Linear Regression

Wednesday, 8 December 2021 7:51 PM

Cars24 : automate the process of estimate the price of used car

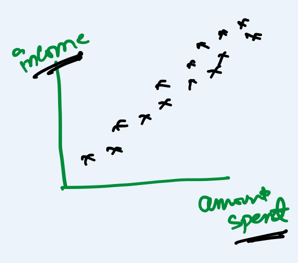
Data Available

x = [Brand, Model, Year, kms, Engine] [selling price]

x_1
x_2
x_3
x_4
x_5

Supervised ML (x, y)

max power → x, selling price → y



Regression : $y \in R$

Training Data → ML Algo → h → $\hat{y}$ (Prediction)

x new → h

$x \xrightarrow{h} y$

y $\quad \boxed{w_0} + \boxed{w_1} x$ $\quad$ -7 + 0.14 × 1

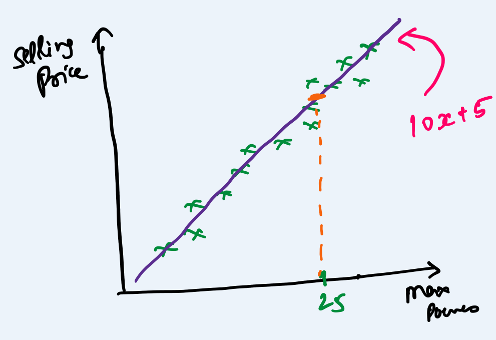
1, —
2, —
3, —
4, —

$0.14 \times \begin{bmatrix} 1 \\ 5 \\ 10 \\ 15 \end{bmatrix} + (-7) \Rightarrow \begin{bmatrix} \\ \end{bmatrix}$ y

X-new

$\boxed{y = mx + c}$

$\boxed{w_1 . x_1 + w_0 = y}$ / $\theta_1 x_1 + \theta_0 \Rightarrow y$



$w_0 = 5$
$w_1 = 10$

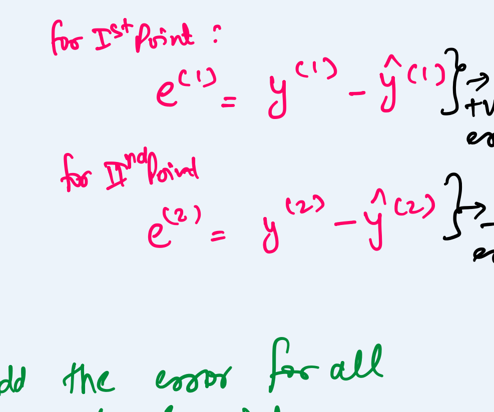
$x\_test = [25]$

y hat ($\hat{y}$) Prediction $\Rightarrow 10 \times 25 + 5 \Rightarrow$
$\Rightarrow$ Rs 255

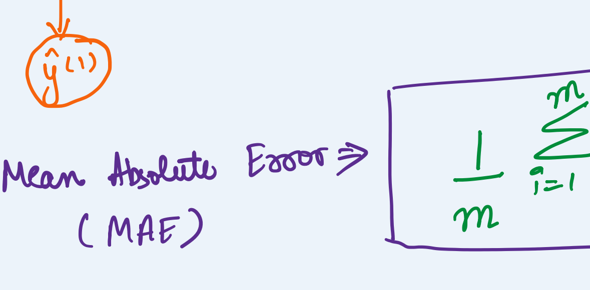
θ_1 w_1 ?
θ_0 w_0 ?

→ maximum points should exactly on the line or near to line.

why line 3 ?



Error function



for I^{st} Point :

$e^{(1)} = y^{(1)} - \hat{y}^{(1)} \Big\} \to$ +ve error

for II^{nd} Point

$e^{(2)} = y^{(2)} - \hat{y}^{(2)} \Big\} \to$ -ve error

$x^{(1)}$ $y^{(1)}$

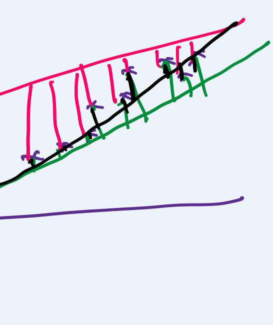
$\hat{y}^{(1)}$

Add the error for all 'm' points

Mean Absolute Error $\Rightarrow$ $\boxed{\frac{1}{m} \sum_{i=1}^{m} \left| y^{(i)} - \hat{y}^{(i)} \right|}$

(MAE)

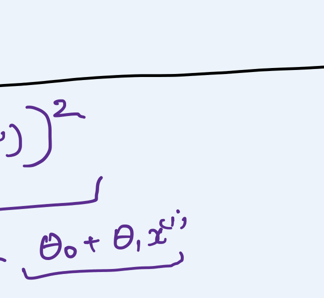
$|x| \to$



Instead of abs, we take square.

$$\boxed{MSE = \frac{1}{2m} \sum_{i=1}^{m} \left(y^{(i)} - \hat{y}^{(i)} \right)^2}$$

Loss function / Error
Cost function



Goal : Minimize Loss function
Error

modify my hypothesis function such that the Loss is minimized.

$\boxed{h_{\theta}(x) = \theta_0 + \theta_1 x}$ Hypothesis function

---

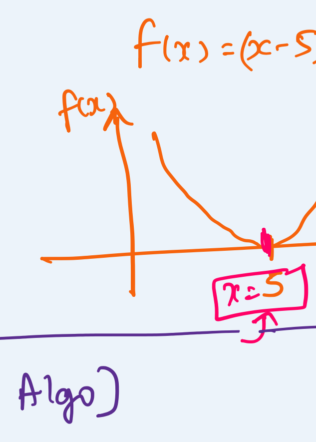
$$\underbrace{\mathcal{L}(y, \hat{y}) = \frac{1}{2m} \sum_{i=1}^{m} \left(y^{(i)} - (\boxed{\theta_0} + \boxed{\theta_1} x^{(i)}) \right)^2}_{Loss}$$

$\hat{y}^{(i)} = \theta_0 + \theta_1 x^{(i)}$

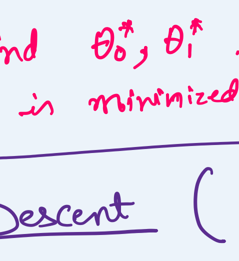
$y = x^2$

$\frac{1}{2m} \sum (\quad Z \quad)^2$

Variables → θ_0 θ_1

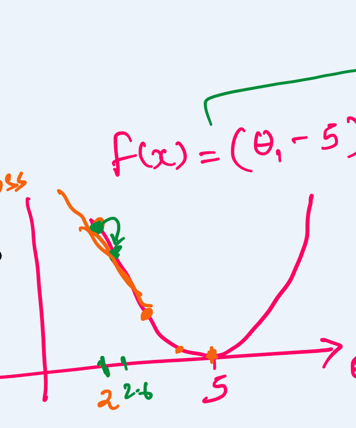


x^2



θ_0^*
θ_1^* $\Rightarrow$ $\boxed{\underset{(\theta_0, \theta_1)}{\operatorname{argmin}} \; \mathcal{L}}$ → objective

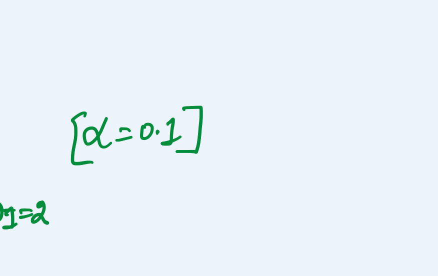
I need to find θ_0^*, θ_1^*. Such that is minimized.



Gradient Descent (Optimization Algo)

$$\boxed{\theta_0^*, \theta_1^* = \underset{\theta_0, \theta_1}{\operatorname{argmin}} \; \mathcal{L}}$$

1. Random init θ's
   $\theta_1 = 2$

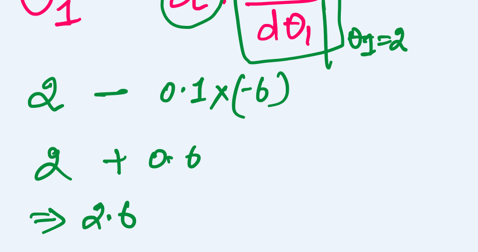


2. $\frac{d\mathcal{L}}{d\theta_1}$ $\quad \mathcal{L} \to (\theta_1 - 5)^2$

   $\Rightarrow 2(\theta_1 - 5)$

   $\Rightarrow 2(2 - 5)$

   $\frac{d\mathcal{L}}{d\theta_1} \Rightarrow -6$



3. take a step.

   $\theta_1 = \theta_1 - \alpha \boxed{\frac{d\mathcal{L}}{d\theta_1}}\Big|_{\theta_1 = 2}$ $\quad [\alpha = 0.1]$

   α → step size → 0.1

   $= 2 - 0.1 \times (-6)$

   $2 + 0.6$

   $\theta_1 \Rightarrow 2.6$

4. Repeate step-2 & 3 until Convergence.

---

1. Randomly init θ
   $\theta_0 = $ —
   $\theta_1 = $ —

2. $\frac{\partial \mathcal{L}}{\partial \theta_0}\Big|_{\theta_0}$, $\frac{\partial \mathcal{L}}{\partial \theta_1}\Big|_{\theta_1 =}$

3. $$\left\{ \begin{aligned} \theta_0 &= \theta_0 - \alpha . \frac{\partial \mathcal{L}}{\partial \theta_0} \\ \theta_1 &= \theta_1 - \alpha . \frac{\partial \mathcal{L}}{\partial \theta_1} \end{aligned} \right\}$$ Update Rule / Eqn

   α → step-size

4. Repeate step 2 & 3 until convergence / fixed times $\boxed{50}$

$$\mathcal{L} = \frac{1}{2m} \sum_{i=1}^{m} \underbrace{\left(y^i - (\theta_0 + \theta_1 x^i) \right)}_{Z}{}^2$$

$\frac{dz}{dk} . \frac{dk}{d\theta_0}$

$\frac{d}{dz}(Z)^2 \Rightarrow 2Z$

$$\left\{ \frac{\partial \mathcal{L}}{\partial \theta_0} = \boxed{\frac{dz^2}{dz}} . \boxed{\frac{dz}{d\theta_0}} \right.$$

$$\underbrace{\frac{\cancel{2}}{\cancel{2}m} \sum_{i=1}^{m} \left(y^i - (\theta_0 + \theta_1 x^i) \right)^{\cancel{2}}}. \underbrace{\frac{d\left[\cancel{y^i} - \theta_0 - \cancel{\theta_1 x^i}\right]}{d\theta_0}}$$

$$\Rightarrow \frac{1}{m} \sum_{i=1}^{m} \left(y^i - \underbrace{(\theta_0 + \theta_1 x^i)}_{\hat{y}^i} \right) . 1$$

$$\checkmark \boxed{\frac{\partial \mathcal{L}}{\partial \theta_0} = \frac{1}{m} \sum_{i=1}^{m} (y^i - \hat{y}^i) . 1}$$

$$\frac{\partial \mathcal{L}}{\partial \theta_1} \Rightarrow \frac{2}{2m} \sum_{i=1}^{m} \left(y^i - (\theta_0 + \theta_1 x^i) \right) \times \frac{d}{d\theta_1}\left[\cancel{y^i} - \cancel{\theta_0} - \theta_1 x^i\right]$$

$\times \frac{d}{d\theta_1}\left[-\theta_1 x^i\right]$

$\times \boxed{x^i} \frac{d\,\theta_1}{d\theta_1}$

$$\checkmark \boxed{\frac{\partial \mathcal{L}}{\partial \theta_1} = \frac{1}{m} \sum_{i=1}^{m} (y^i - \hat{y}^i) . x^i}$$

$h_{\theta}(x) = 1.\theta_0 + \theta_1 \boxed{x}$

$$\boxed{\begin{aligned} \theta_0 &= \theta_0 - \alpha . \left[\frac{1}{m} \sum_{i=1}^{m} (y^i - \hat{y}^i) \right] \\ \theta_1 &= \theta_1 - \alpha \left[\frac{1}{m} \sum_{i=1}^{m} (y^i - \hat{y}^i) . x^i \right] \end{aligned}}$$