Simple Linear Regression

Supervised ML → Regression — O/P or Dependent feature → Continuous

Supervised ML → Classification — O/P " → Categorical

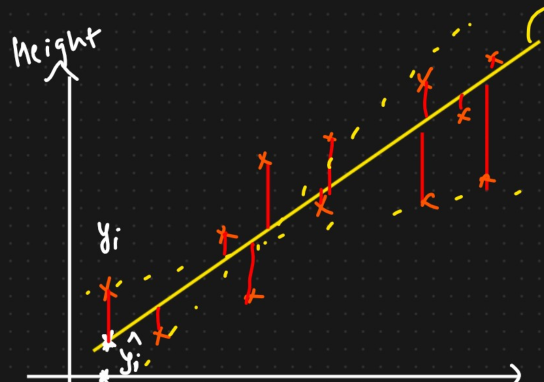
Dataset

Weight (Independent feature)	Height (Dependent feature, O/P)
74	170
80	180
75	175.5
—	—
—	—
—	—

TRAIN DATASET ↓ Model

new Weight → Model → Height (prediction)

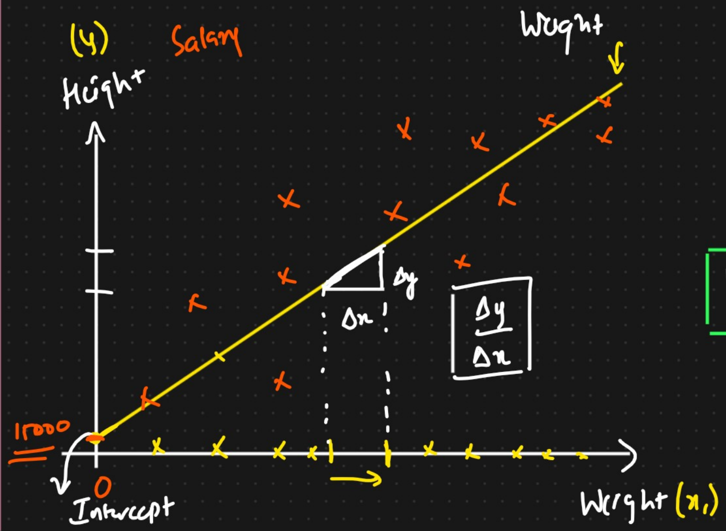
Model ↓ Acc ↑↑



→ best fit line → predictions

Σ Error ↓↓ Minimal

Geometrical Intuition



$\hat{y} = mx + c$

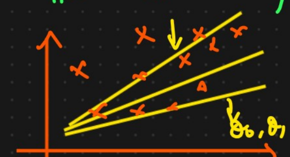
$\hat{y} = \beta_0 + \beta_1 x_1$

$$\boxed{h_\theta(x) = \theta_0 + \theta_1 x_1}$$

$\theta_0 \Rightarrow$ Intercept

$\theta_1 \Rightarrow$ Slope or Coefficient

Simple Linear Regression ⇑ Model ⇓ Hypothesis Testing



$\theta_0, \theta_1 \rightarrow$ change

Error ↓↓

Experience $x_1 \Rightarrow$ Datapoint

Cost function [Error].

$$J(\theta_0,\theta_1) = \frac{1}{n}\sum_{i=1}^{n}(\underset{\text{Actual}}{y_i} - \underset{\text{predicted}}{h_\theta(x)_i})^2$$ [Mean Squared Error]

$h_\theta(x)_i \uparrow \hat{y}_i$

n: no. of datapoints

y_i = Actual value

$h_\theta(x)$ = predicted value

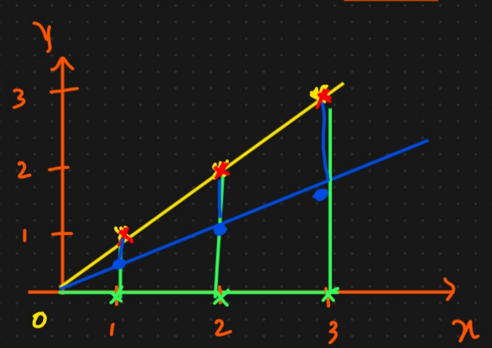
Final Aim [In order to get best fit line]

Minimize θ_0,θ_1 $\quad J(\theta_0,\theta_1) = \frac{1}{n}\sum_{i=1}^{n}(y_i - h_\theta(x)_i)^2 \downarrow\downarrow\downarrow$

Dataset

x	y	$h_\theta(x)$
1	1	
2	2	
3	3	

Minimization Error $\downarrow\downarrow$



$h_\theta(x) = \overset{0}{\theta_0} + \theta_1 x_1$

Lets consider $\theta_0 = 0$

$h_\theta(x) = \theta_1 x_1$

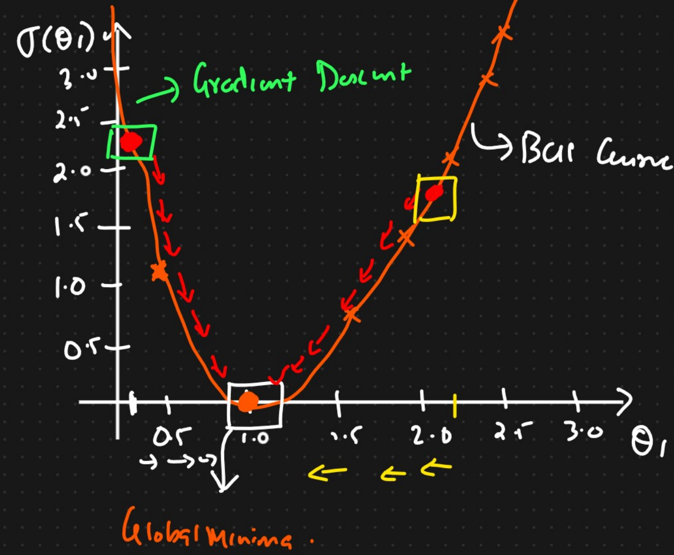
		Let $\theta_1 = 1$	Let $\theta_1 = 0.5$	Let $\theta_1 = 0$
predicted	$x=1$	$h_\theta(x) = 1(1) = 1$	$h_\theta(x) = 0.5$	$h_\theta(x) = 0$
	$x=2$	$h_\theta(x) = 1(2) = 2$	$h_\theta(x) = 1.0$	$h_\theta(x) = 0$
	$x=3$	$h_\theta(x) = 1(3) = 3$	$h_\theta(x) = 1.5$	$h_\theta(x) = 0$

Cost fn $\theta_1 = 1$

$$J(\theta_1) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - h_\theta(x)_i\right)^2$$

$$= \frac{1}{3}\left[(1-1)^2 + (2-2)^2 + (3-3)^2\right]$$

$$J(1) = 0$$



Cost fn $\theta_1 = 0.5$

$$J(\theta_1) = \frac{1}{3}\left[(1-0.5)^2 + (2-1)^2 + (3-1.5)^2\right]$$

$$J(\theta_1) = \underline{\underline{1.16}}$$

Cost fn $\theta_1 = 0$

$$J(\theta_1) = \frac{1}{3}\left[(1-0)^2 + (2-0)^2 + (3-0)^2\right]$$

$$J(\theta_1) = 4.66$$

Convergence Algorithm

{Optimize the change of θ_0, θ_1 to Global Minima}.

Repeat until Convergence

$\{$ $j = 0, 1$

$$\theta_j := \theta_j - \alpha \boxed{\frac{\partial}{\partial \theta_j} J(\theta_j)} \Rightarrow \text{Derivative}$$

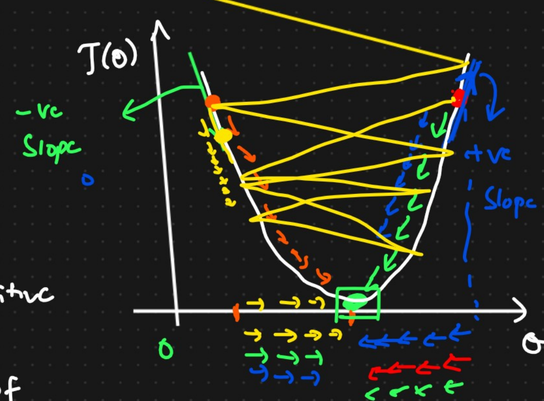
$\Downarrow$

Slope of

$\}$

$\Downarrow$

$\boxed{\alpha = 0.01}$



a point

Learning Rate

① $\theta_1 = \theta_1 - \alpha(-ve)$

$\theta_{1new} = \theta_{1old} + (value)$

$\boxed{\theta_{1new} >> \theta_{1old}}$

Learning Rate ⇒ speed of Convergence

② $\theta_{1new} = \theta_{1old} - \alpha(+ve)$

$\theta_{1new} = \theta_{1old} - (value)$

$\boxed{\theta_{1new} << \theta_{1old}}$

Conclusion

Repeat until Convergence

{

$\theta_j : \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$

}

$J(\theta_0, \theta_1) = \frac{1}{n} \sum_{i=1}^{n} (y_i - h_\theta(x)_i)^2$

⇓

Mean Square Error