deeplearning.ai

Setting up your optimization problem

Numerical approximation of gradients

Checking your derivative computation

J

$f(\theta) = \theta^3$

$\theta \in \mathbb{R}.$

$g(\theta) = \frac{d}{d\theta} f(\theta) = f'(\theta)$

dw

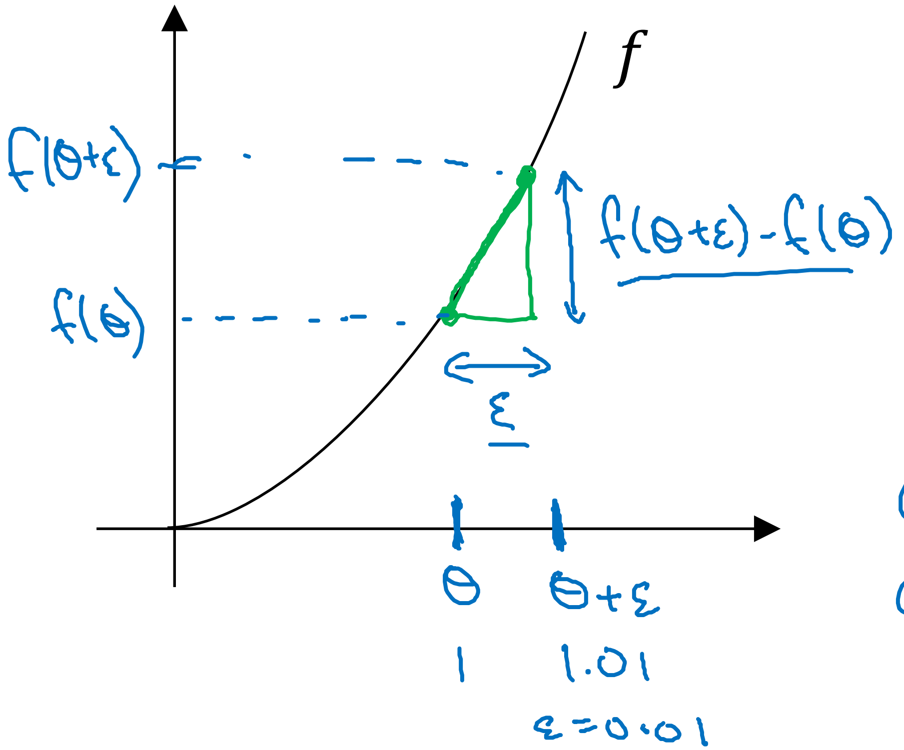
db

$g(\theta) = 3\theta^2.$

$g(\theta) = 3\cdot(1)^2 = 3$

when $\theta = 1$

$$\boxed{\frac{f(\theta+\varepsilon) - f(\theta)}{\varepsilon} \approx g(\theta)}$$



$$\frac{(1.01)^3 - 1^3}{0.01} = 3.0301 \approx 3$$

$\theta = 1$

$\theta + \varepsilon = 1.01$

0.0301

3.1

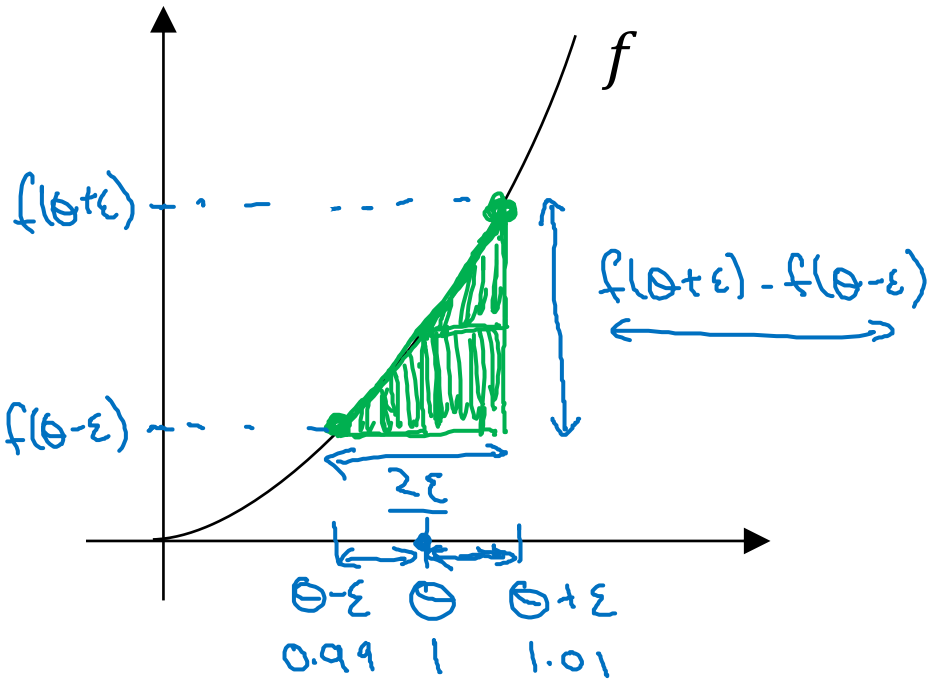
3.2

Andrew Ng

Checking your derivative computation

$f(\theta) = \theta^3$

Using two sided difference way of approximating the derivative



$\varepsilon = 0.01$

$$\frac{f(\theta+\varepsilon) - f(\theta-\varepsilon)}{2\varepsilon} \approx g(\theta)$$

$$\frac{(1.01)^3 - (0.99)^3}{2(0.01)} = 3.0001 \approx 3$$

$$g(\theta) = 3\theta^2 = 3$$

approx error: 0.0001

(prev slide: 3.0301. error: 0.03)

Two sided difference formula is much more accurate.

$$f'(\theta) = \lim_{\varepsilon \to 0} \frac{f(\theta+\varepsilon) - f(\theta-\varepsilon)}{2\varepsilon} \qquad O(\varepsilon^2)$$

0.01

0.0001

$$\frac{f(\theta+\varepsilon) - f(\theta)}{\varepsilon} \qquad \text{error}: O(\varepsilon)$$

0.01

Andrew Ng