gradient $L = \frac{1}{2} \cdot \left((2\theta_1 - 2)^2 + (\theta_0 - 3)^2 \right)$

$$\nabla_\theta L = \begin{bmatrix} \frac{\partial L}{\partial \theta_0} \\ \frac{\partial L}{\partial \theta_1} \end{bmatrix} = \begin{bmatrix} \theta_0 - 3 \\ (2\theta_1 - 2) * 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

vector.

$$\frac{dy}{dx} = \frac{dy}{d\theta} \cdot \frac{d\theta}{dx}.$$

$$\frac{\partial L}{\partial \theta_0} = \frac{1}{2} \cdot \frac{\partial}{\partial \theta_0} \left(\frac{c}{\theta_1} + (\theta_0 - 3)^2 \right)$$

$$= \frac{1}{2} \frac{\partial}{\partial (\theta_0 - 3)} \cdot (\theta_0 - 3)^2 \cdot \frac{d}{d\theta_0} \cdot (\theta_0 - 3)$$

$$= \frac{1}{2} \cdot \frac{d}{x} \cdot x^2 \cdot \frac{d}{d\theta_0} (\theta_0 - 3)$$

$$= \frac{1}{2} \cdot 2 \cdot (\theta_0 - 3) \cdot 1$$

$$= (\theta_0 - 3)$$

$$\nabla_\theta L = \begin{bmatrix} \theta_0 - 3 \\ (2\theta_1 - 2) 2 \end{bmatrix} \qquad L = \frac{1}{2} \left((2\theta_1 - 2)^2 + (\theta_0 - 3)^2 \right)$$

$$\theta^{(1)} = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \qquad L_{\theta^{(1)}} = 0$$

$$\nabla_{\theta^{(1)}} L = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Gradient Descent Solution

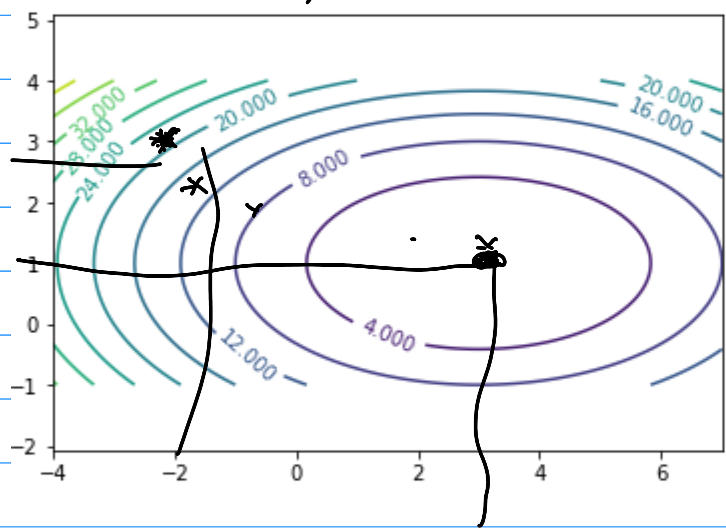
$$\theta = (X^T X)^{-1} X^T y$$

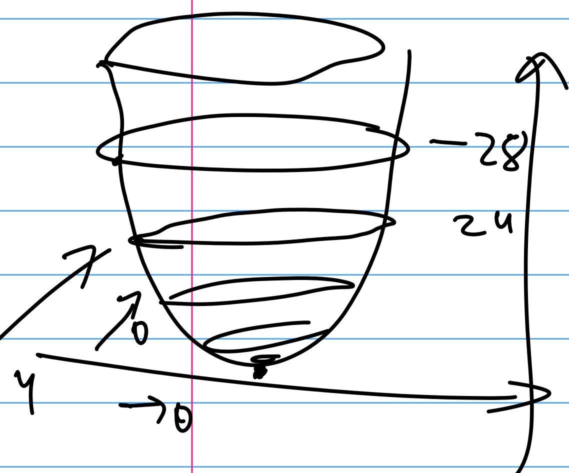
$D = \{(x_1, y_1) \ldots (x_n, y_n)\}$

$= \{(\mathbf{x}_1, y_1) \ldots (\mathbf{x}_n, y_n)\}$ $\quad \mathbf{x}^{(1)} = \begin{bmatrix} x_0^{(1)} \\ x_1^{(1)} \\ \vdots \end{bmatrix}$

$L = (y - (\theta_0 x_0 + \theta_1 x_1 \ldots \theta_d \ldots))^2$

$L = f(\theta_0, \theta_1)$





$$y = \theta_0 x_0 + \theta_1 x_1 + \theta_d x_d$$

$$ax^2 + bx + c = 0 \Rightarrow 2x^2 - 3x + 4 = 0$$

find $x = ?$

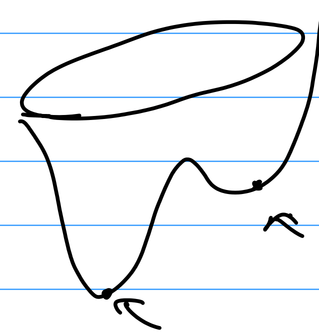
$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

closed form soln.

$x = ?$

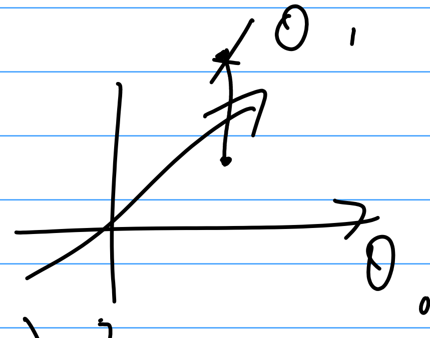
$$ax^3 + bx^2 + cx + d = 0 \quad x = ?$$

There is no closed form solution for this



$$L = f(\theta_0, \theta_1) = \frac{1}{2}\left((2\theta_1 - 2)^2 + (\theta_0 - 3)^2\right)$$

$$\nabla_\theta L = \begin{bmatrix} \frac{\partial L}{\partial \theta_0} \\ \frac{\partial L}{\partial \theta_1} \end{bmatrix} = \begin{bmatrix} \theta_0 - 3 \\ 2(2\theta_1 - 2) \end{bmatrix} \quad \text{— } 1$$



$$\min_\theta L \; ; \; \theta^* = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad L(\theta^*) = 0$$

$$\left[\theta_i = \theta_{i-1} - \alpha * \underbrace{\nabla_\theta L(\theta_{i-1})}\right]$$

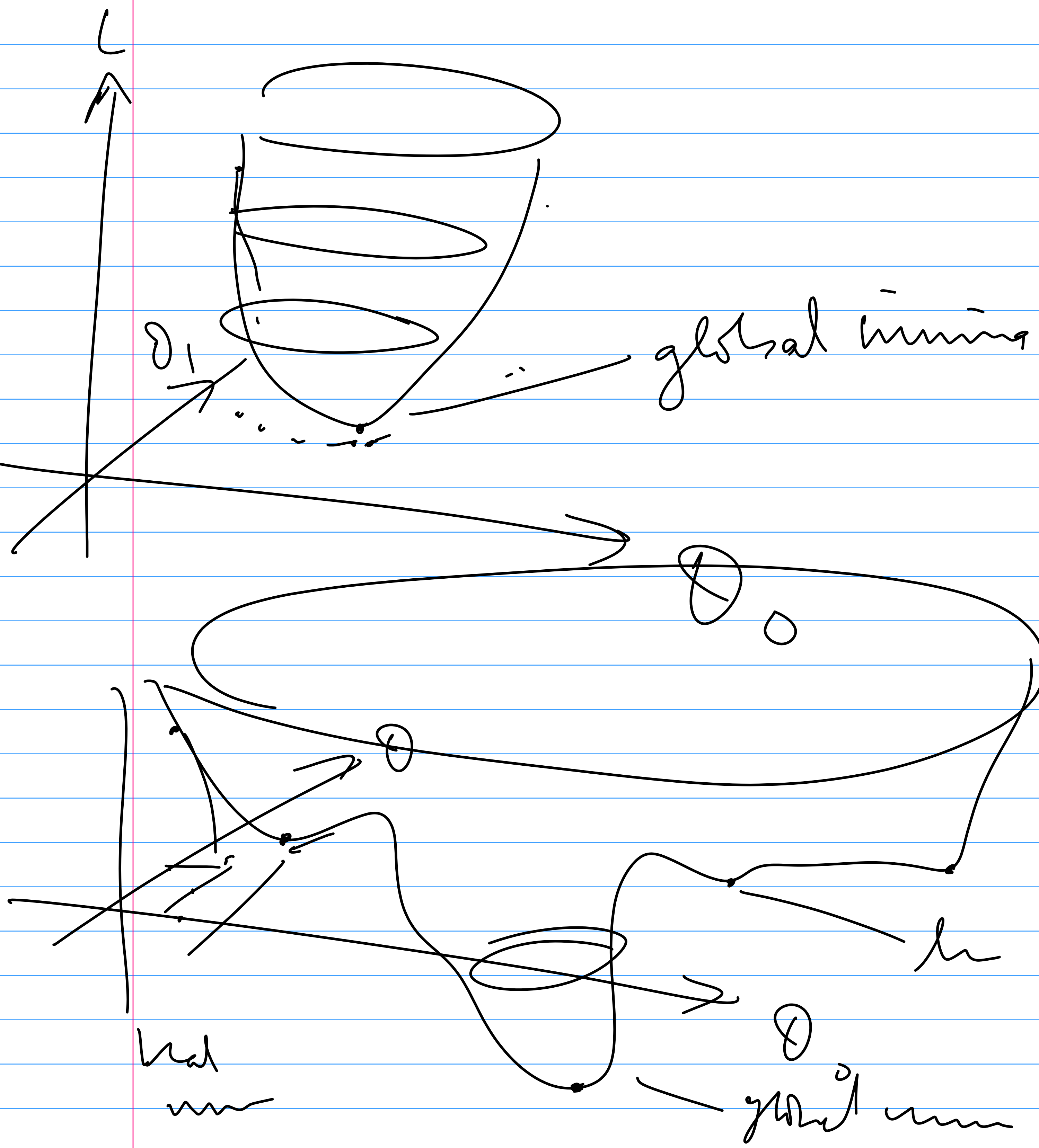
$$\theta^{(0)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} ; \; L = 4 \; ; \; \nabla_{\underbrace{\theta^{(0)}}} L = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

$$\frac{\partial L}{\partial \theta_0} = -2 \; ; \; \theta_0 + \varepsilon \rightarrow L + \varepsilon \cdot \frac{\partial L}{\partial \theta_0} \qquad \alpha = 0.1.$$

$$\varepsilon \rightarrow 0.$$

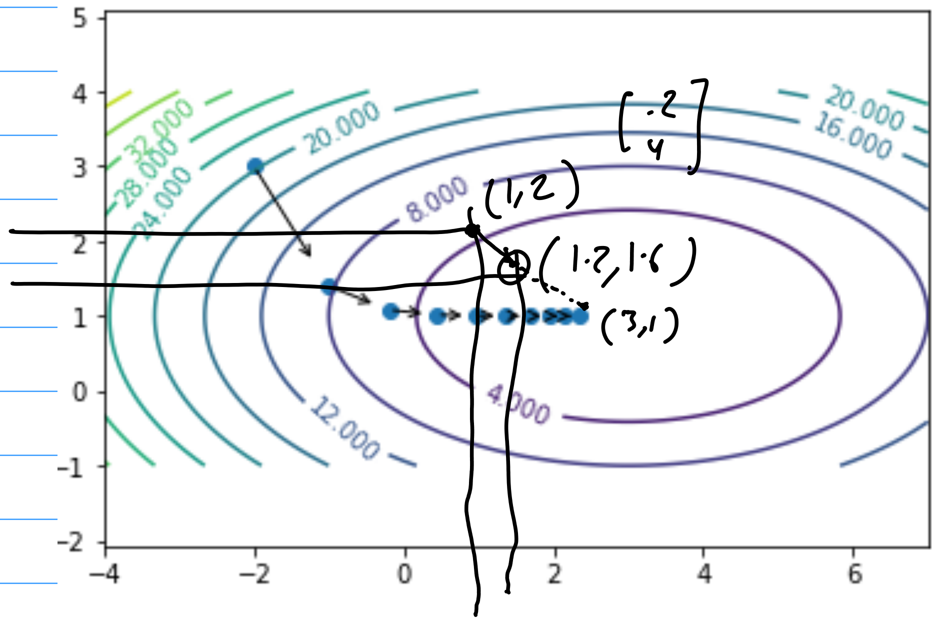
$$\theta^{(1)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - 0.1 \times \begin{bmatrix} -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1.2 \\ 1.6 \end{bmatrix}$$

$$L(\theta^{(1)}) = 4.68 = 2.34$$


L
θ_1
global minima
θ_0
local min
h
θ
global min

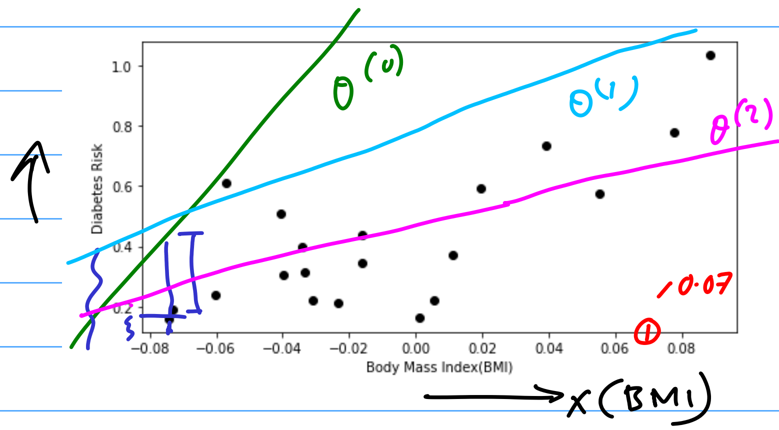
$$\theta^{(i)} = \theta^{(i-1)} - \alpha \cdot \nabla_\theta L(\theta^{(i-1)}) \longrightarrow \text{①} \text{ Gradient } \underline{\text{Descent}}$$

$\alpha = 0.1$



$y = \theta^T x$

$= \theta_0 + \theta_1 x$



$\theta^{(0)} = \begin{bmatrix} \theta_0^{(0)} \\ \theta_1^{(0)} \end{bmatrix}$

$= \begin{bmatrix} 0.35 \\ 0.8 \end{bmatrix}$

$L = MSE$ $\quad \nabla_\theta L \quad$ $\theta^{(1)} = \theta^{(0)} - \alpha \cdot \nabla_\theta (L)$