

# Gradient Descent 경사하강법

cost function value to minimum

by taking the derivative of cost function.

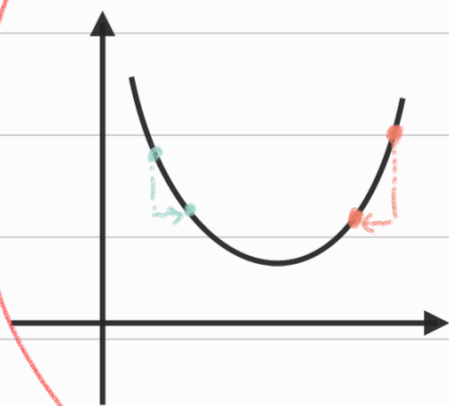
Slope of the tangent is the derivative at that point.

repeat until convergence:

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$

where  $j=0,1$  represents the feature idx number.

(수렴할 때까지 할당 반복, 가울기 0이면 수렴)



미분값 (가울기) 이 positive 이면

$$\theta_1 := \theta_1 - \alpha \cdot (+)$$

미분값 이 negative 이면

$$\theta_1 := \theta_1 - \alpha \cdot (-)$$

따라서,

미분값 이 '0' 일때

$$\theta_1 := \theta_1 - \alpha \cdot 0$$

Converge!

learning rate

- 너무 작으면 학습이 느리고

- 너무 크면 converge 실패

- 수렴할수록 가울기가 줄기 때문에

$\alpha$  줄일 필요 X