GRADIENT DESCENT

Tech.

Halotech Academy

Training Linear Regression

Linear Regression

W dan **b**

Cost Function

Merupakan Fungsi yang digunakan untuk menghitung tingkat error pada suatu algoritma machine learning linear

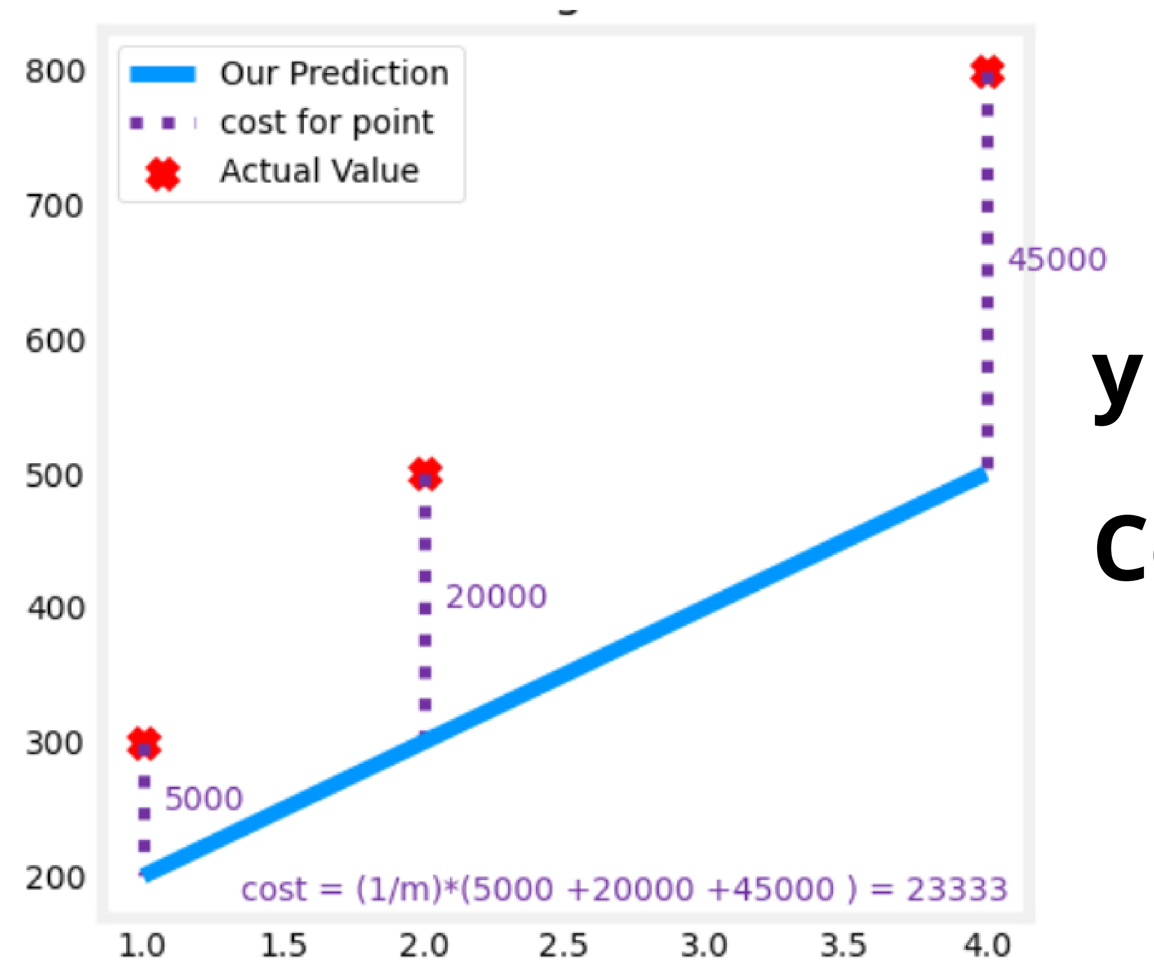
Gradient Descent

Algoritma optimasi yang digunakan untuk menemukan nilai-nilai parameter (koefisien) dari suatu fungsi (f) yang meminimalkan fungsi cost.

Merupakan algoritma yang digunakan untuk memperkecil nilai Cost Function

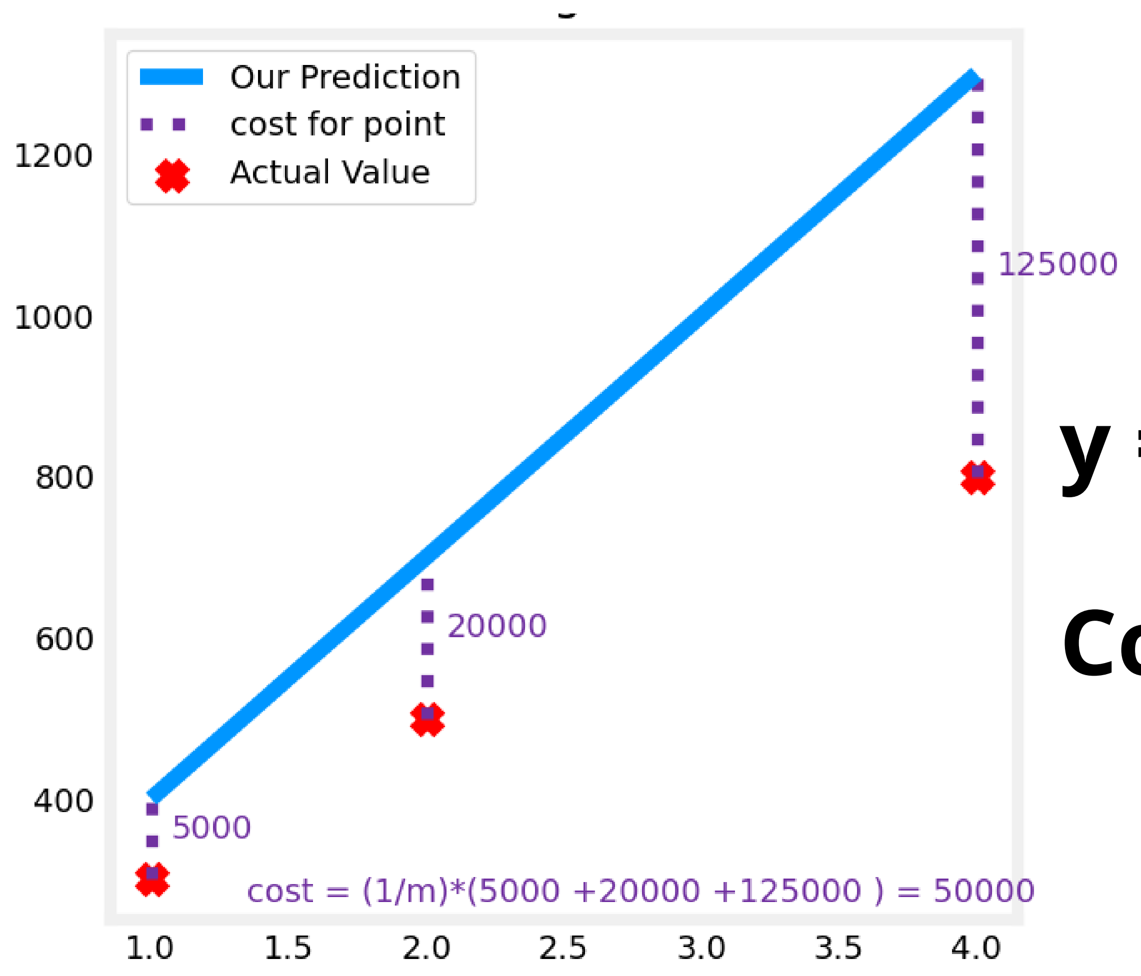
Halotech Academy

Source : Wikipedia



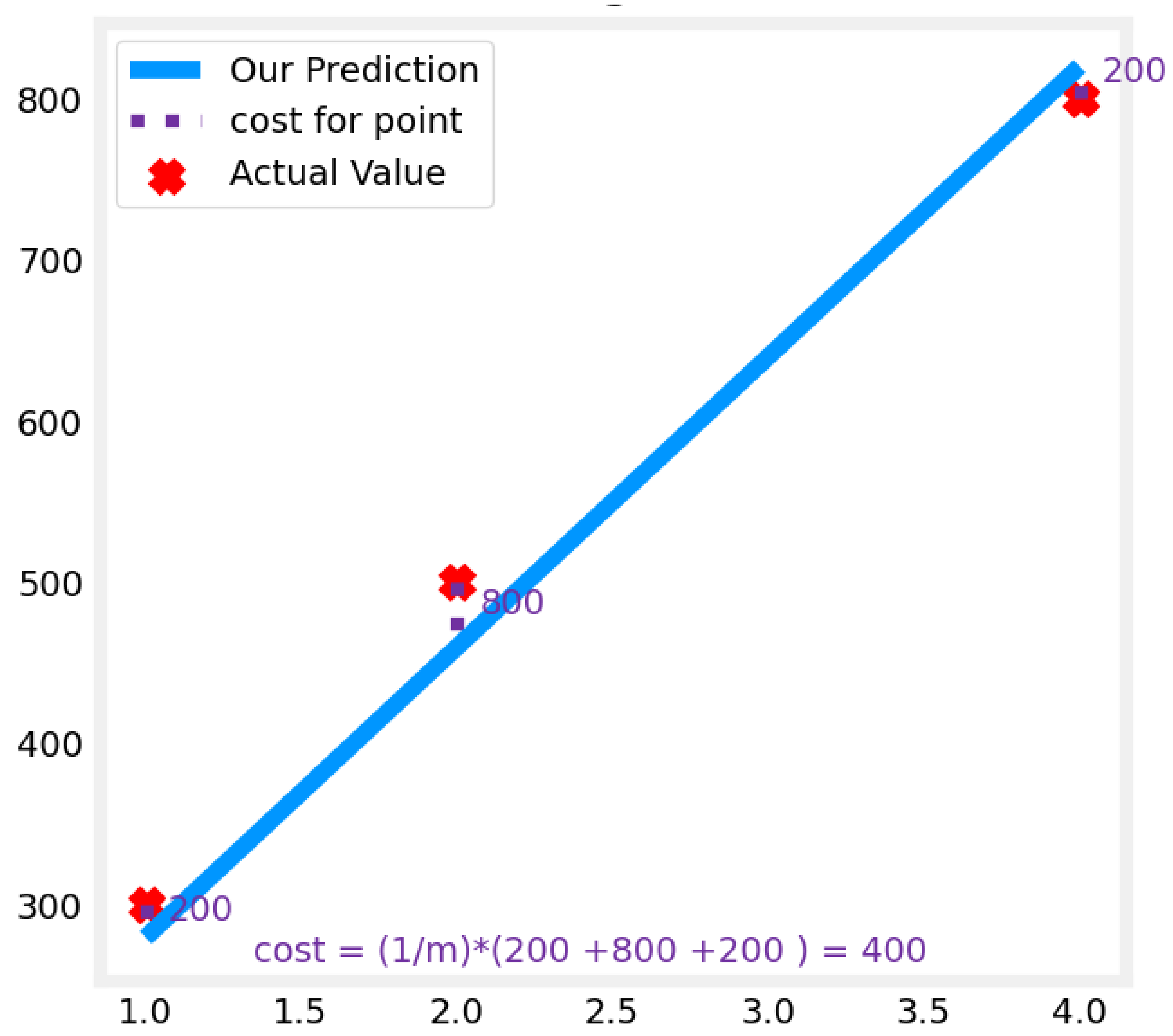
y = 100x + 100

Cost = 23.333



y = 300x + 100

Cost = 50.000



y = 180x + 100 **Cost = 400**

Source Image : Course Supervised Machine Learning: Regression and Classification, Coursera

GRADIENT DESCENT

Repeat Until Convergent {

$$\theta_j := \theta_j + \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) \qquad (for\ j = 0\ and\ j = 1)$$

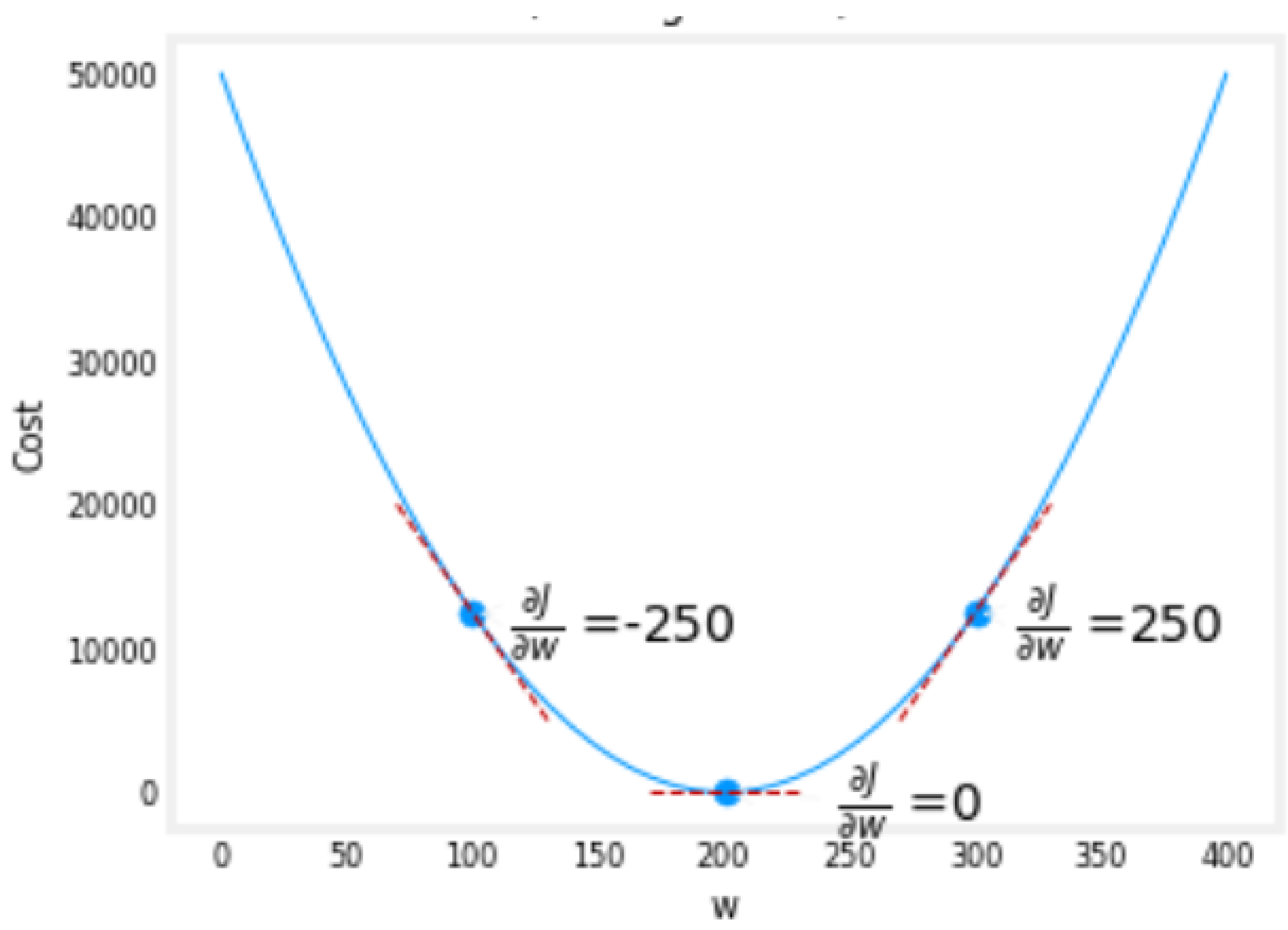
)

θ_j	**Parameter** ,Nilai yang di optimasi)
$:=$	Assigment Operator, Operator menggantikan value , (a:=b , ini bermasud b menggantikan value a)
α	**Learning Rate** **,**Mengkontrol seberapa besar step yang dilakukan.
$\frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$	**Nilai Turunan / Derivative Term** , yang memberi arah menuju minimum/optimal

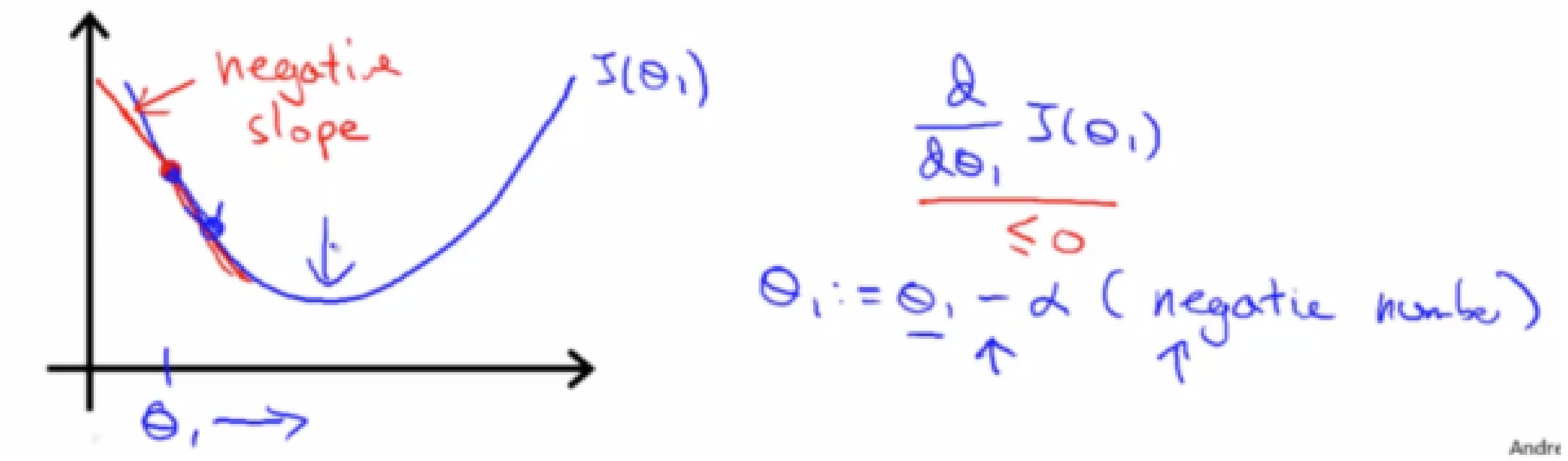
Halotech
Academy

NILAI TURUNAN / DERIVATIVE TERM

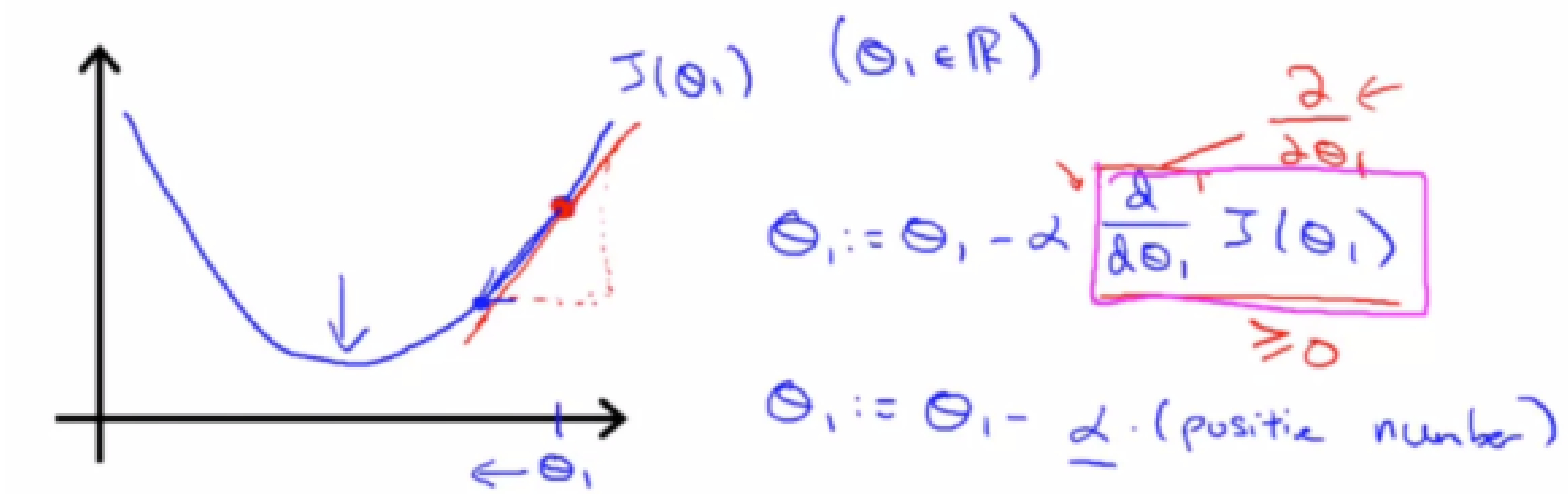
$$w := w - \alpha \frac{\partial J(w,b)}{\partial w}$$



Negative Slope



Positive Slope

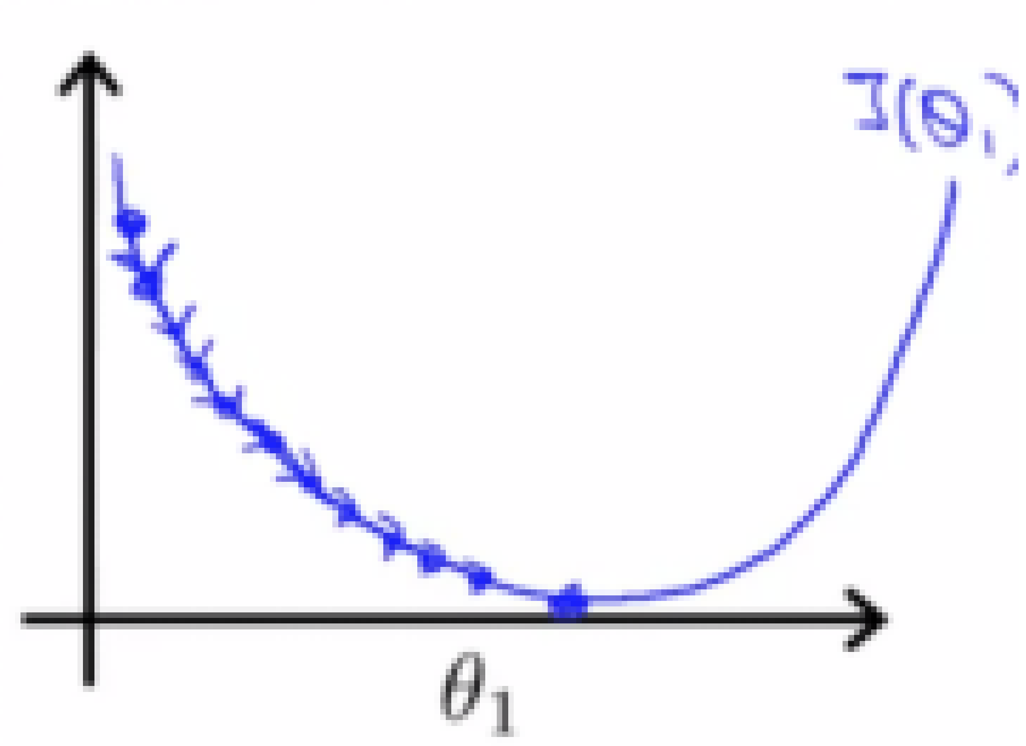


Halotech Academy
Tech.

Source Image : Course Supervised Machine Learning: Regression and Classification, Coursera

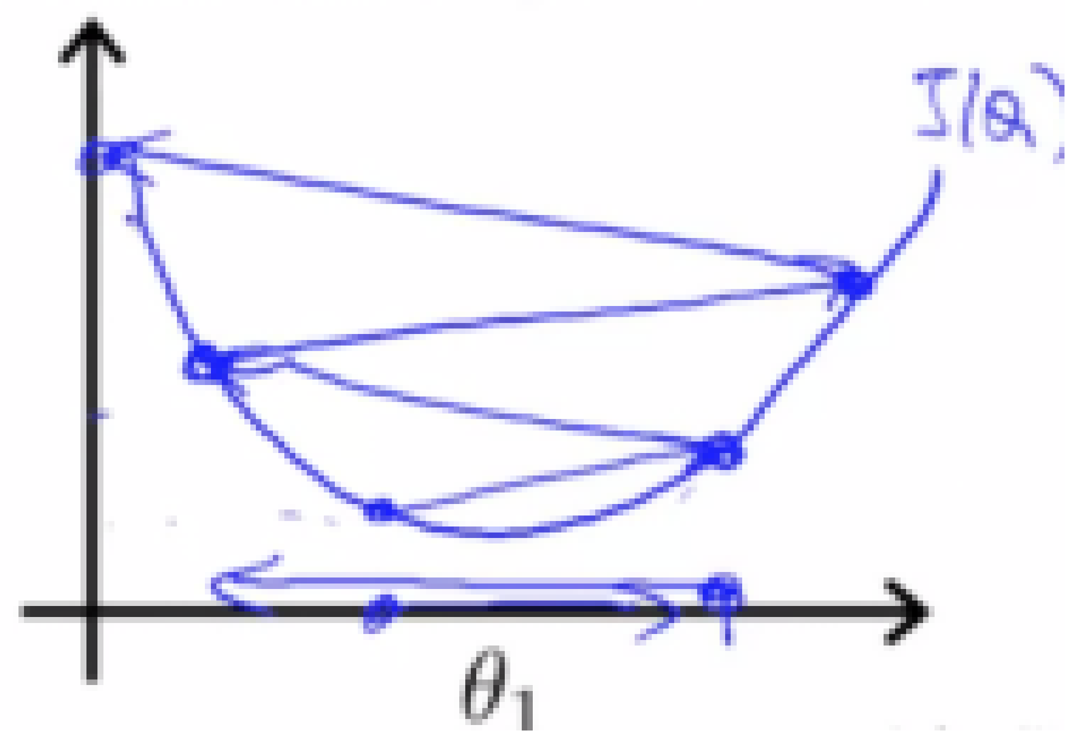
LEARNING RATE

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



Very Small

$\alpha <<$



Very Big

$\alpha >>$

Source Image : Course Supervised Machine Learning: Regression and Classification, Coursera

HOW TO UPDATE

Correct: Simultaneous update

$$\text{temp0} := \theta_0 - \alpha \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$$
$$\text{temp1} := \theta_1 - \alpha \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)$$
$$\theta_0 := \text{temp0}$$
$$\theta_1 := \text{temp1}$$

Incorrect:

$$\text{temp0} := \theta_0 - \alpha \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$$
$$\theta_0 := \text{temp0}$$
$$\text{temp1} := \theta_1 - \boxed{\alpha \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)}$$
$$\theta_1 := \text{temp1}$$

Halotech Academy

Source Image : Course Supervised Machine Learning: Regression and Classification, Coursera

GRADIENT DESCENT

Fungsi --> $f_{w,b}(x^{(i)}) = \mathbf{w}x^{(i)} + \mathbf{b}$

Cost Function --> $J(\mathbf{w},\mathbf{b}) = \frac{1}{2m} \sum_{i=0}^{m-1} (f_{w,b}(x^{(i)}) - y^{(i)})^2$

Gradient Descent will be

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

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

Then

$$w := w - \alpha \frac{1}{m} \sum_{i=0}^{m-1} (f_{w,b}(x^{(i)}) - y^{(i)})\, x^{(i)}$$

$$b := b - \alpha \frac{1}{m} \sum_{i=0}^{m-1} (f_{w,b}(x^{(i)}) - y^{(i)})$$

GRADIENT DESCENT

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

$$\frac{\partial J(w,b)}{\partial w} = \frac{\partial}{\partial w}\left(\frac{1}{2m}\sum_{i=0}^{m-1}\left(f_{w,b}(x^{(i)}) - y^{(i)}\right)^2\right)$$

$$\frac{\partial J(w,b)}{\partial w} = \frac{1}{2m}\sum_{i=0}^{m-1}\frac{\partial}{\partial w}\left(f_{w,b}(x^{(i)}) - y^{(i)}\right)^2$$

$$\frac{\partial J(w,b)}{\partial w} = \frac{1}{2m}\sum_{i=0}^{m-1}\frac{\partial}{\partial w}\left(\mathbf{w}x^{(i)} + \mathbf{b} - y^{(i)}\right)^2$$

$$\frac{\partial J(w,b)}{\partial w} = \frac{1}{2m}\sum_{i=0}^{m-1}\left(2\left(\mathbf{w}x^{(i)} + \mathbf{b} - y^{(i)}\right)x^{(i)}\right)$$

$$\frac{\partial J(w,b)}{\partial w} = \frac{1}{m}\sum_{i=0}^{m-1}\left(\mathbf{w}x^{(i)} + \mathbf{b} - y^{(i)}\right)x^{(i)}$$

$$\frac{\partial J(w,b)}{\partial w} = \frac{1}{m}\sum_{i=0}^{m-1}\left(f_{w,b}(x^{(i)}) - y^{(i)}\right)x^{(i)}$$

Halotech
Academy
Tech.

GRADIENT DESCENT

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

$$\frac{\partial J(w,b)}{\partial b} = \frac{\partial}{\partial b}\left(\frac{1}{2m}\sum_{i=0}^{m-1}(f_{w,b}(x^{(i)}) - y^{(i)})^2\right)$$

$$\frac{\partial J(w,b)}{\partial b} = \frac{1}{2m}\sum_{i=0}^{m-1}\frac{\partial}{\partial b}(\mathbf{w}x^{(i)} + \mathbf{b} - y^{(i)})^2$$

$$\frac{\partial J(w,b)}{\partial b} = \frac{1}{m}\sum_{i=0}^{m-1}(\mathbf{w}x^{(i)} + \mathbf{b} - y^{(i)})$$

$$\frac{\partial J(w,b)}{\partial b} = \frac{1}{m}\sum_{i=0}^{m-1}(f_{w,b}(x^{(i)}) - y^{(i)})$$

Halotech Academy

THANKS FOR WATCHING

Halotechin

halotechin@gmail.com

Halotech Academy