

Setting up your optimization problem

Gradient Checking

Gradient check for a neural network

Take $W^{[1]}$, $b^{[1]}$, ..., $W^{[L]}$, $b^{[L]}$ and reshape into a big vector θ . $\mathcal{J}(\omega^{CI}, b^{CI}, \omega^{CI}, b^{CI})^2 = \mathcal{J}(\theta)$

Take $dW^{[1]}$, $db^{[1]}$, ..., $dW^{[L]}$, $db^{[L]}$ and reshape into a big vector $d\theta$.

Is do the gradet of J(0)?

Gradient checking (Grad check)

For each
$$\bar{z}$$
:

 $\Rightarrow \underline{100_{\text{opp}}} \, \bar{z} = \underline{100_{\text{opp}}} \, \bar{z} =$



Setting up your optimization problem

Gradient Checking implementation notes

Gradient checking implementation notes

- Don't use in training – only to debug

- If algorithm fails grad check, look at components to try to identify bug.

- Remember regularization.

- Doesn't work with dropout.

- Run at random initialization; perhaps again after some training.

