{dz}^{[l]}$$

$$\underline{da}^{[l-1]} = W^{[l]T} \cdot \underline{dz}^{[l]}$$

$$\underline{dz}^{[l-1]} = W^{[l+1]T} \underline{dz}^{[l]} * g^{[l]'}(z^{[l]})$$

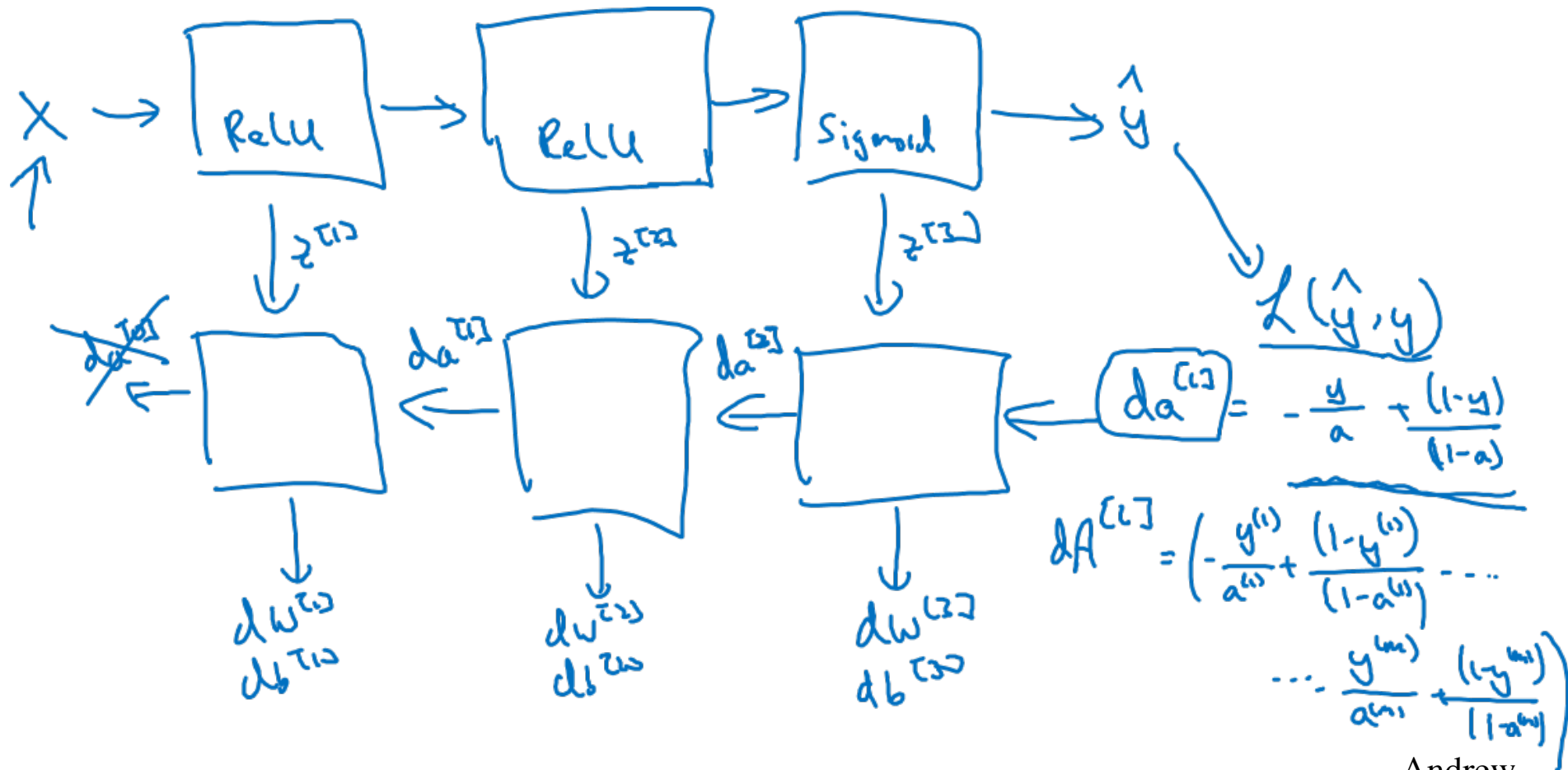
$$\underline{dz}^{[l]} = \underline{dA}^{[l]} * g^{[l]'}(z^{[l]})$$

$$\underline{dw}^{[l]} = \frac{1}{n} \underline{dz}^{[l]} \cdot A^{[l-1]T}$$

$$\underline{db}^{[l]} = \frac{1}{n} \text{np.sum}(\underline{dz}^{[l]}, \text{axis}=1, \text{keepdims}=\text{True})$$

$$\underline{dA}^{[l-1]} = W^{[l]T} \cdot \underline{dz}^{[l]}$$

Summary





deeplearning.ai

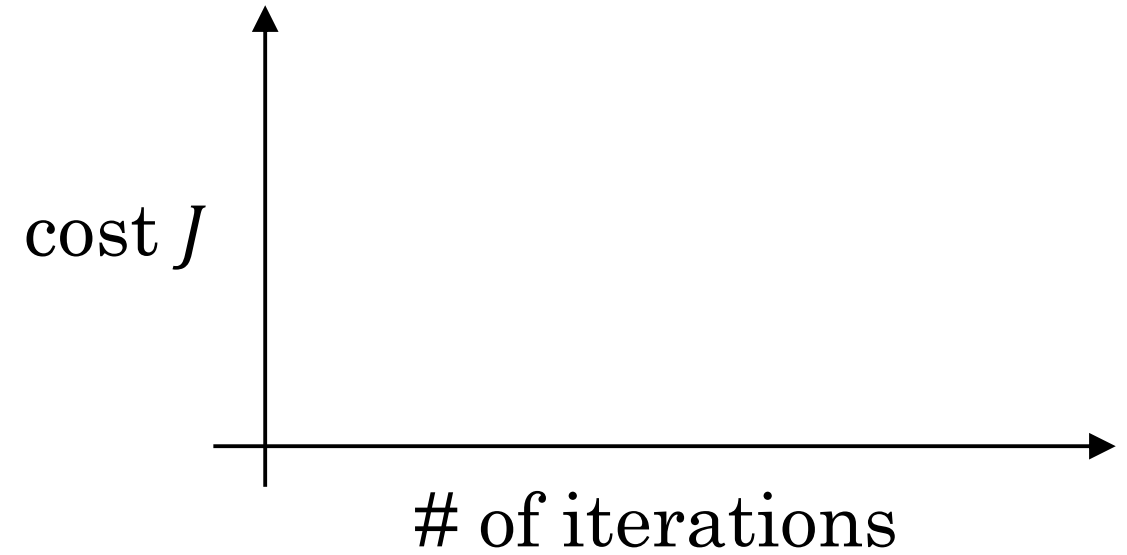
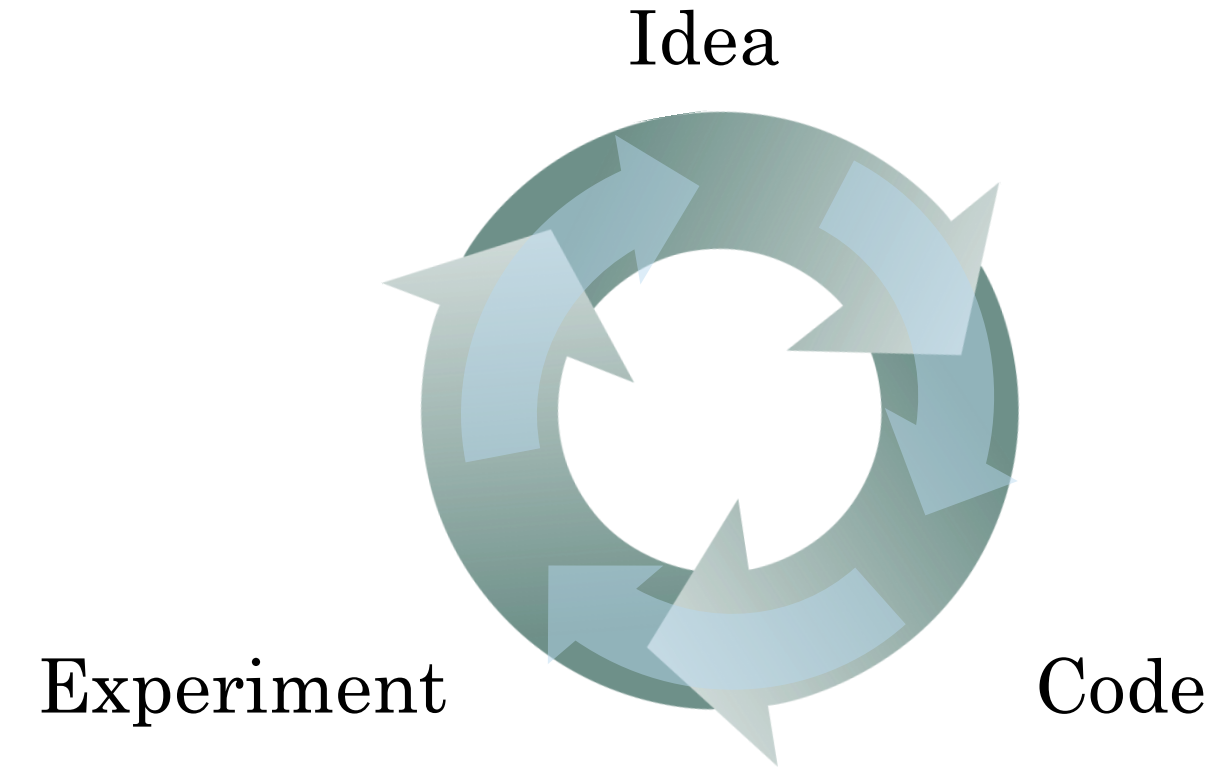
Deep Neural Networks

Parameters vs Hyperparameters

What are hyperparameters?

Parameters: $W^{[1]}$, $b^{[1]}$, $W^{[2]}$, $b^{[2]}$, $W^{[3]}$, $b^{[3]}$...

Applied deep learning is a very empirical process





deeplearning.ai

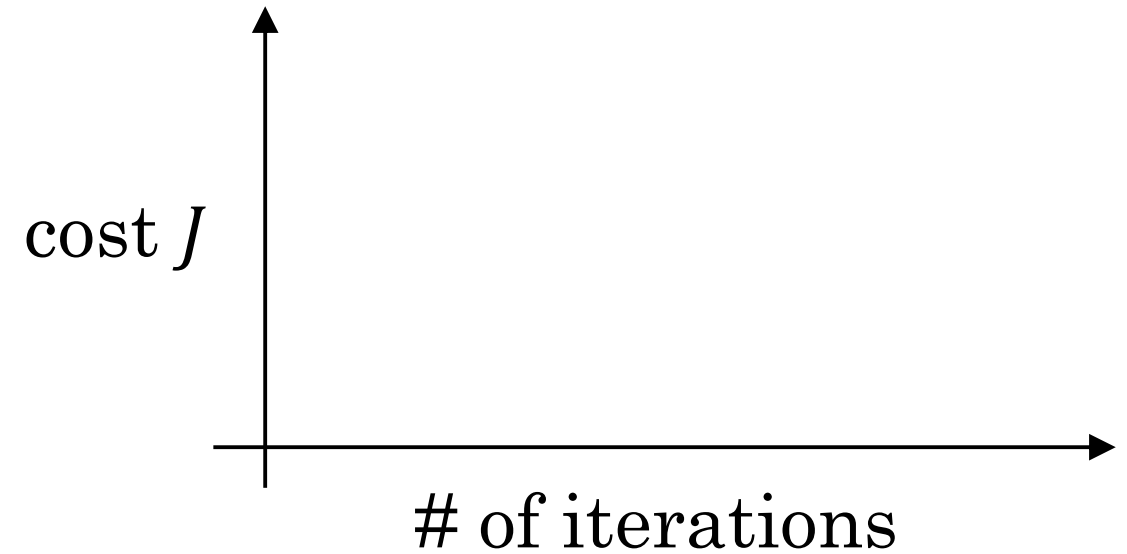
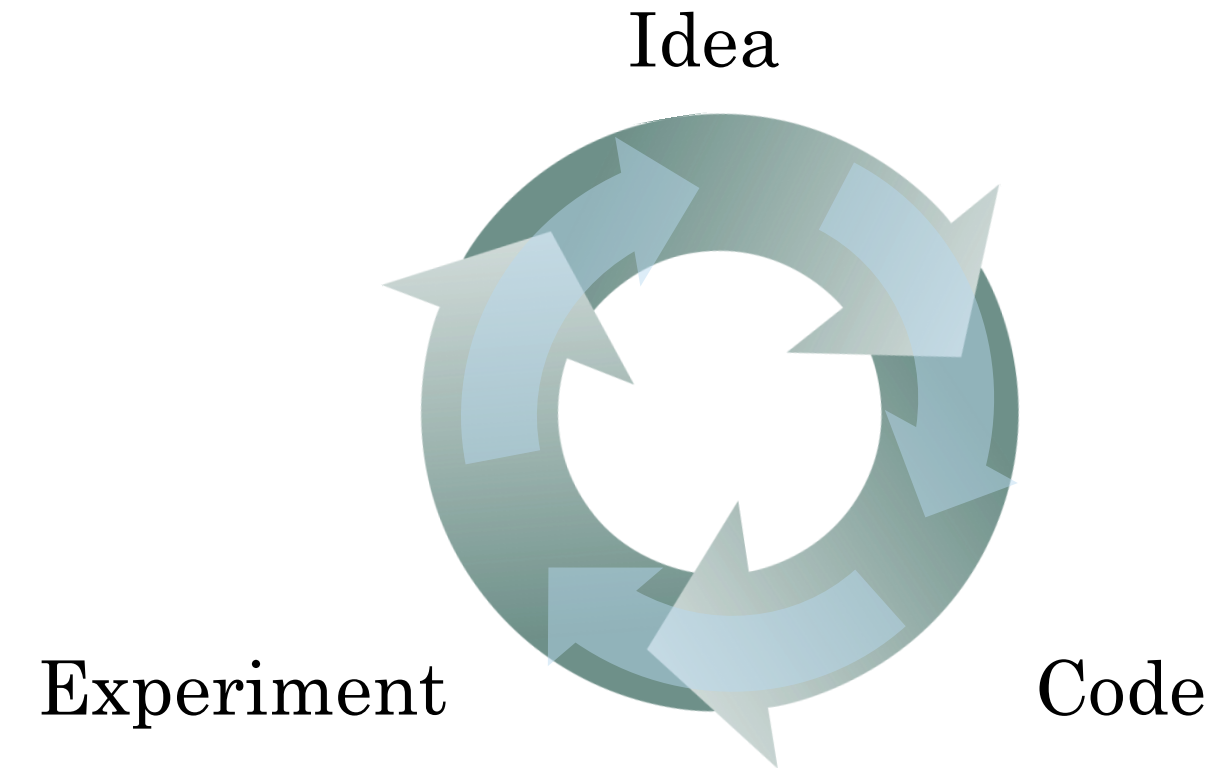
Deep Neural Networks

Parameters vs Hyperparameters

What are hyperparameters?

Parameters: $W^{[1]}$, $b^{[1]}$, $W^{[2]}$, $b^{[2]}$, $W^{[3]}$, $b^{[3]}$...

Applied deep learning is a very empirical process





deeplearning.ai

Deep Neural Networks

What does this
have to do with
the brain?

Forward and backward propagation

$$Z^{[1]} = W^{[1]}X + b^{[1]}$$

$$A^{[1]} = g^{[1]}(Z^{[1]})$$

$$Z^{[2]} = W^{[2]}A^{[1]} + b^{[2]}$$

$$A^{[2]} = g^{[2]}(Z^{[2]})$$

$$\vdots$$

$$A^{[L]} = g^{[L]}(Z^{[L]}) = \hat{Y}$$

$$dZ^{[L]} = A^{[L]} - Y$$

$$dW^{[L]} = \frac{1}{m} dZ^{[L]} A^{[L]T}$$

$$db^{[L]} = \frac{1}{m} \text{np.sum}(dZ^{[L]}, \text{axis} = 1, \text{keepdims} = \text{True})$$

$$dZ^{[L-1]} = dW^{[L]T} dZ^{[L]} g'^{[L]}(Z^{[L-1]})$$

$$\vdots$$

$$dZ^{[1]} = dW^{[L]T} dZ^{[2]} g'^{[1]}(Z^{[1]})$$

$$dW^{[1]} = \frac{1}{m} dZ^{[1]} A^{[1]T}$$

$$db^{[1]} = \frac{1}{m} \text{np.sum}(dZ^{[1]}, \text{axis} = 1, \text{keepdims} = \text{True})$$

