

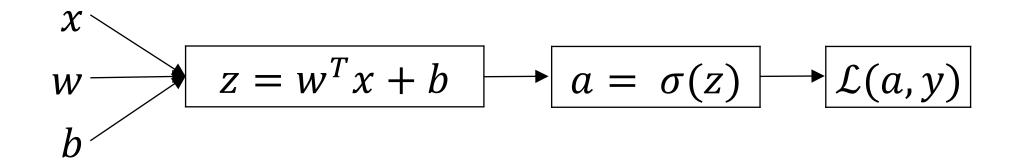
deeplearning.ai

One hidden layer Neural Network

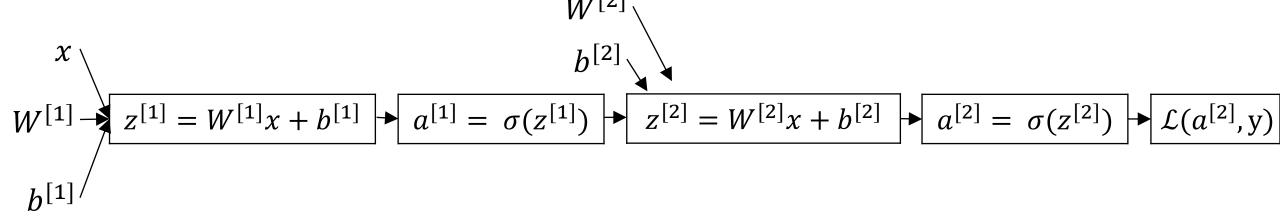
Backpropagation intuition (Optional)

Computing gradients

Logistic regression



Neural network gradients $W^{[2]}$



Summary of gradient descent

$$dz^{[2]} = a^{[2]} - y$$
 $dW^{[2]} = dz^{[2]}a^{[1]^T}$
 $db^{[2]} = dz^{[2]}$
 $dz^{[1]} = W^{[2]T}dz^{[2]} * g^{[1]'}(z^{[1]})$
 $dW^{[1]} = dz^{[1]}x^T$
 $db^{[1]} = dz^{[1]}$

Summary of gradient descent

$$\begin{aligned} dz^{[2]} &= a^{[2]} - y \\ dW^{[2]} &= dz^{[2]}a^{[1]^T} \\ db^{[2]} &= dz^{[2]} \end{aligned} \qquad \begin{aligned} dW^{[2]} &= \frac{1}{m}dZ^{[2]}A^{[1]^T} \\ db^{[2]} &= dz^{[2]} \\ dz^{[1]} &= W^{[2]T}dz^{[2]} * g^{[1]'}(z^{[1]}) \end{aligned} \qquad \begin{aligned} dZ^{[1]} &= W^{[2]T}dZ^{[2]} * g^{[1]'}(Z^{[1]}) \\ dW^{[1]} &= dz^{[1]}x^T \end{aligned} \qquad \begin{aligned} dW^{[1]} &= \frac{1}{m}dZ^{[1]}X^T \\ db^{[1]} &= dz^{[1]} \end{aligned} \qquad \end{aligned} \end{aligned}$$

$$\begin{split} dZ^{[2]} &= A^{[2]} - Y \\ dW^{[2]} &= \frac{1}{m} dZ^{[2]} A^{[1]^T} \\ db^{[2]} &= \frac{1}{m} np. \, sum(dZ^{[2]}, axis = 1, keepdims = True) \\ dZ^{[1]} &= W^{[2]T} dZ^{[2]} * g^{[1]'}(Z^{[1]}) \\ dW^{[1]} &= \frac{1}{m} dZ^{[1]} X^T \\ db^{[1]} &= \frac{1}{m} np. \, sum(dZ^{[1]}, axis = 1, keepdims = True) \end{split}$$