



deeplearning.ai

Setting up your
optimization problem

Numerical approximation
of gradients

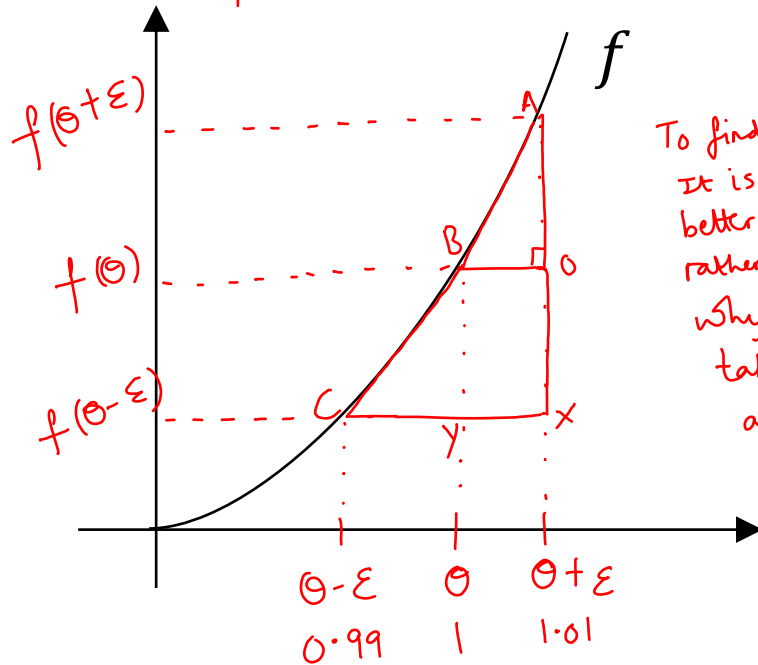
Checking your derivative computation

gradient checking \rightarrow method to check if all your back prop gradients are correct

Conclusion

we should use the 2 sided formula (Bigger Δ), not the smaller Δ

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



To find gradient at θ
It is found to be better to take ΔAXC rather than ΔAOB
Why? Because you're taking both ΔCBY and ΔABO into account

Bigger Δ

$$\frac{f(\theta + \epsilon) - f(\theta - \epsilon)}{\theta + \epsilon - (\theta - \epsilon)} = g(\theta) = \frac{(1.01)^3 - (0.99)^3}{2 \times (0.01)} = \underline{3.0001}$$

The actual derivative for $f(\theta) = \theta^3$ is $3\theta^2$
 $= 3(1)^2 = 3$

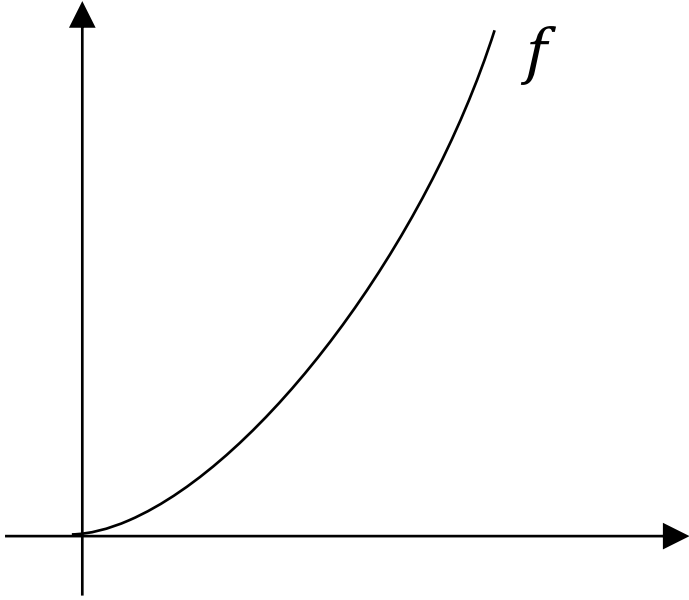
Smaller Δ

$$\frac{f(\theta + \epsilon) - f(\theta)}{\theta + \epsilon - \theta} = \frac{(1.01)^3 - (1)^3}{1.01 - 1} = \underline{3.0301}$$

\Rightarrow error with bigger $\Delta <$ error with smaller Δ

\Rightarrow check both sides of θ

Checking your derivative computation



~~PTO~~