Deep Neural Network

Sidharth Baskaran

July 2021

Deep L-layer neural network

- Logistic regression is shallow model, and a deeper network has more hidden layers
- Notation
  - L - number of layers in network
  - $n^{[l]}$ - number of units in layer l
    - $n^{[0]} = n_x$
  - $a^{[l]} = g^{[l]}(z^{[l]})$
    - $a^{[0]} = x, a^{[L]} = \hat{y}$

Forward propagation

- Steps
  - $z^{[l]} = w^{[l]}a^{[l-1]} + b^{[l]}$
  - $a^{[l]} = g^{[l]}(z^{[l]})$
- Vectorized across m examples
  - $Z^{[l]} = W^{[l]}A^{[l-1]} + b^{[l]}$ where $X = A^{[0]}$
  - Z, A, X are stacked columnwise, i.e. $Z^{[1](1)}, \ldots, Z^{[L](m)}$

Matrix Dimension Debugging

- Forward propagation step

$$z^{[l]} = W^{[l]}a^{[l-1]} + b^{[l]}$$
$$(n^{[l]}, 1) = (n^{[l]}, n^{[l-1]})(n^{[l-1]}, 1) + (n^{[l]}, 1)$$

- If vectorized, must modify

$$Z^{[l]} = W^{[l]}A^{[l-1]} + b^{[l]}$$
$$(n^{[l]}, m) = (n^{[l]}, n^{[l-1]})(n^{[l-1]}, m) + \underbrace{(n^{[l]}, 1)}_{\text{broadcasted}}$$