Optimization 3

Recap :

$$\max_{\vec{w}, w_0} \sum_{i=1}^{n} y_i * \left(\frac{w^T x_i + w_0}{\|\vec{w}\|} \right)$$

Gain $\longrightarrow$ Max

$\boxed{\text{Loss} \longrightarrow \text{min}}$ ✓

$$\boxed{\min_{\vec{w}, w_0} - \sum y_i * \left(\frac{w^T x + w_0}{\|\vec{w}\|} \right)}$$

$$\min f(\underbrace{w_1, w_2 \ldots w_n}_{})$$

$w_1, w_2 \ldots w_n, w_0$

eg:-

$$f(x) = 2x^2 + 4x + 3$$

$$f'(x) = \underline{4x+4}$$

$$y \longrightarrow f(\underline{x})$$

$$y = f(x_1, x_2, x_3 \ldots x_n)$$

$$\frac{dy}{dx} = \frac{df(x_1, x_2 \ldots x_n)}{dx}$$

Real-life eg :-

x_1	x_2	x_3	y
Age	Height	weight	Cholestrol

$x_1, \qquad x_2$

$$\boxed{x_2 = g(x_1)}$$

$$Z = f(x_1, x_2)$$

$$= f(x_1, g(x_1))$$

eg :-

$$z = f(x) = x^2 + y^2$$

$$y = g(x)$$

$$z = x^2 + (g(x))^2$$

$$\frac{dz}{dx} = 2x + 2 * \underbrace{g(x) * g'(x)}_{\text{Chain Rule.}}$$

* Partial Derivatives

$$y = f(x)$$

$$\frac{dy}{dx} \longrightarrow \frac{\partial y}{\partial x}$$

eg :- $z = f(x_1, x_2)$

$$\frac{\partial z}{\partial x_1} = f'(\underbrace{x_1}_{\text{var}}, \underbrace{x_2}_{\text{const}})$$

$$\frac{\partial z}{\partial x_2} = f'(\underbrace{x_1}_{\text{const}}, \underbrace{x_2}_{\text{var}})$$

eq2 :- $f(x,y) = 2x^2 y + 3y^3 x^2 + 3y$

$$\frac{\partial f}{\partial x} = 2y(2x) + 3y^3(2x) + 0$$

$$\boxed{\frac{\partial f}{\partial x} = 4xy + 6xy^3}$$

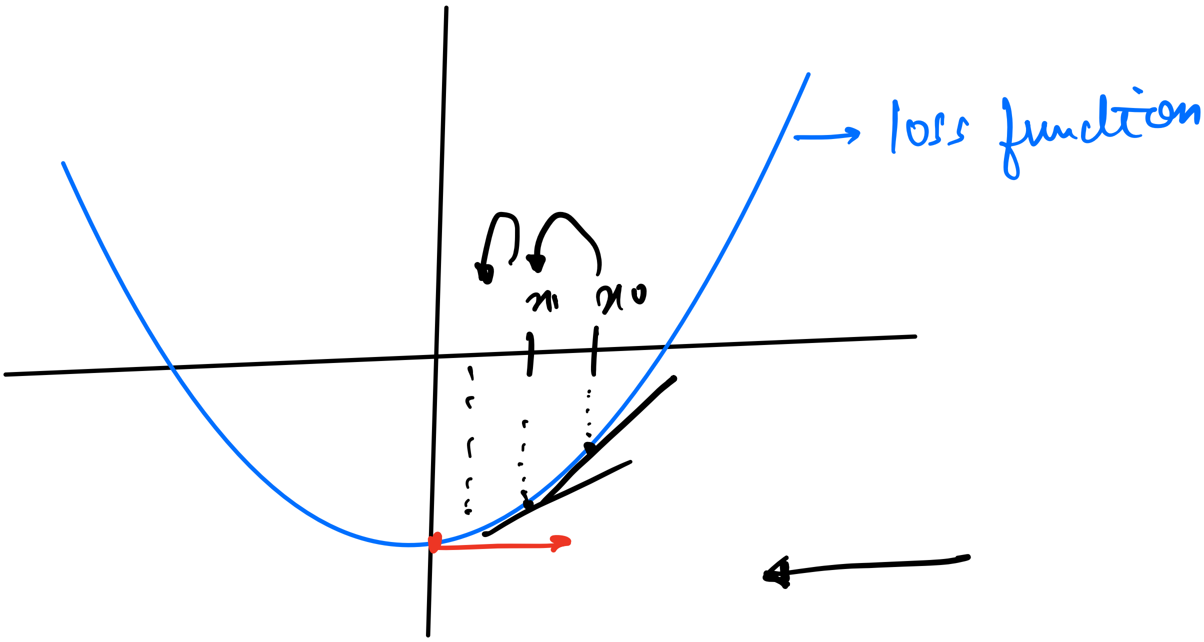
$$\frac{\partial f}{\partial y} = 2x^2(1) + 3x^2(3y^2) + 3$$

$$\boxed{\frac{\partial f}{\partial y} = 2x^2 + 9x^2y^2 + 3}$$

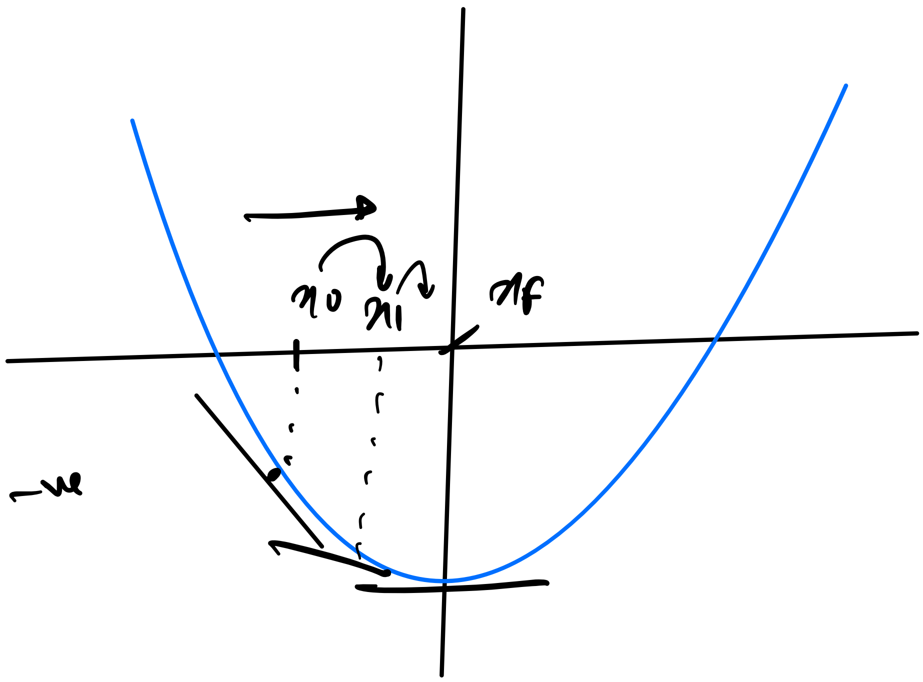
goal :- $\min_{w_1, w_2 \ldots w_n, w_0} \underbrace{f(w_1, w_2, w_3 \ldots w_n)}$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial w_0} \\ \frac{\partial f}{\partial w_1} \\ \frac{\partial f}{\partial w_2} \\ \vdots \\ \frac{\partial f}{\partial w_n} \end{bmatrix} \leftarrow \text{gradient}$$

(+) Gradient Descent Intuition :-



$$x_1 = x_0 - \boxed{f'(x)}$$



$$x_1 = x_0 - \boxed{f'(x)}$$

d +ve

Case1 :-

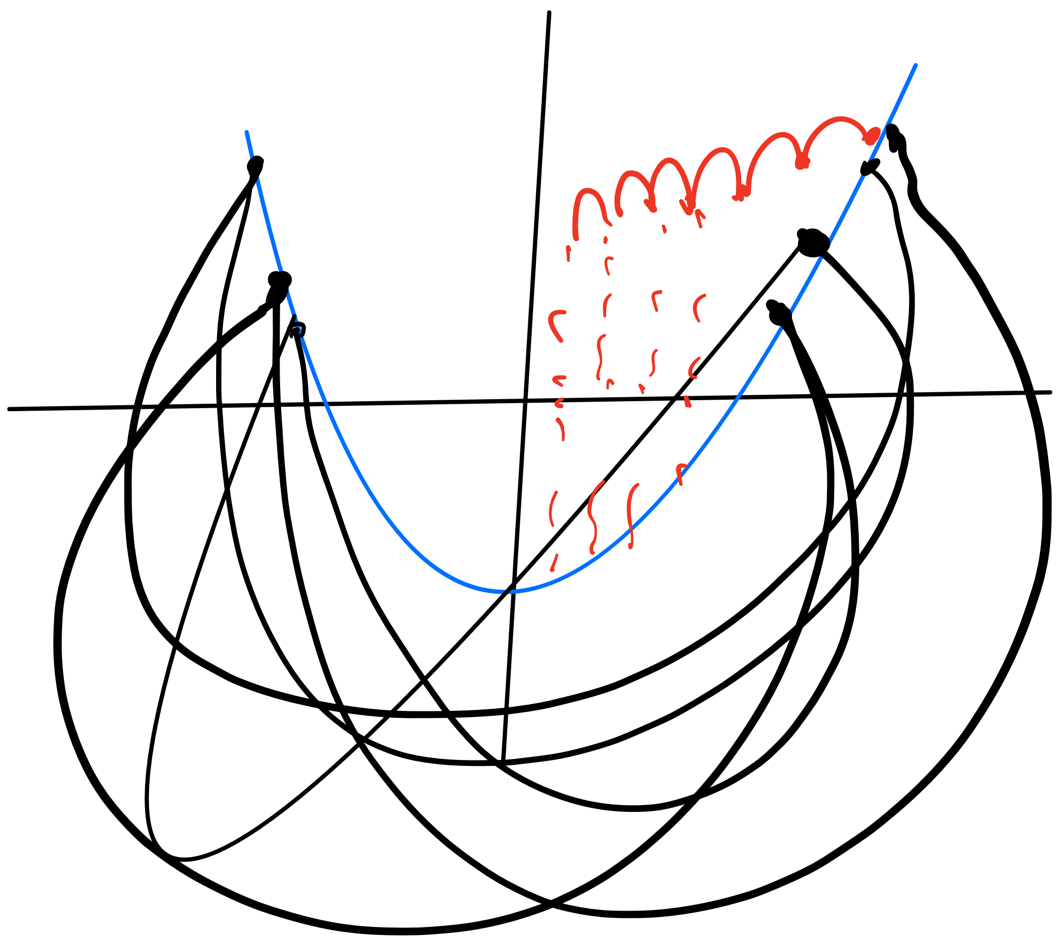
$f'(x) \Rightarrow +ve$

$$x_1 = x_0 - \underline{(+ve)}$$

Case 2 :- $f'(x) \Rightarrow -ve$

$$x_1 = x_0 - (-ve)$$

$$x_1 = x_0 + (+ve)$$



$$x_1 = x_0 - \underbrace{f'(x)}$$

↓

v. v. large

Learning rate

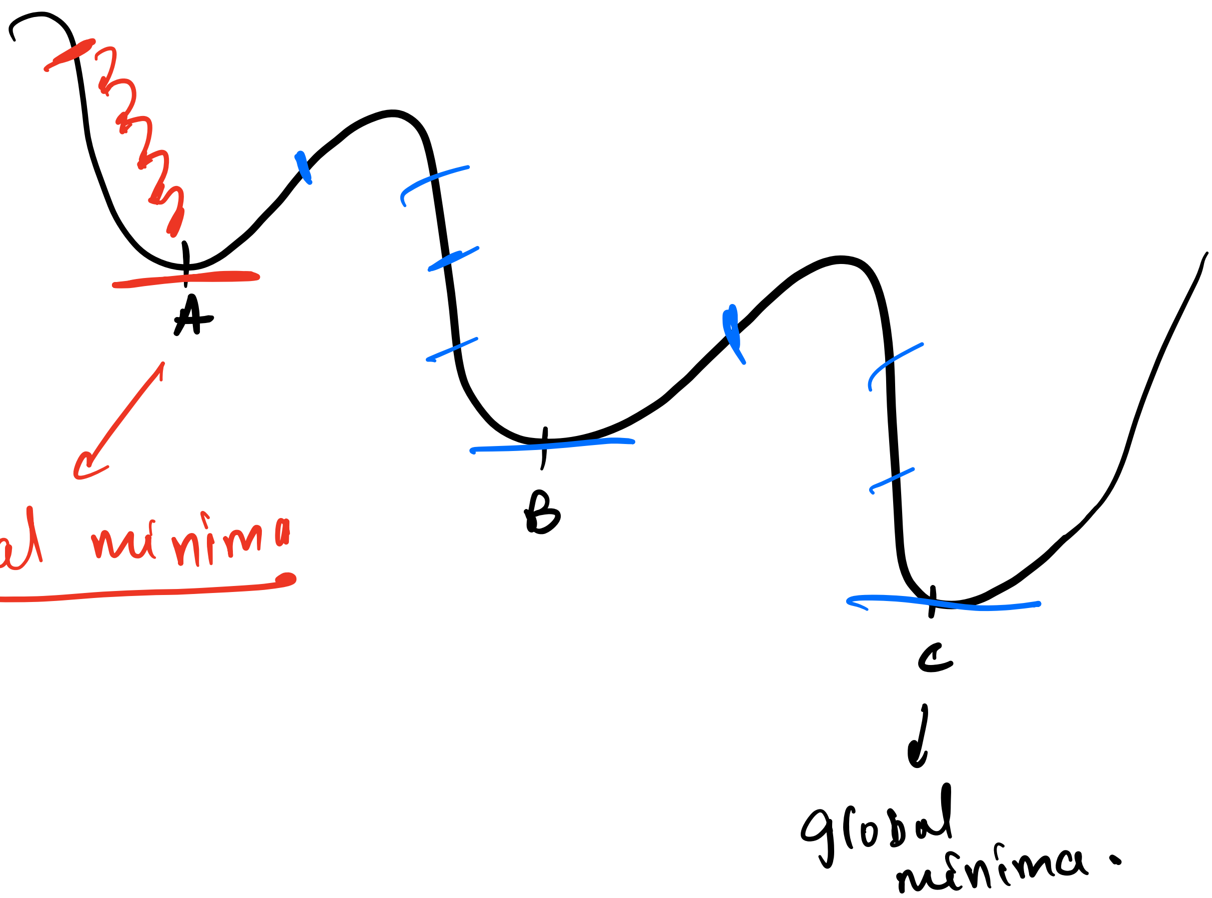
$$\boxed{\underline{x_1} = \underline{x_0} - \eta * \underline{f'(x)}}$$

Case 1 :- $f'(x) \longrightarrow$ v. v. large

$\eta \longrightarrow 0.0001$

Case 2 :- $f'(x) \longrightarrow$ v. v. small

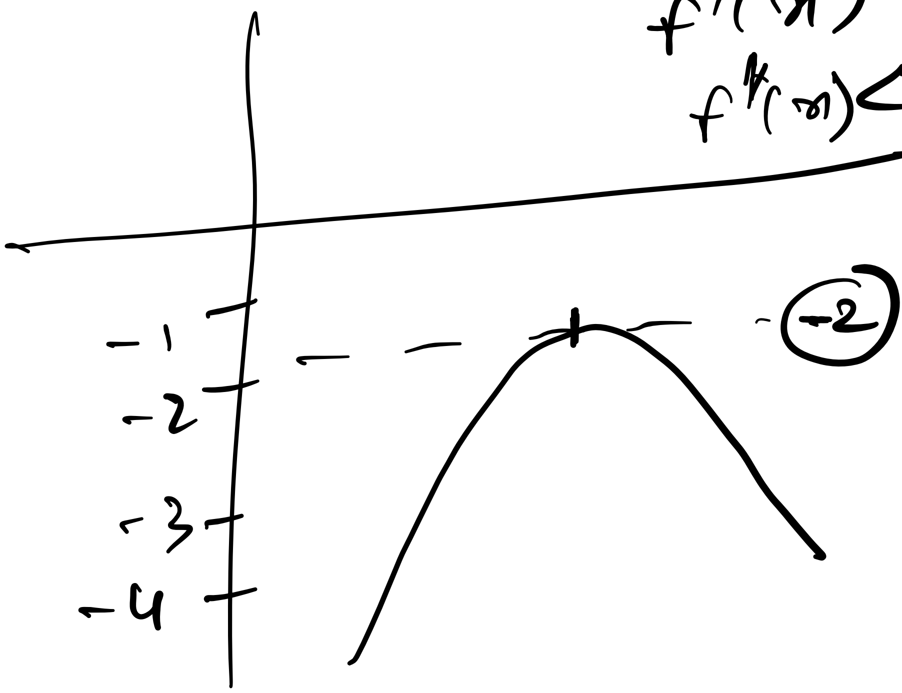
$\eta \longrightarrow$ large



eg

$f''(x) > 0$

$f''(x) < 0$



YOE	salary
⋮	⋮

classification

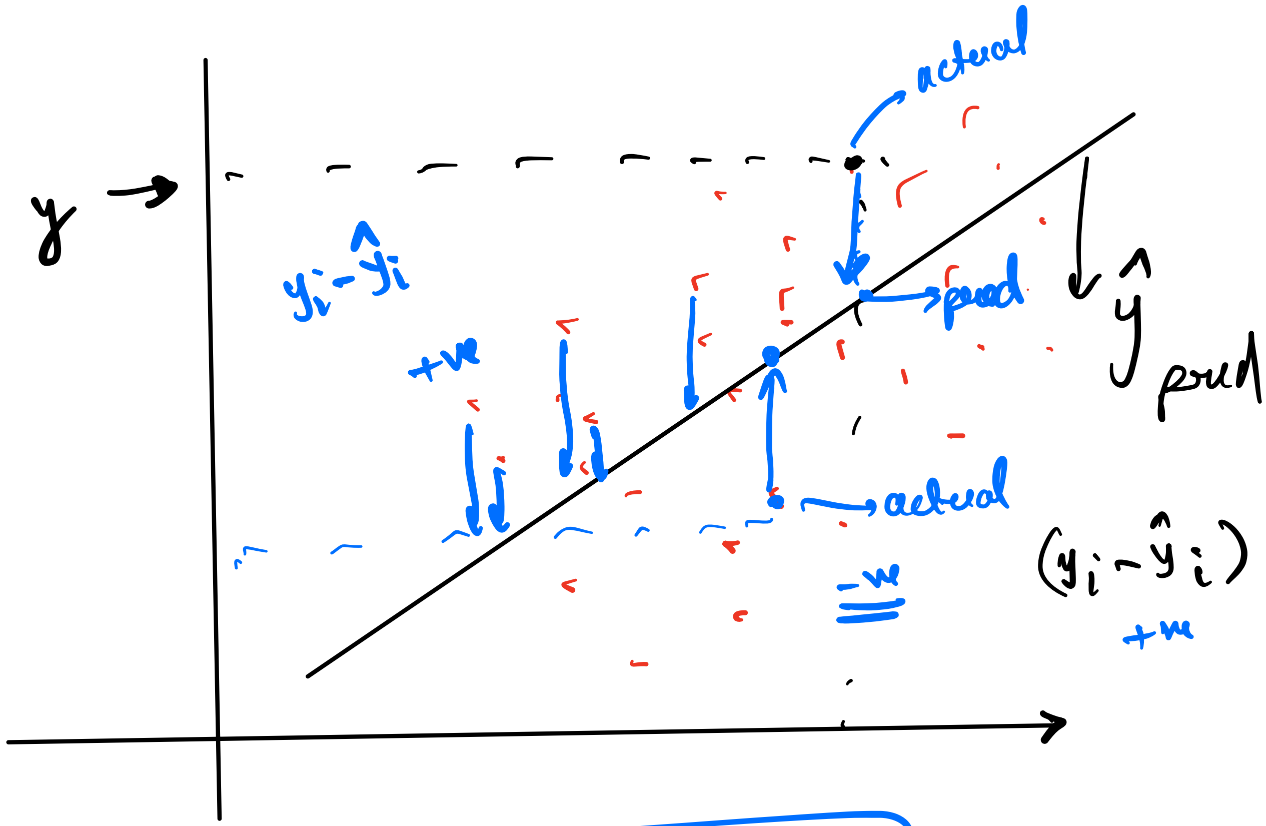
spam

not spam

image → Cat, dog, Goat

Regression

Rooms	locality . . .	→ Price
		y



$$\boxed{\hat{y}_i = m x_i + c}$$

$$J(m, c) = \frac{1}{N} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$\vec{w}$, w_0

$$J(m,c) = \frac{1}{N} \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2 \qquad \hat{y}_i = mx_i + c$$

$$J(m,c) = \frac{1}{N} \sum_{i=1}^{n} \left(y_i - (mx_i + c) \right)^2$$

$$\boxed{\frac{d}{dx}\left(\left(g(x)\right)^2 \right) \longrightarrow 2 * g(x) * g'(x)}$$

$$\frac{\partial J}{\partial m} = \frac{1}{n} \sum_{i=1}^{n} 2 * \underbrace{\left(y_i - (mx_i + c) \right)} * (-x_i)$$

$$= \frac{-2}{n} \sum_{i=1}^{n} \left(y_i - (mx_i + c) \right) * x_i$$

$$J(m,c) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - (mx_i + c)\right)^2$$

$$\frac{\partial J}{\partial c} = \frac{1}{n}\sum_{i=1}^{n}\left[2 * \left(y_i - (mx_i + c)\right) * (-1)\right]$$

$$\boxed{\frac{\partial J}{\partial c} = \frac{-2}{n}\sum_{i=1}^{n}\left(y_i - (mx_i + c)\right)}$$

$x = x + 1$

for i in range(5)

$x = x + 1$

$$w_1[next] = w_1[current] - \eta * \frac{\partial f}{\partial w_1}$$

$$w_2[next] = w_2[current] - \eta * \frac{\partial f}{\partial w_2}$$

$$w_3[\quad]$$

$$\vdots$$

$$w_n[next] = w_n[current] - \eta * \frac{\partial f}{\partial w_n}$$

Discussion

$i = 5000$