

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 . $\mathcal{J}(\omega^{CO}, b^{CO}, \omega^{CO})^2 = \mathcal{J}(\theta)$

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
$$\bar{c}$$
:

 $\Rightarrow \underline{AOCiJ} = \underline{J(O_1,O_2,...,O_i-E_1,...)} - \underline{J(O_1,O_2,...,O_i-E_1,...)}$
 $\Rightarrow \underline{AOCiJ} = \underline{JJ}$

Check

 $||AO_{apper} - AoI||_2$
 $\Rightarrow ||AO_{apper} - AoI||_2$



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.

