

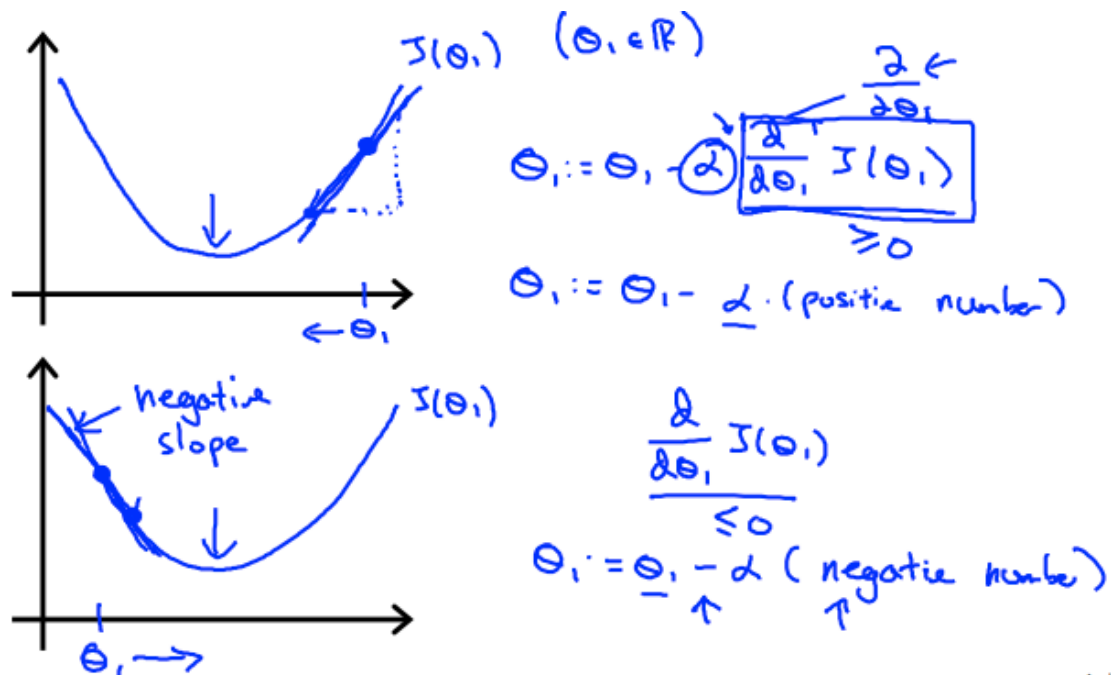
# Gradient Descent Intuition

In this video we explored the scenario where we used one parameter  $\theta_1$  and plotted its cost function to implement a gradient descent. Our formula for a single parameter was :

Repeat until convergence:

$$\theta_1 := \theta_1 - \alpha \frac{d}{d\theta_1} J(\theta_1)$$

Regardless of the slope's sign for  $\frac{d}{d\theta_1} J(\theta_1)$ ,  $\theta_1$  eventually converges to its minimum value. The following graph shows that when the slope is negative, the value of  $\theta_1$  increases and when it is positive, the value of  $\theta_1$  decreases.



On a side note, we should adjust our parameter  $\alpha$  to ensure that the gradient descent algorithm converges in a reasonable time. Failure to converge or too much time to obtain the minimum value imply that our step size is wrong.