Linear Regression Machine Learning Algorithms

Data Scientist → Linear Regression

Agenda

1. What problem we are solving
2. Geometric Intuition
3. Mathematical Intuition

①

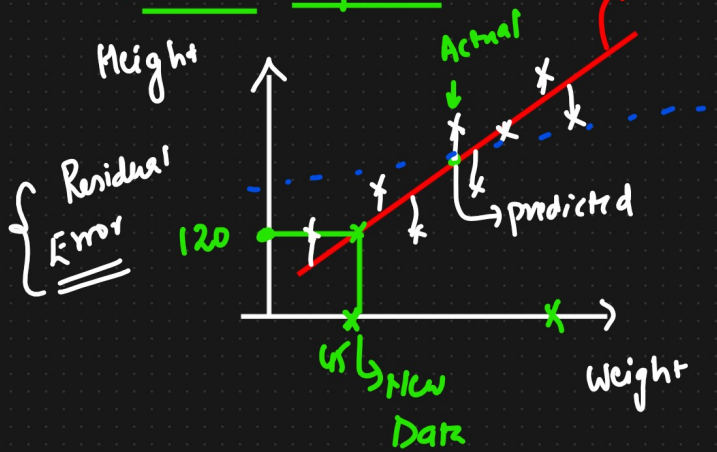
- Machine Learning
  - Supervised
    - Regression
    - Classification
  - Unsupervised

Linear Regression → Regression

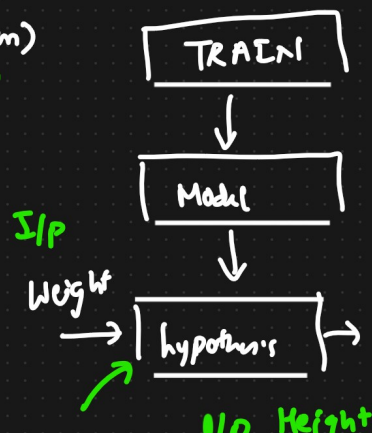
Continuous

Aim : To find best fit line

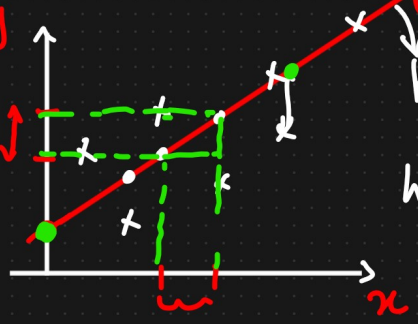
Linear Regression ✓



Weight (kg)	Height (cm)
73	150
63	155
78	170
82	165



Maths {Basic Maths}



$$\hat{y} = mx + c$$

m = Slope or Coefficient

c = Intercept

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

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \cdots \theta_n x_n$$

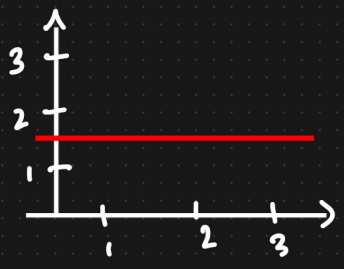
Andrew NG

Hypothesis {y is a linear function of x}

$\rightarrow h_\theta(x) = \theta_0 + \theta_1 x_1$ (θ_1 → coefficient, θ_0 → intercept)

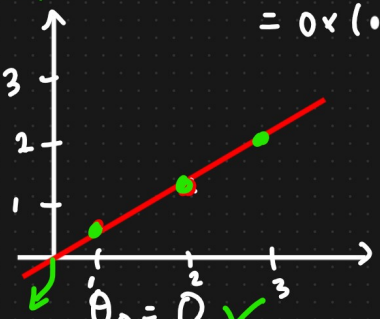
$= 0 \times (0.5)\,1 = 0.5$

$= 0 \times (0.5)\,2 = 1$



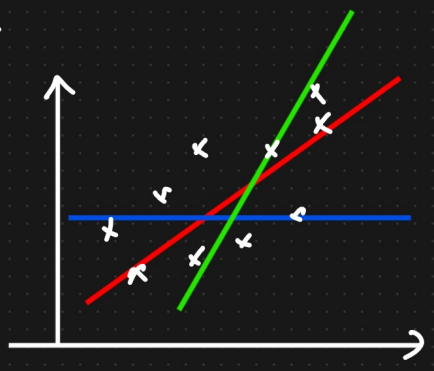
$\theta_0 = 1.5$

$\theta_1 = 0$



$\theta_0 = 0$ ✓

$\theta_1 = 0.5$



$m \rightarrow$ No. of datapoints

Solve {Cost function}

Minimize θ_0, θ_1

$$\sum_{i=1}^{m} \frac{1}{2m} \left(h_\theta(x)^{(i)} - y^{(i)} \right)^2 \downarrow\downarrow$$

$$\frac{\partial}{\partial x}(x^2) = 2x^{2-1} = 2x$$

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x)^{(i)} - y^{(i)} \right)^2 \downarrow\downarrow\downarrow$$

↳ Squared Error Function

Hypothesis

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

{When $\theta_1 = 1$} (1,1) (2,2) (3,3)

Lets $\theta_0 = 0$

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

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

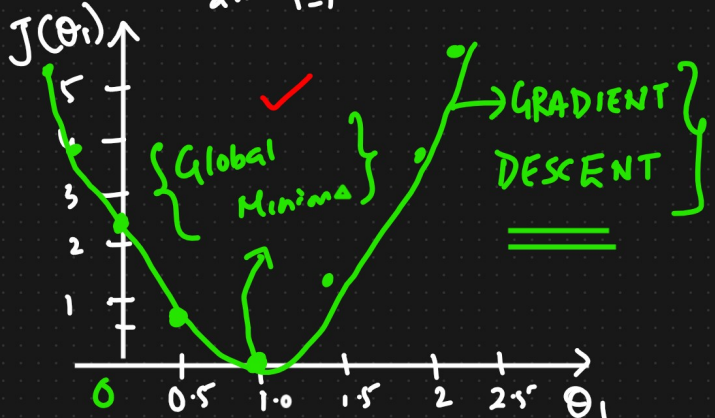
$h_\theta(x) = 1 \times 3$

When $\theta_1 = 0.5$

When $\theta_1 = 0.0$

Cost function

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



$$J(\theta_1) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x)^{(i)} - y^{(i)}\right)^2$$ When $\theta_1 = 1$

$$= \frac{1}{6}\left[(0)^2 + (0)^2 + (0)^2\right]$$

$$= 0$$

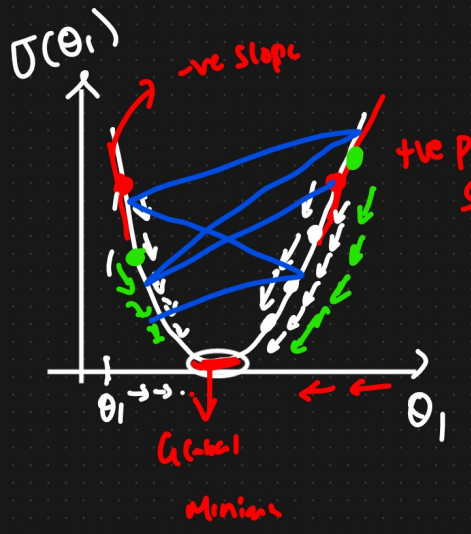
$J(\theta_1)$ When $\theta_1 = 0.5$

$$= \frac{1}{6}\left[(0.5-1)^2 + (1-2)^2 + (1.5-3)^2\right]$$

$$= \frac{1}{6}(3.5) = \approx 0.58$$

$$J(\theta_1) = \frac{1}{6}\left[(0-1)^2 + (0-2)^2 + (0-3)^2\right]$$

$$= \frac{14}{6} \approx 2.3$$



$\theta_1 \rightarrow$ Update $\rightarrow$ find derivation or slope

$\alpha = 1000$

{ Learning Rate }

$\alpha = 0.01$

Repeat convergence theorem

$$\{\quad \theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j}(J(\theta_1)) \quad \}$$ (Slope)

$\theta_1 := \theta_1 - \alpha(+ve)$ } Decrease ✓

$\theta_1 := \theta_1 - \alpha(-ve)$ } Increase ✓ (Slope)

$\theta_1 = \theta_1 +$

Outline

1. Start with θ_0 & θ_1
2. Keep changing θ_0, θ_1 to reduce $J(\theta_0, \theta_1)$ until we reach near global Minima.
3. Convergence Theorm
   {

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

$j := 0$ and 1

}