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

Set the first derivative to 0

$$f'(x) = 0$$

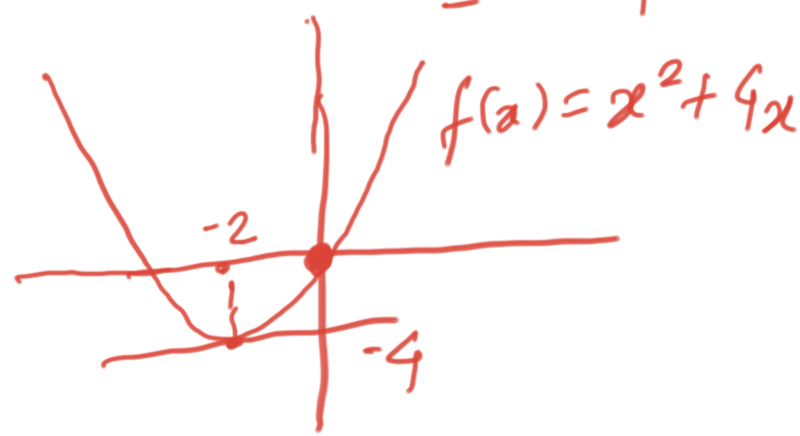
$$\Rightarrow 2x + 4 = 0$$

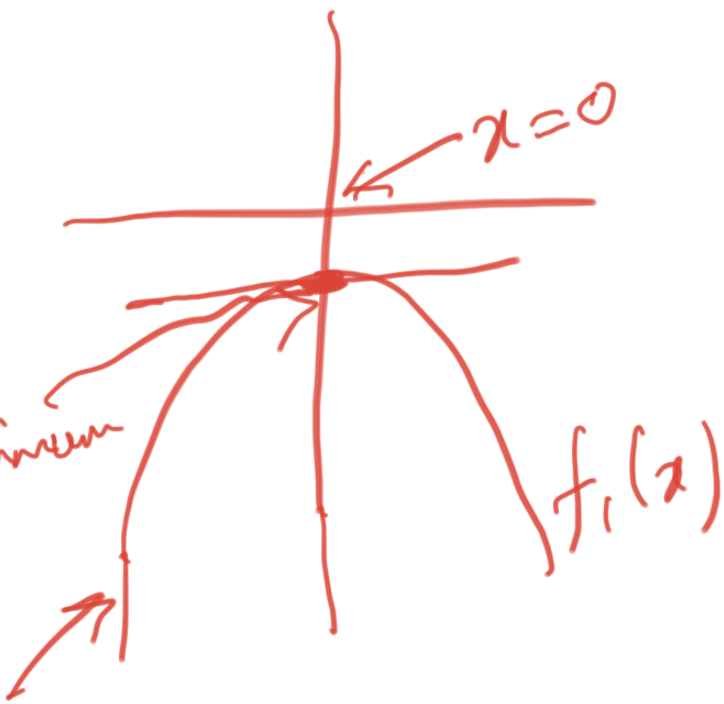
$$\Rightarrow x = -2.$$

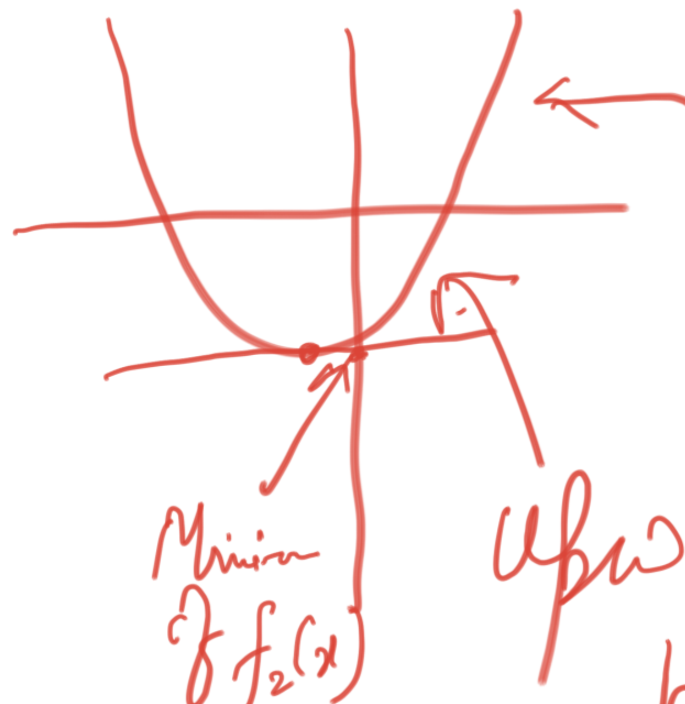
$$f''(x) = 2 > 0$$

$$f_{min} = (-2)^2 + 4(-2)$$

$$= -4$$







$f'(x)=0 \Rightarrow x \leftarrow$ for these x we will get the extrema

Actual

Predicted

Regression

$\uparrow$

DV

$$X \xrightarrow{f(\theta_1, \theta_2)} Y$$

DV

Dependent variable

$$Y = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$

Actual (DV)

Predicted (f)

$\boxed{f = \theta_0 + \theta_1 x}$

$x \longrightarrow 10$

$f(x) \longrightarrow 9.5$

$\mathcal{L} = g(A, P)$

$f = e^{\theta_0 + \theta_1 x}$

$f(\theta_0, \theta_1)$

Loss function $= \frac{1}{n} \sum_{i=1}^{n} \left(f(x_i) - A(x_i)\right)^2$

$\mathcal{L} = \frac{1}{n} \sum \left[\theta_0 + \theta_1 x_i - A(x_i)\right]^2$

Loss function $\mathcal{L} = \frac{1}{n} \sum \left[\theta_0 + \theta_1 x_i - A(x_i)\right]^2$

Model parameters

$\frac{\partial \mathcal{L}}{\partial \theta_0} = 0$, $\frac{\partial \mathcal{L}}{\partial \theta_1} = 0$

$\frac{d\mathcal{L}}{d\theta} = 0$

(X, Y)

$X \xrightarrow{f(\theta_0, \theta_1)} Y$

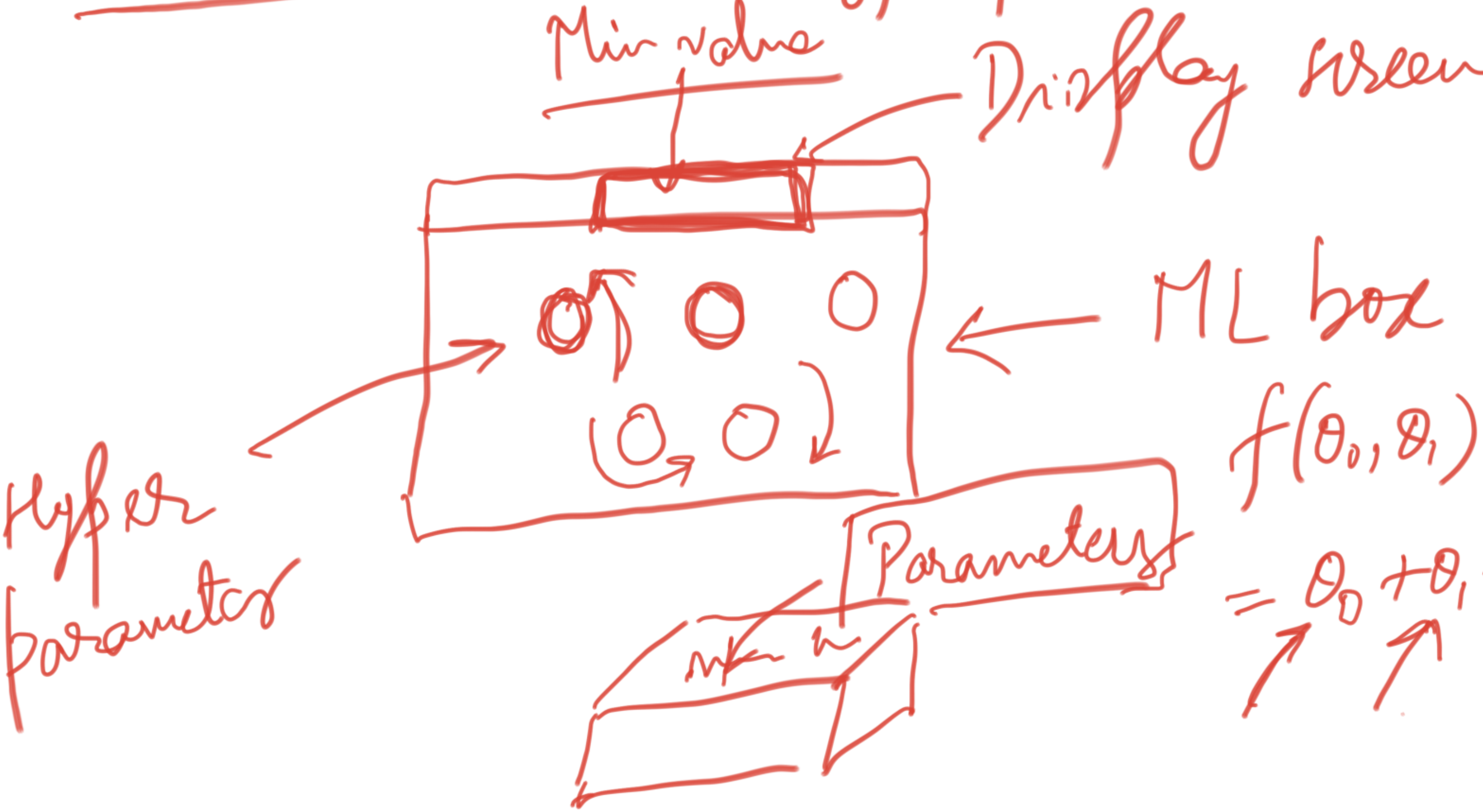
$\theta_0 = 2$

$\theta_1 = 5$

$\rightarrow (2 + \cancel{5x2}) =$

Parameters

Hyperparameters



$f(\theta_0, \theta_1) = \theta_0 + \theta_1 x$

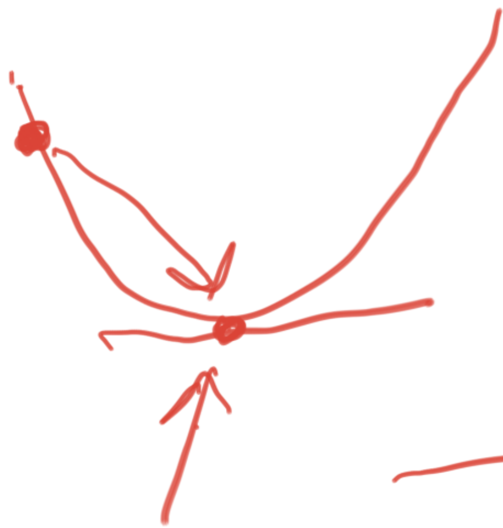
$$f(x) = x^2 + 4x$$

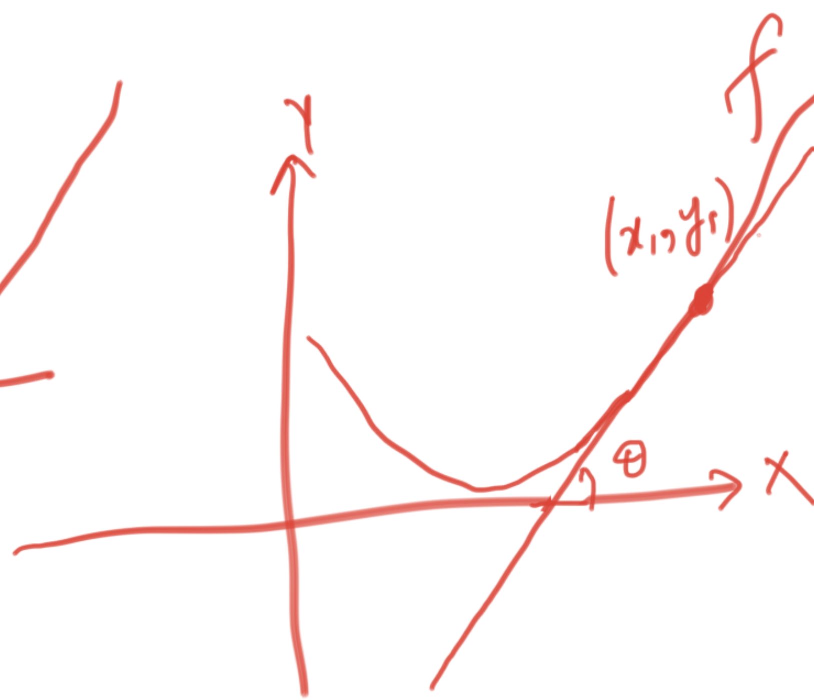
Analytically

$$f'(x) = 0$$

Numerically (Iterative procedure)

Gradient Descent Algorithm





derivative of f at point (x_1, y_1) is $\tan\theta$

<u>Gradient</u>

<u>optimization</u>

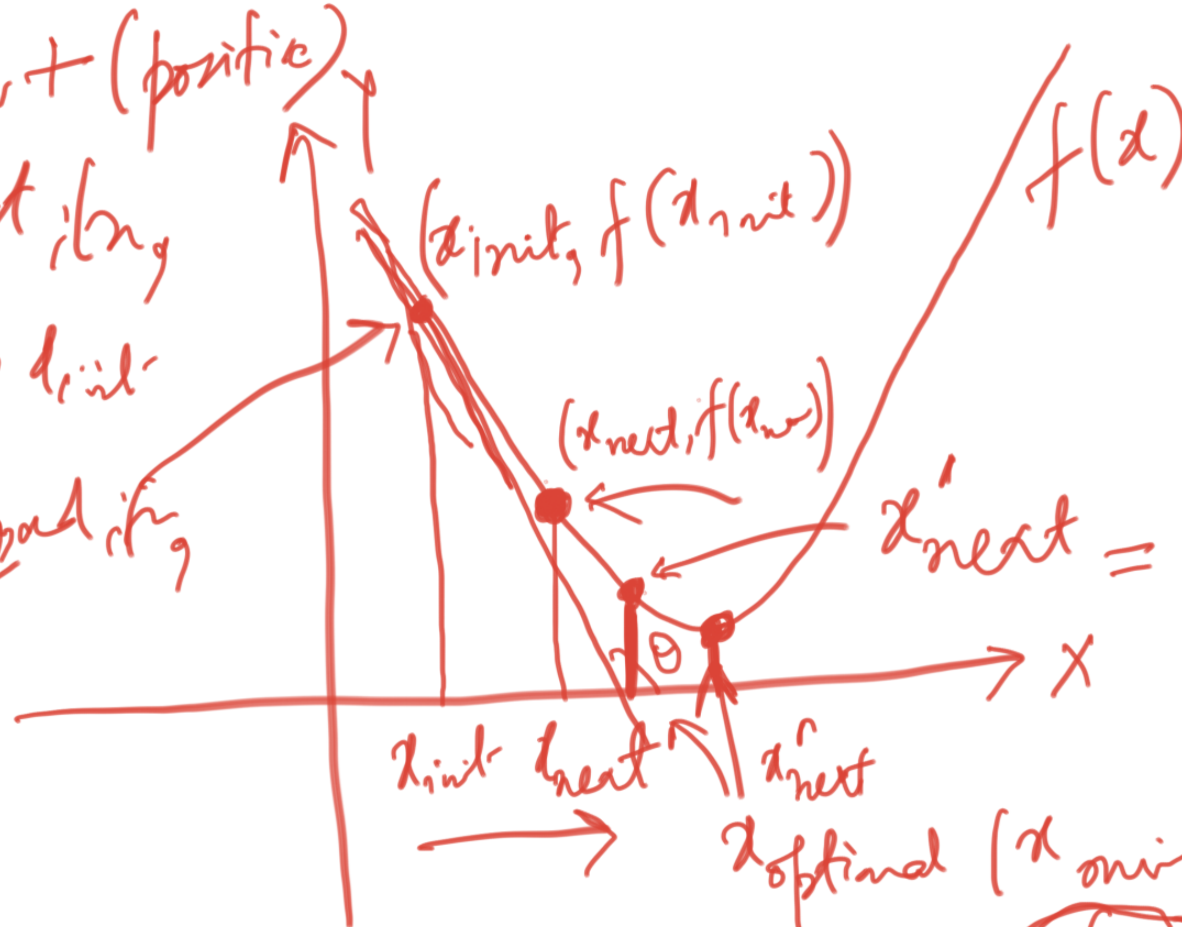
<u>Descend</u> ← Descent ?

$x_{next} = x_{init} + (\text{positive})$

In the 1st itn,

$x_{next} > x_{init}$

In the 2nd itn,



$x_{next} = x_{prev} - \lambda \times \left(\frac{\partial f}{\partial x}\right)$

$\tan\theta$ $\quad x_{prev}$

$x'_{next} = x_{next} + (+ve)$

$x'_{next} > x_{next}$

-ve

GD eqn $\rightarrow$ $x_{next} = x_{init} - \lambda \times \frac{\partial f}{\partial x}\Big|_{x_{init}}$

+ve

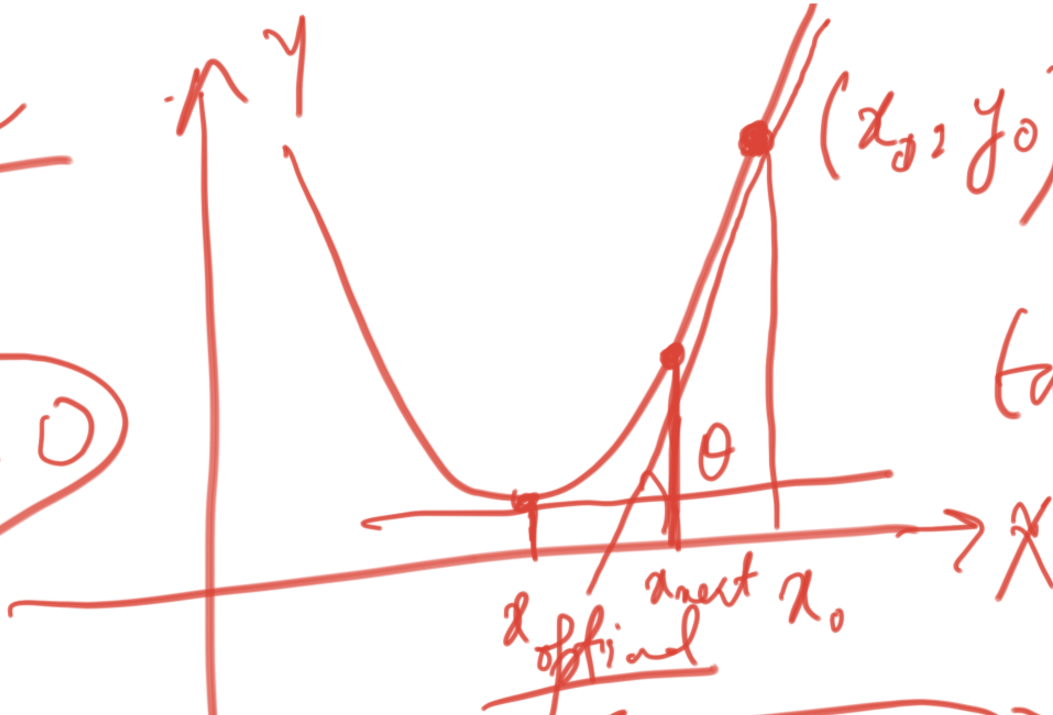
We will stop iterating when the update (diff between x_{next} & x_{prev}) is smaller than ϵ

$$|x_{next} - x_{prev}| < \epsilon$$

<u>Convergence</u>

→ Convex fn

$f''(x) > 0$



$\tan\theta > 0$

+ve

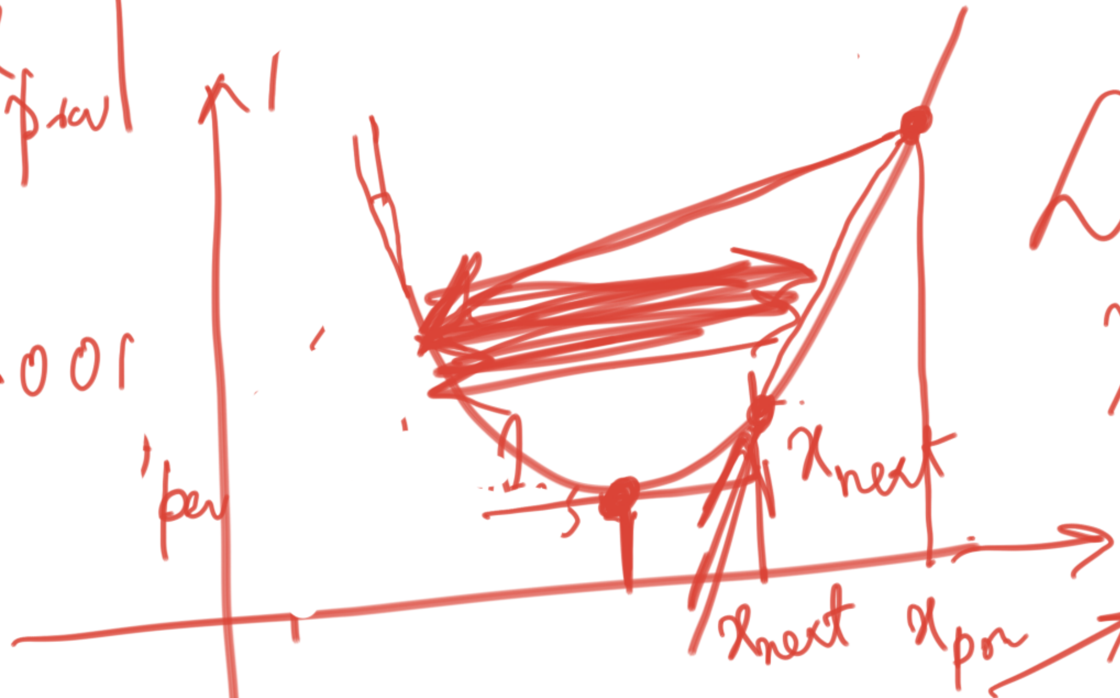
$$x_{next} = x_{prev} - \lambda \times \left(\frac{\partial f}{\partial x}\right)_{x_{prev}}$$

$x_{next} < x_{prev}$

Hessian matrix definite
→ positive semi-

$$|x_{next} - x_{prev}| < \epsilon$$

0.001



$L = f(\theta)$

$\lambda > 5$

θ

Model parameter

$$\boxed{x_{next} = x_{prev} - \left.\lambda \times \frac{\partial f}{\partial x}\right|_{x_{prev}}}$$

1^{st} itn

$$\theta_1 = \theta_0 - \lambda \times \frac{\partial f}{\partial \theta}\Big|_{\theta_0}$$

θ_{n-1}

0.256

$\theta_{n-1} \approx \theta_n$

$$|\theta_1 - \theta_0| < \epsilon$$

θ_n

0.256121

2nd itn

$$\theta_2 = \theta_1 - \lambda \frac{\partial f}{\partial \theta}\Big|_{\theta_1}$$

$\frac{1}{1000}$

$$|\theta_2 - \theta_1| < \epsilon$$

3rd itn

$\vdots$

$$\rightarrow \quad |\theta_n - \theta_{n-1}| < \epsilon$$

$\theta_0, f(\theta)$

single

$f(\theta_n)$

θ_{prev} θ_n takes a value

Convergence

Multivariable GD

$t \to t^{th}$ iteration in ML algo

$\theta_0, \theta_1, \theta_2, \theta_3, \theta_n$

$$\left\{ \begin{bmatrix} \theta_0^{t+1} \\ \theta_1^{t+1} \\ \theta_2^{t+1} \\ \vdots \\ \theta_n^{t+1} \end{bmatrix} = \begin{bmatrix} \theta_0^{t} \\ \theta_1^{t} \\ \vdots \\ \theta_n^{t} \end{bmatrix} - \lambda \times \begin{bmatrix} \frac{\partial f}{\partial \theta}\Big|_{\theta_0^t} \\ \frac{\partial f}{\partial \theta}\Big|_{\theta_1^t} \\ \frac{\partial f}{\partial \theta}\Big|_{\theta_n^t} \end{bmatrix} \right.$$

$$\sqrt{\underbrace{(\theta_0^{t+1} - \theta_0^t)^2}_{} + \underbrace{\ldots + (\theta_n^{t+1} - \theta_n^t)^2}_{}} < \epsilon$$

$$
\begin{cases}
\theta_0^{t+1} = \theta_0^{t} - \lambda \dfrac{\partial f}{\partial \theta}\Big|_{\theta_0^{t}} & \text{— } 1 \\[2ex]
\theta_1^{t+1} = \theta_1^{t} - \lambda \dfrac{\partial f}{\partial \theta}\Big|_{\theta_1^{t}} & \text{— } 2 \\
\vdots \\
\theta_n^{t+1} = \theta_n^{t} - \lambda \times \dfrac{\partial f}{\partial \theta}\Big|_{\theta_2^{t}} & \text{— } n
\end{cases}
$$

$$f(\theta_0, \theta_1) = \theta_0^2 + \theta_1^2 + 1$$

Multivariate fn.

$$f_{min} = 0^2 + 0^2 + 1$$
$$= 1$$

$$\frac{\partial f}{\partial \theta_0} = 2\theta_0 \quad , \quad \frac{\partial f}{\partial \theta_1} = 2\theta_1$$

$$\Rightarrow 2\theta_0 = 0 \quad , \quad 2\theta_1 = 0$$

$$\theta_0 = 0 \quad , \quad \theta_1 = 0$$