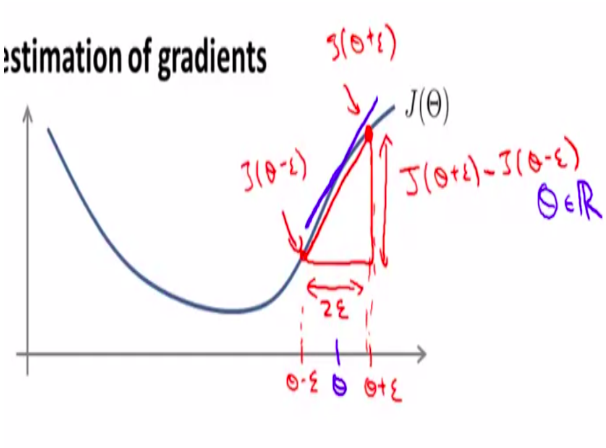
# 1 Derivative's numerical approximation

Let's consider the following function's graph:




Assuming the function $J(x)$ is differentiable at point $\theta$, for small enough $\epsilon$, the expression $\frac{J(x+\epsilon)-J(x-\epsilon)}{2*\epsilon}$ is good approximation of $\frac{dJ}{dx}|_{x=\theta}$

# 2 Partial derivative numerical approximation

The approximation in the previous section can be extended to multi-variable fuctions. For $\theta \in \mathbb{R}^n$ and $f(\theta)$ we can approximate partial derivatives as follows:

$$\frac{\delta}{\delta\theta_j}J(\theta) = \frac{J(\theta_1, ..., \theta_j + \epsilon ..., \theta_n) - J(\theta_1, ..., \theta_j - \epsilon ..., \theta_n)}{2 \cdot \epsilon}$$

# 3 Usage

When using gradient descent, or any other advanced numerical optimization algorithms, one way to verify the correctness of the implementation is to have small differences between the numerical approximation of the gradient and the computed gradient.