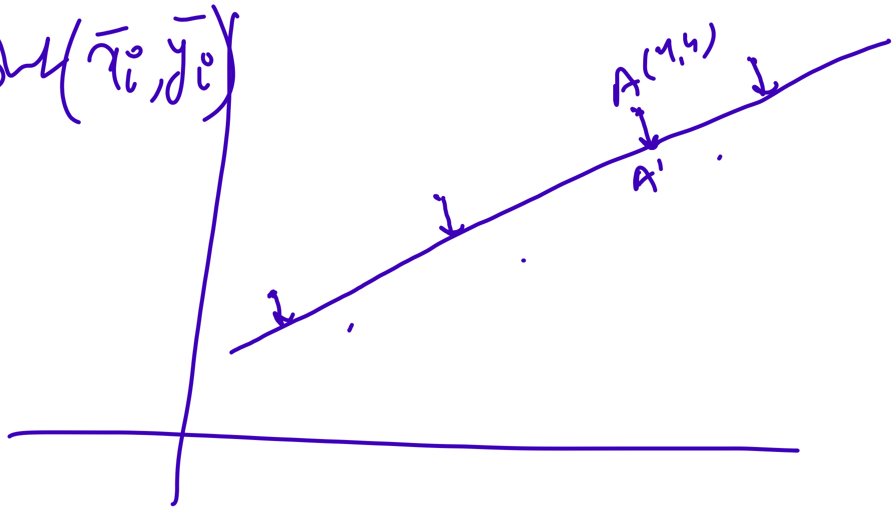
17/11/21

A : Actual Point (x_i, y_i)

A' : Prediction Point $(\bar{x}_i, \bar{y}_i)$



Equation of line

$$\bar{y}_i = m\bar{x}_i + C$$

$\bar{y}_i$ ← Prediction

$\bar{x}_i$ ← Testing point (400)

unavailable : m & C

find m : slope ✓

c : intercept ✓

Linear Regression — ① Hypothesis function
② Cost function

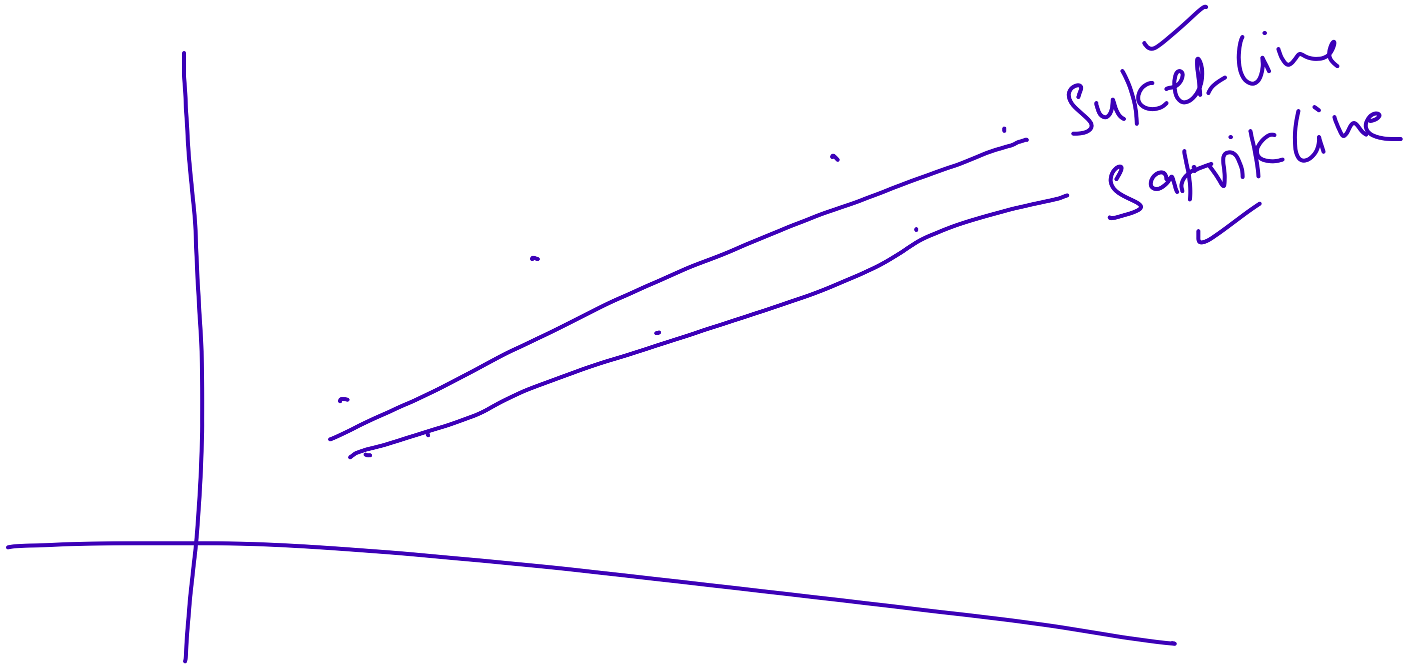
① Hypothesis function -

$y = mx + c$

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

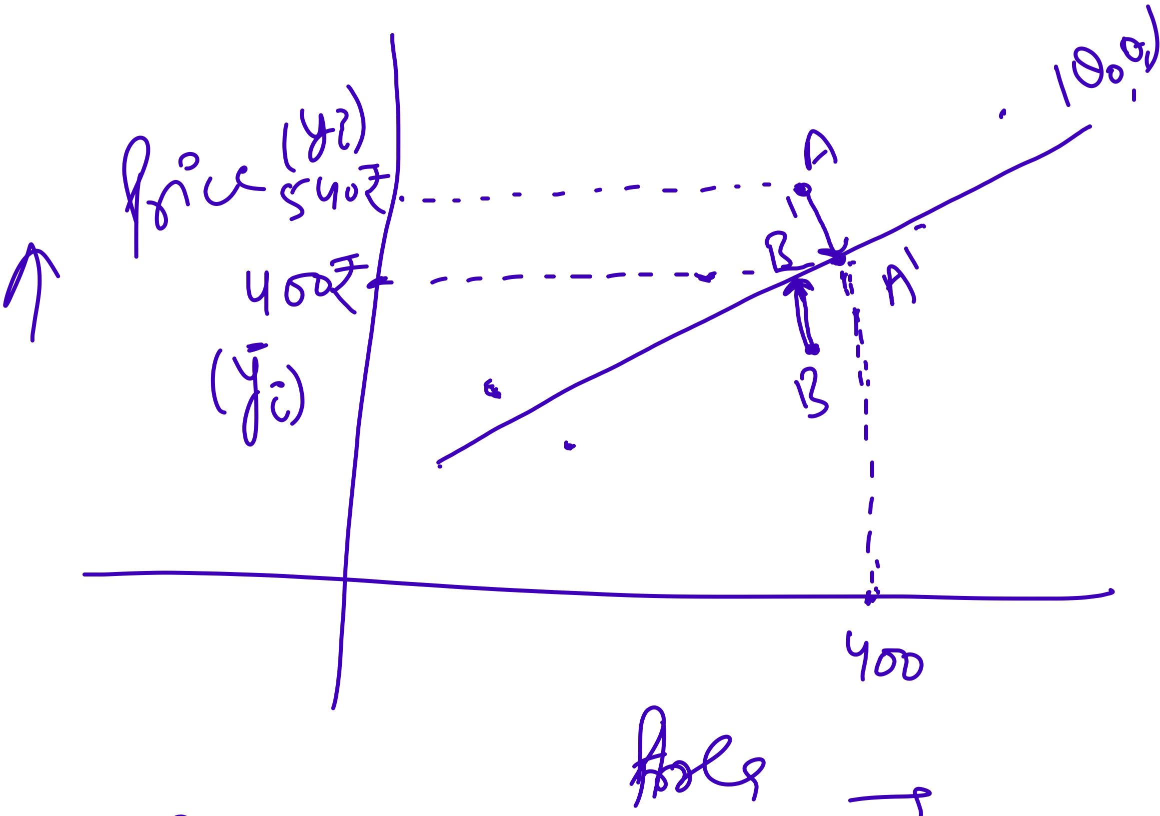
$m(\theta_0)$, $c(\theta_1)$

② Cost function :- $J(\theta_0, \theta_1)$



Cost function : MSE

MSE : Minimum squared distance for linear Regression.



n - Points

Prediction line Equation

$$\bar{y}_i = h_\theta(x) = \theta_0 x + \theta_1$$

$$\boxed{J(\theta_0, \theta_1) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - \bar{y}_i)^2}$$

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

eg

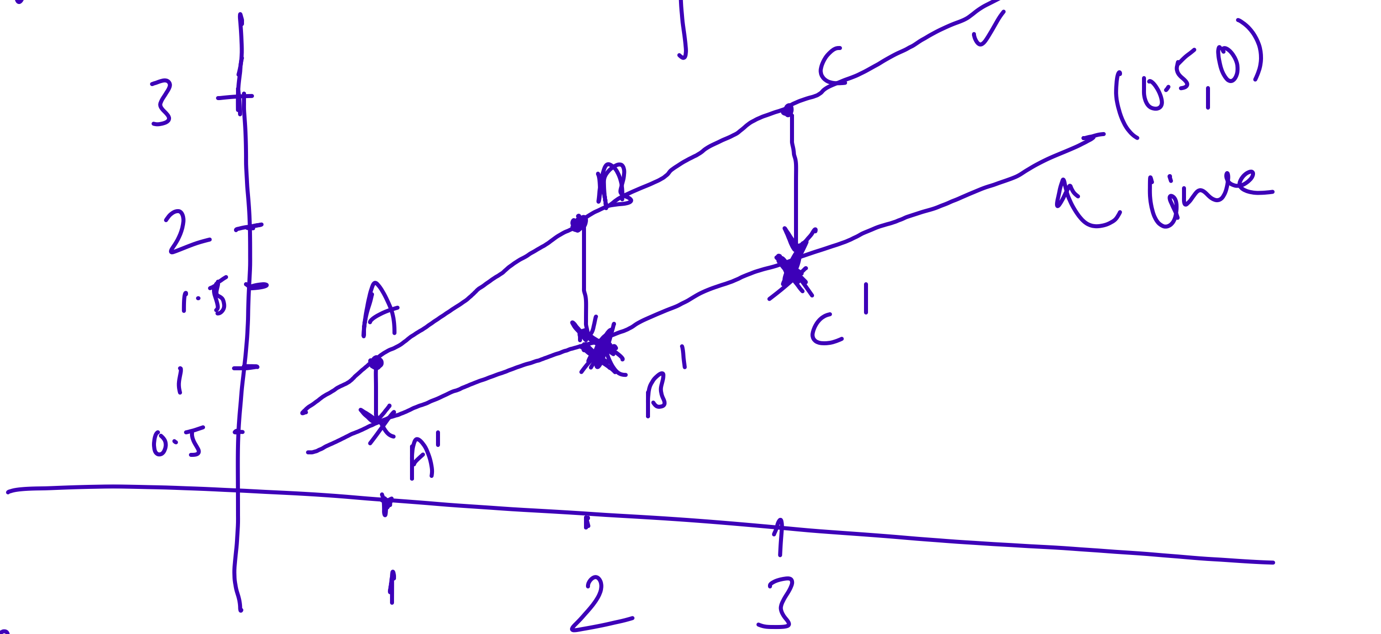
Dataset

Area (X_i) (sq ft)	Price (Y_i) (Rupee)
1	1
2	2
3	3

What is Testing data

Area = 4 sqft

Price = ?



Case I :-

Assume Consider $\theta_0 = 0.5$, $\theta_1 = 0$

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

$$= 0.5 \times x + 0$$

$$h_\theta(x) = \underline{0.5x}$$

$x=1$, $h_\theta(x) = 0.5 \times 1 = 0.5$ $(1, 0.5)$

$x=2$, $h_\theta(2) = 0.5 \times 2 = 1$ $(2, 1)$

$x=3$, $h_\theta(3) = 0.5 \times 3 = 1.5$ $(3, 1.5)$

Cost function :-

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

$$J(0.5, 0) = \frac{1}{2 \times 3} \left[(0.5 - \underset{y_i}{1})^2 + (1 - \underset{y_i}{2})^2 + (1.5 - 3)^2 \right]$$

$$= \frac{1}{6} \left[(-0.5)^2 + (-1)^2 + (-1.5)^2 \right]$$

$$= 0.58$$

~~Case~~ -

~~θ_0~~ Case : Ideal Case

$\theta_0 = 1$, $\theta_1 = 0$

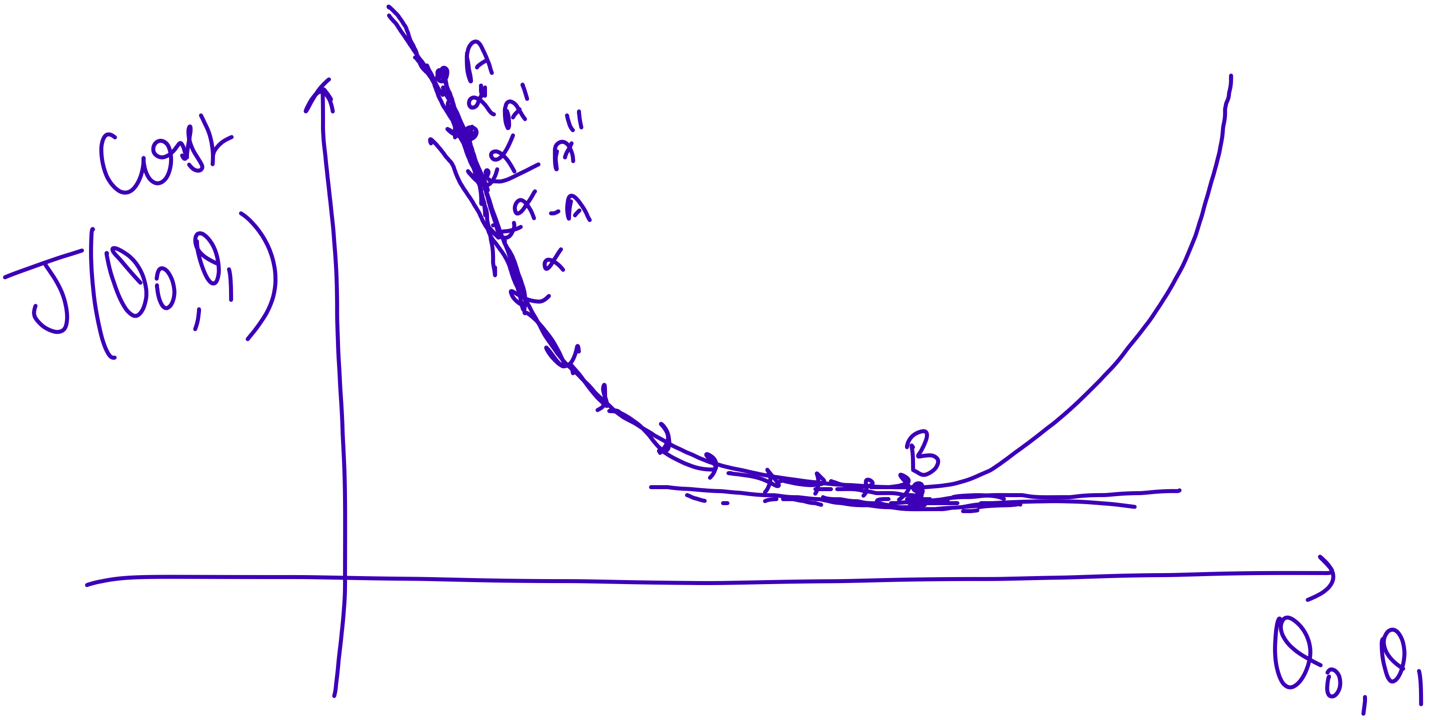
$$h_\theta(x) = \theta_0 x + \theta_1$$
$$= 1 \times x + 0$$
$$h_\theta(x) = x$$
$$J(\theta_0, \theta_1) = \frac{1}{2 \times 3}\left[(1-1)^2 + (2-2)^2 + (3-3)^2\right]$$
$$= 0$$
cost = 0,

<u>Problem</u> :- How many value will consider for (θ_0, θ_1).

Where will we stop?

Gradient — slope

descent — going down



α : learning rate

Algorithm for Gradient descent
(one line algo).

repeat untill Convergence {

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

}

$\Downarrow$

expansion (θ_0, θ_1)

repeat untill convergence {

$$\theta_0 := \theta_0 - \alpha * \frac{\partial}{\partial \theta_0} [J(\theta_0, \theta_1)]$$

$$\theta_1 := \theta_1 - \alpha * \frac{\partial}{\partial \theta_1} [J(\theta_0, \theta_1)]$$

Linear Regression using Gradient descent :-

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

$$\frac{\partial}{\partial \theta_0} [J(\theta_0, \theta_1)] = \frac{1}{2n} \sum_{i=1}^{n} (\theta_0 x_i + \theta_1 - y_i)^2$$

$$= \frac{2}{2n} \sum_{i=1}^{n} (\theta_0 x_i + \theta_1 - y_i) \times x_i$$

✓ $$\frac{\partial}{\partial \theta_0} [J(\theta_0, \theta_1)] = \frac{1}{n} \sum_{i=1}^{n} (h_\theta(x) - y_i) \times x_i$$

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

~~Algo~~ LR using GD

repeat until convergence {

$$\theta_0 := \theta_0 - \alpha * \frac{1}{n}\sum_{i=1}^{n}(h_\theta(x) - y_i)x_i$$

$$\theta_1 := \theta_1 - \alpha * \frac{1}{n}\sum_{i=1}^{n}(h_\theta(x) - y_i)$$

}

$:=$

☆ if not using GD

→ $\theta_0 = 0.5,\ \theta_1 = 0$

→ slope $= \frac{(x_i - x_{mean})(y_i - y_{mean})}{(x_i - x_{mean})^2}$

~~intecep~~ ~~intecept~~

✓ intercept = ymean − (slope) * xmean

✓ $m = \dfrac{y_2 - y_1}{x_2 - x_1} \Rightarrow \dfrac{y_i - y_{mean}}{x_i - x_{mean}}$

✓ $c = y_{mean} - m * x_{mean}$

$: =$

$a = 2$

$a = a + 1$

$a = 4$

$a = 3$

$a := a + 1$

$a : 4$

temp