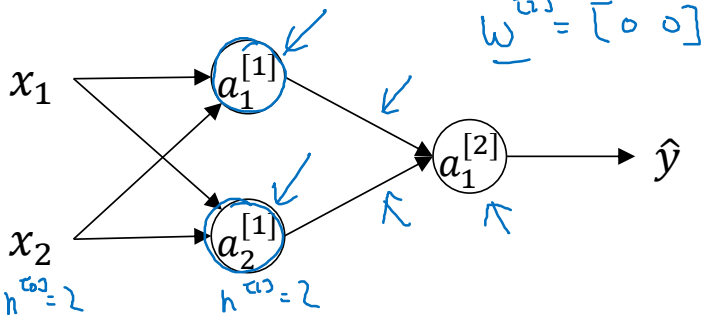
deeplearning.ai

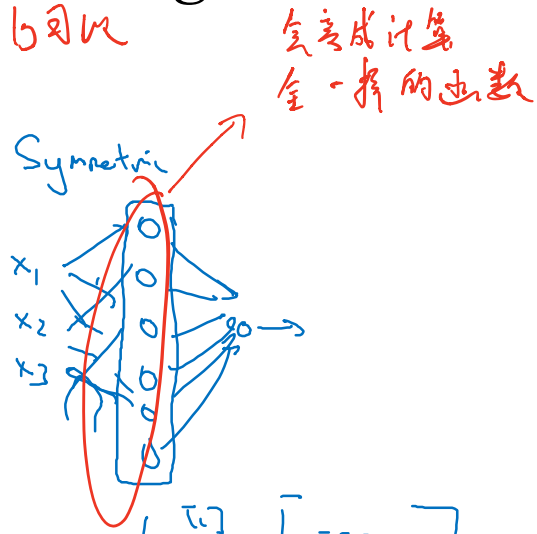
One hidden layer Neural Network

Random Initialization

What happens if you initialize weights to zero?



$w^{[1]} = [0\ 0]$



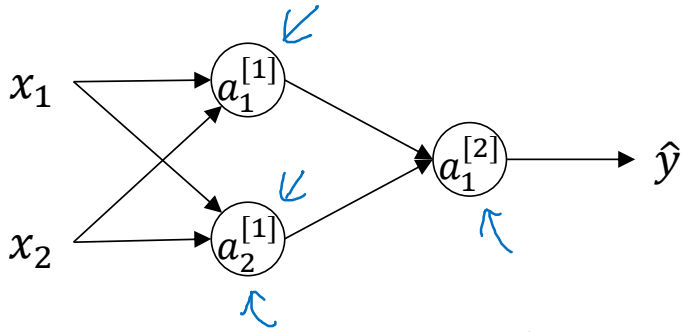
$w^{[1]} = \begin{bmatrix} \cdots \\ \cdots \end{bmatrix}$

$w^{[1]} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ $\quad b^{[1]} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$a_1^{[1]} = a_2^{[1]}$ $\quad dz_1^{[1]} = dz_2^{[1]}$

$dw = \begin{bmatrix} u & v \\ u & v \end{bmatrix}$ $\quad w^{[1]} = w^{[1]} - \alpha\, dw$

Random initialization



$W^{[1]} = np.random.randn((2,2)) * 0.01$

$b^{[1]} = np.zeros((2,1))$

$W^{[2]} = np.random.randn((1,2)) * 0.01$

$b^{[2]} = 0$

如果权重太大，z ↑↓

0.01 / 100 ?

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

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