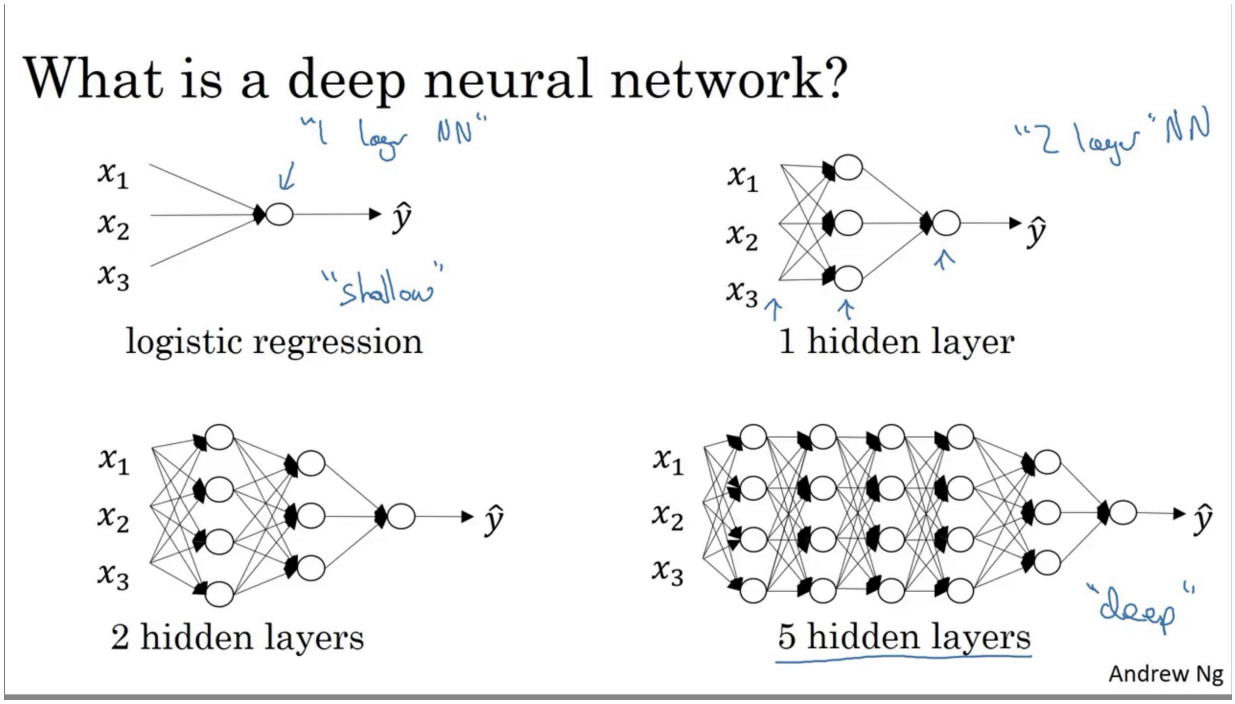
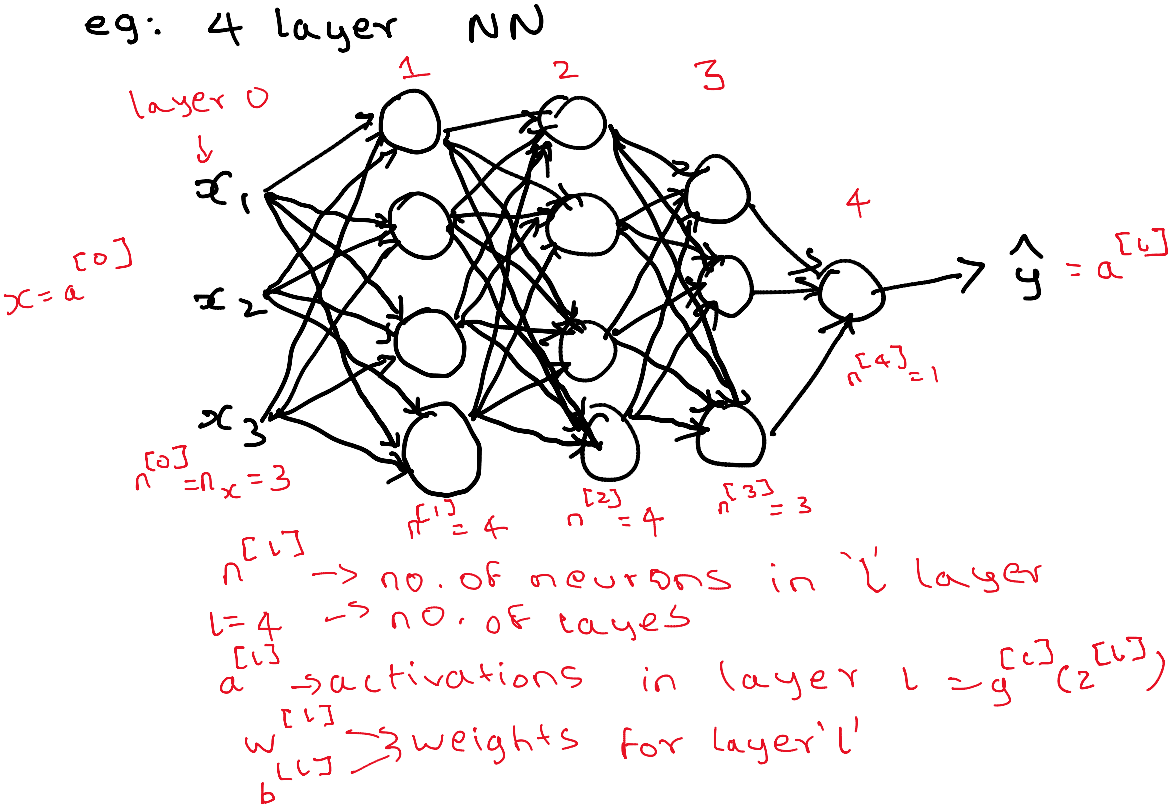
Week 4 - Deep Neural Networks

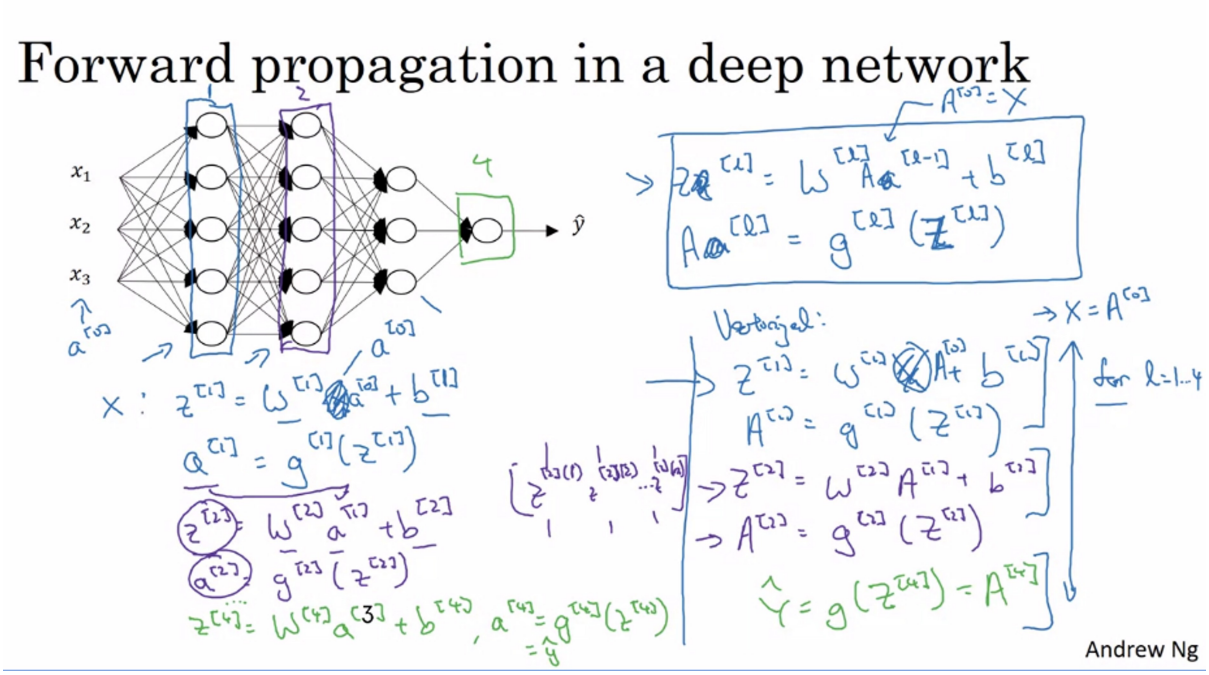
Wednesday, August 12, 2020 1:02 PM

① Deep L-layer Neural Network





② Forward propagation in Deep Neural Network

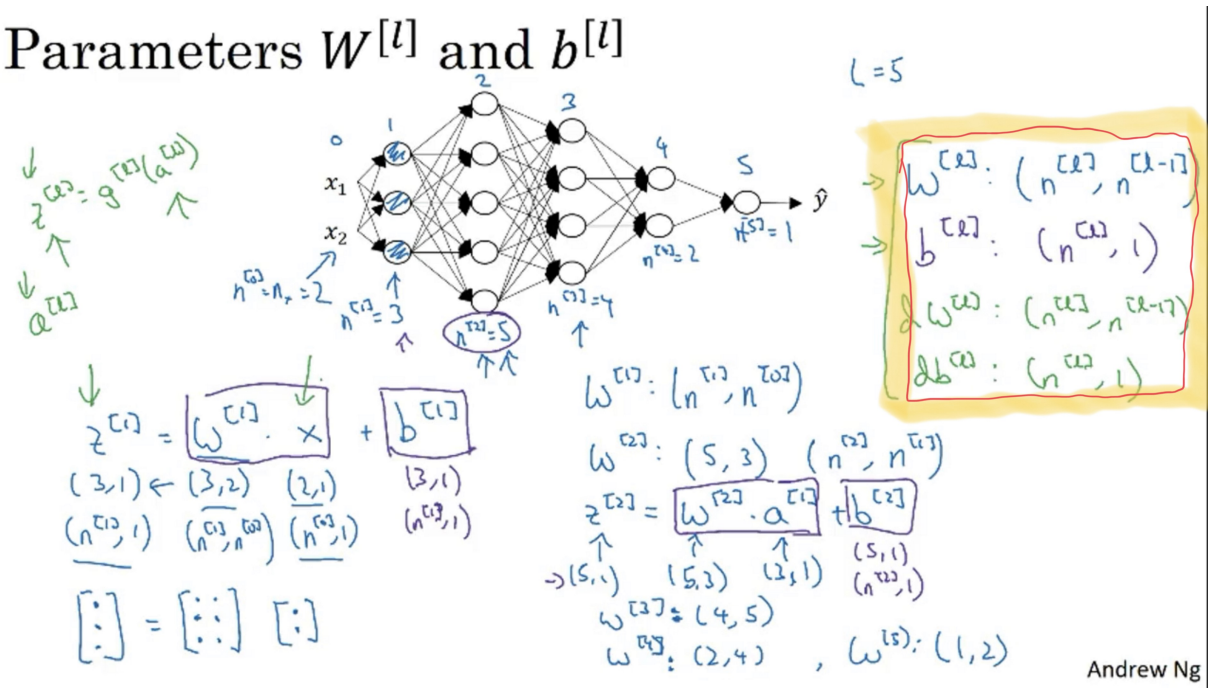


→ General formula

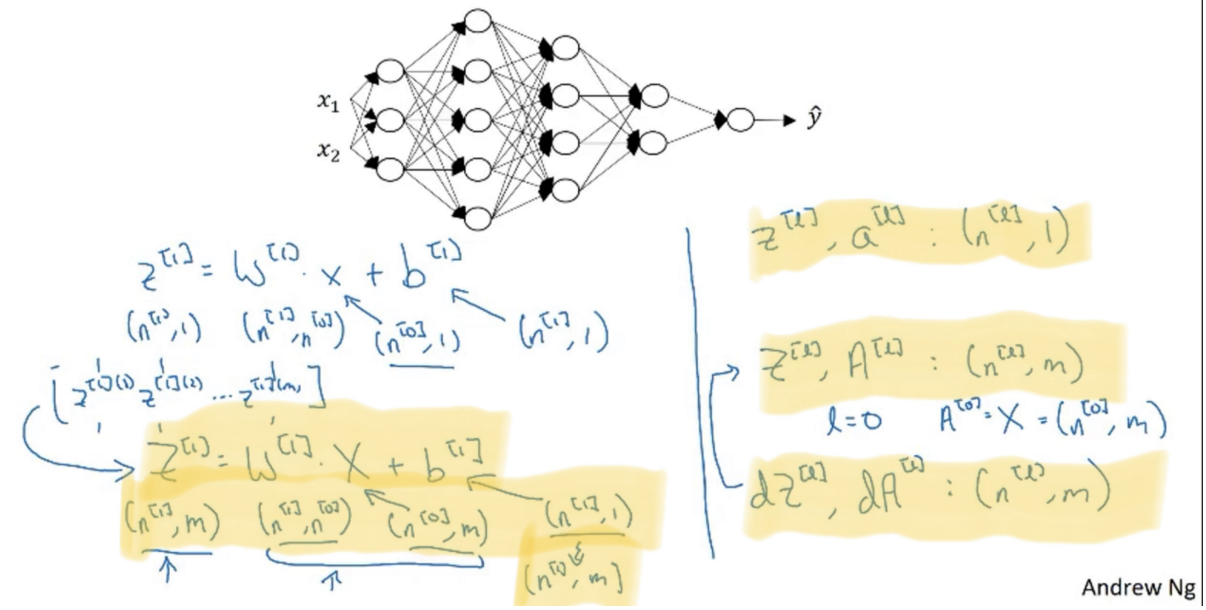
$$Z^{[l]} = W^{[l]} A^{[l-1]} + b^{[l]}$$
$$A^{[l]} = g^{[l]}(Z^{[l]})$$

↓ looped through 1 to L (no. of layers)

(3) Getting the matrix dimensions right



Vectorized implementation



④ Why deep representations

- The earlier layers compute simple things
- The deeper layer compute complex objects.

⑤ Building blocks of deep neural networks

- Forward & backward functions

-> layer l: $w^{[l]}, b^{[l]}$

-> Forward function : Input -> $a^{[l-1]}$

Output -> $a^{[l]}$

$z^{[l]} = w^{[l]} a^{[l-1]} + b^{[l]}$ -> cache $z^{[l]}$ for backprop

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

-> Backward function : Input -> $da^{[l]}$

cache $z^{[l]}$ Output -> $da^{[l-1]}$, $dw^{[l]}$, $db^{[l]}$

1 layer 1 iter

L layers 1 iteration

(6) Forward and Backward propagation

- Forward propagation for layer 'l'
  - i. Input $A^{[l-1]}$
  - ii. Output $A^{[l]}$, cache $(z^{[l]})$ ($w^{[l]}, b^{[l]}$)

$$z^{[l]} = W^{[l]} A^{[l-1]} + b^{[l]}$$

$$\boxed{A^{[l]} = g^{[l]}(z^{[l]})}$$

- Backward propagation for layer 'l'
  - i. Input $dA^{[l]}$

.. [l-1] . [l] .. [l]

i. Input $dA^{[l]}$

ii. Output $da^{[l-1]}, dW^{[l]}, db^{[l]}$

$$dZ^{[l]} = dA^{[l]} * g^{[l]'}(Z^{[l]})$$

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

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

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

$$dZ^{[l]} = \left(W^{[l+1]}\right)^T dZ^{(l+1)} * g^{[l]'}(Z^{[l]})$$

⑦ Parameters vs Hyperparameters

- Parameters: $W^{[l]}, b^{[l]}$
- Hyperparameters: → Learning rate (α)
  - → no. of iterations (epochs)
  - → no. of hidden layers (L)
  - → no. of hidden units ($n^{[l]}$)
  - → activation fn. choice ($g^{[l]}$)
  - → others: Momentum, minibatch size, regularization parameters...
- Deep learning is a very empirical process

idea → code → experiment

experiment code

- The values may change over time

⑧ Deep learning and Human Brain

- While they may look similar they are fundamentally very different
- Analogy is misused and is an oversimplification