

Gradient descent algorithm

Repeat until convergence

$$\begin{cases} \underline{w} = w - \alpha \frac{\partial}{\partial w} J(w, b) \\ \underline{b} = b - \alpha \frac{\partial}{\partial b} J(w, b) \end{cases}$$

Learning rate
Derivative

Simultaneously
update w and b

Assignment

$$a = c$$

$$a = a + 1$$

Code

Truth assertion

$$a = c$$

$$a = a + 1$$

Math

$$a == c$$

Correct: Simultaneous update

$$tmp_w = w - \alpha \frac{\partial}{\partial w} J(w, b)$$

$$tmp_b = b - \alpha \frac{\partial}{\partial b} J(w, b)$$

$$w = tmp_w$$

$$b = tmp_b$$

Incorrect

$$tmp_w = w - \alpha \frac{\partial}{\partial w} J(w, b)$$

$$\underline{w} = tmp_w$$

$$tmp_b = b - \alpha \frac{\partial}{\partial b} J(w, b)$$

$$\underline{b} = tmp_b$$

Stanford ONLINE

DeepLearning.AI

Andrew Ng

Formula:

$$w = w - \alpha \frac{\partial}{\partial w} J(w, b)$$

$$b = b - \alpha \frac{\partial}{\partial b} J(w, b)$$

→ Assignment sign.

→ Learning rate

→ Between 0 and 1

→ It controls how big of a step you take downhill.

→ Large value means large step

toward the local minima.

→ Derivative

→ It tells us in which direction we need to take our steps.

→ We need to repeat these two formula untill we reach to local minima.

→ Update w and b simultaneously

