Machine Learning Algorithms

ML Algorithms → Supervised ML → Regression, Classification

ML Algorithms → Unsupervised ML → Clustering.

Regression

House price prediction ⇓

No. of rooms, Total Size, Location (Independent) → o/p Price (Dependent)

x_i →

Continuous

1.5m $y_i = y'$

1.65m

1. Linear Regression [Regression]
2. Ridge Regression ["]
3. Lasso " [ ]
4. ElasticNet ["]
5. Logistic Regression [classif.]

Classification

Binary or Multiclass Classificat.

No. of Play hour	No. of sleep hours	No. of study	Pass/Fail o/p
5hrs	7hrs	1hr.	Fail
–	–	–	–

Decision Tree

Random Forest

AdaBoost

Xgboost → Classification, Regression

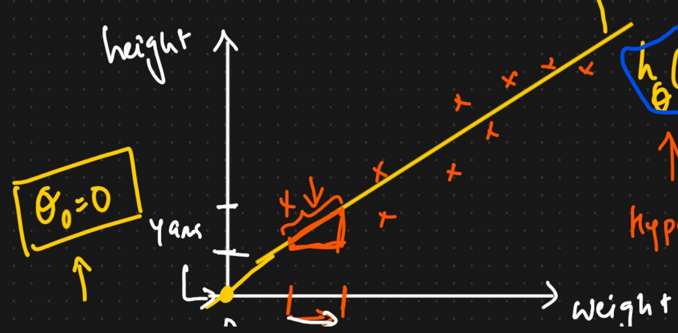
1. Simple Linear Regression

$y = mx + c$

$y = \beta_0 + \beta_1 x_1$

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

Hypothesis



Weight (Independent) — Height (o/p)

TRAIN DATASET (Sample of DATASET) → Model (Training)
↓

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

⇓

Equation of a straight line

New weight → Hypothesis → Height

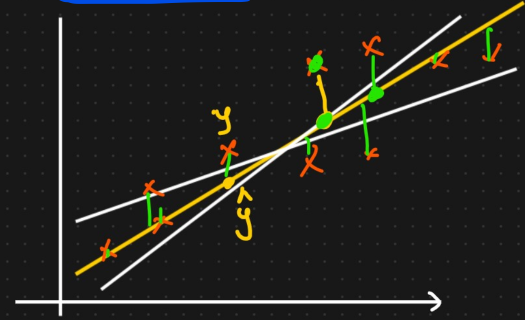
⇑

$y = mx + c$

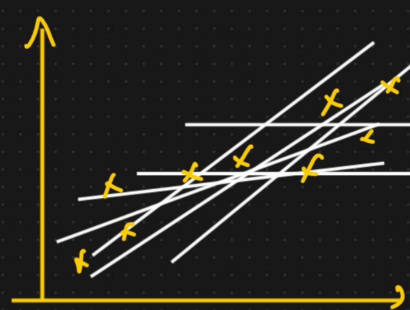
Aim

θ_0 = Intercept

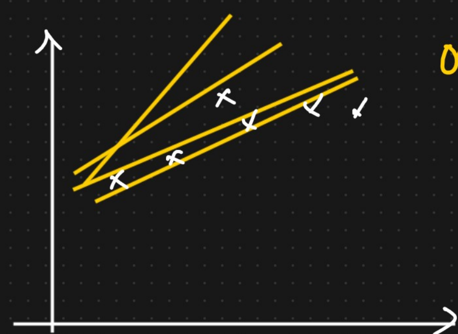
θ_1 = Coefficient or Slope



less ERROR



No.



Optimization

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

$$\boxed{\hat{y} - y}$$

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

Cost function

Datapoint

$\hat{y}$

Error

Mean Squared Error

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

→ MSE

What we need to Solve (AIM) Final Goal

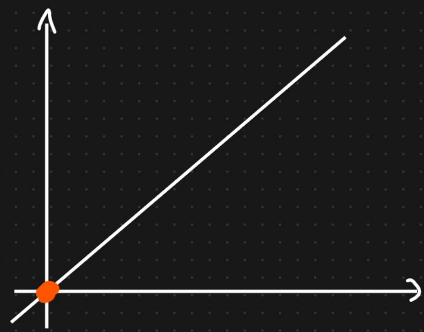
$$\underset{\theta_0, \theta_1}{\text{minimize}}\ J(\theta_0, \theta_1) \quad \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

Change θ_1 value

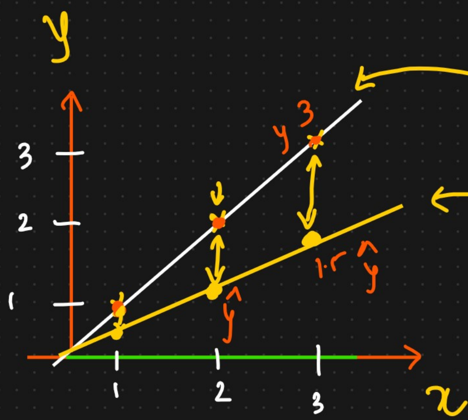
① $h_\theta(x) = \theta_0 + \theta_1 x$

if $\theta_0 = 0$

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



0.5×3
0.5×2
0.5×1
$\downarrow$
$h_\theta(x) = \theta_1 x$



if $\theta_1 = 1$

if $\theta_1 = 0.5$

if $\theta_1 = 0 \rightarrow$

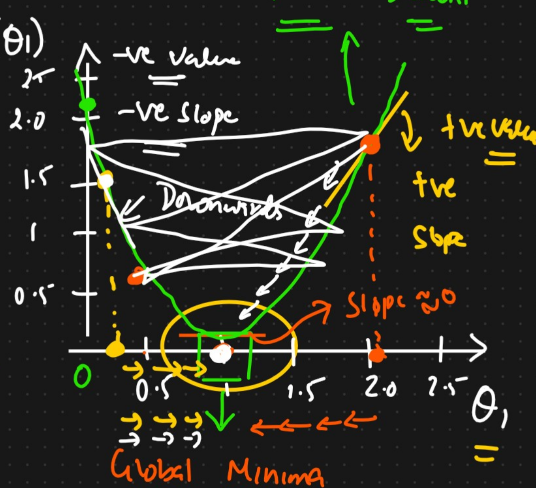
$\alpha R = 100\,000$

MSE

$\alpha R = 0.001$

↗ Cost fn

Gradient Descent



$$J(\theta_1) = \frac{1}{2m} \sum_{i=1}^{3} \left(h_\theta(x)^{(i)} - y^{(i)}\right)^2$$

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

$$= 0$$

$$\Rightarrow \frac{1}{2m} \sum_{i=1}^{m} (\hat{y} - y)^2$$

$\Downarrow$

Quadratic Equation

$\Downarrow$

$$\boxed{ax^2 + bx + c}$$

$$J(\theta_1) = \frac{1}{6} \left[(0.5-1)^2 + (1-2)^2 + (1.5-3)^2\right]$$

$$J(\theta_1) = 0.58$$

$$J(\theta_1) = \frac{1}{6} \left[(0-1)^2 + (0-2)^2 + (0-3)^2\right] \approx \boxed{2.3}$$

Gradient Descent Convergence Algorithm

repeat until convergence

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

$\Rightarrow$ Coefficient updation formula

$LR = 0.01$

$\downarrow$

LR

Calculate Slope

$$\Rightarrow \theta_{new} = \theta_{old} - \alpha \boxed{\frac{\partial (J(\theta_{old}))}{\partial \theta_{old}}}$$

$$\Rightarrow \theta_{new} \approx \theta_{old} - \alpha\,(-ve)$$

$$\theta_{new} \approx \theta_{old} + \text{+ve value}$$

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

$\Rightarrow$

$$\theta_{new} = \theta_{old} - \eta\,(\text{+ve value})$$

$$= \theta_{old} - (\text{+ve})\text{ value}$$

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

Gradient Descent [MSE]

$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad y$

$$h_\theta(x) = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \theta_4 x_4 + \theta_0$$

$\Downarrow$

Multi Linear Regression

Advantages

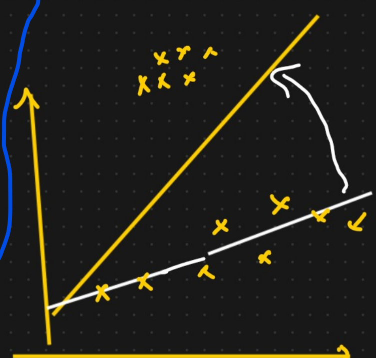
① There is one global Minima $\leftarrow$

② It is differentiable $\leftarrow$

Disadvantage

① Not, Robust to outliers

$$\frac{1}{2n}\sum_{i=1}^{n} \underbrace{(y-\hat{y})}_{\text{Error}}{}^{\overset{\text{Error}}{2}}$$



x^2

$2x$

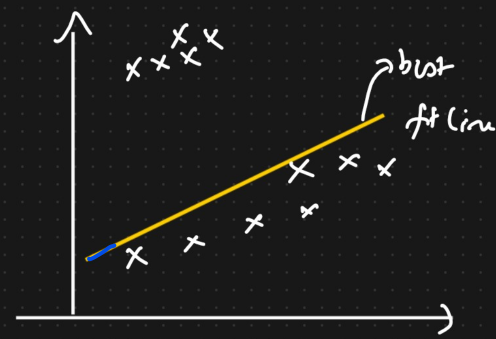
{MSE penalizes the Error}

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

② Mean Absolute Error

$\hat{y} = h_\theta(x^i)$

$$J(\theta_0,\theta_1) = \frac{1}{m}\sum_{i=1}^{m}|\hat{y}-y|$$

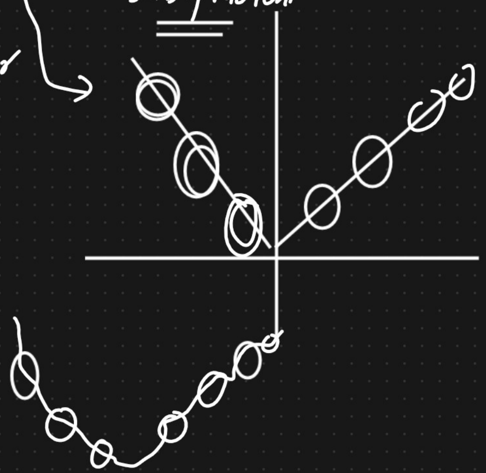


5 number Summary → Remove the outlier

⇓

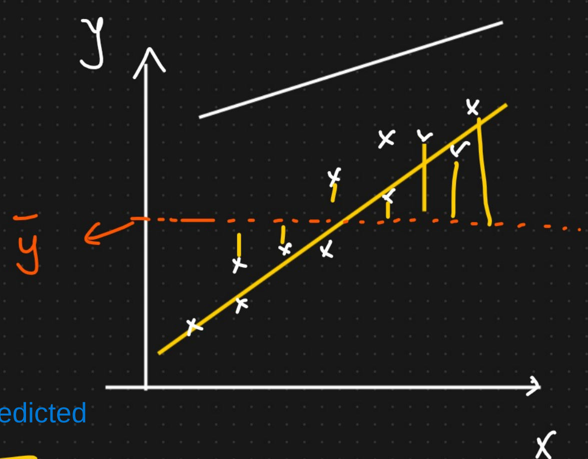
MSE

Subgradient



Performance Metrics

R^2 and Adjusted R^2



$$R^2 = 1 - \frac{SS_{Res}}{SS_{total}}$$

$\boxed{-ve}$??

Predicted

$\boxed{0 \text{ to } 1}$

↑ Error

$$= 1 - \frac{\sum(y_i-\hat{y})^2}{\sum(y_i-\bar{y})^2}$$

→ less

→ High

Average

= 1 - Small number

= Higher Number

Adjusted R^2

House pricing

No. of room	Total size	Location	Gender	Price
—	—		—	
—	—		—	

$R^2 = 89\% \Rightarrow 0.89$

$R^2 \approx 0.95$

$R^2 \approx 0.96$

Adjusted R^2

$R^2 = 0.87$

$R^2 = 0.89$

Adjusted $R^2 = 0.85$

Adjusted R^2

$$R^2_{\text{adjusted}} = 1 - \left[\frac{(1-R^2)(N-1)}{\boxed{N-p-1}}\right]$$

denominator ↓ smaller ↓↓↓

P = No. of predictors

N = No. of data points

$p=2$ $\quad R^2 = 90\%$ $\quad$ Adj $R^2 = 86\%$ (low)

$p=3$ $\quad R^2 = 91\%$ $\quad$ Adj $R^2 = 82\%$

$p=2 \qquad p=3$

$R^2 = 0.9 \qquad p=2 \qquad n=100$

$$1 - \left[\frac{(1-0.9) \times (100-1)}{100-2-1}\right]$$

$$= 1 - \left[\frac{0.1 \times 99}{97}\right]$$

$$= 1 - \frac{9.9}{97} = 0.8979$$

$p=3$, $R^2 = 0.91$

$R^2 = 0.90$

Adjus $= 0.8979$.

$$R^2 = 1 - \left[\frac{(1-0.91)(100-1)}{100-3-1}\right]$$

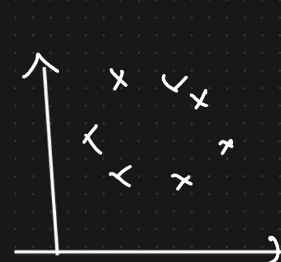
>> Adjusted R^2

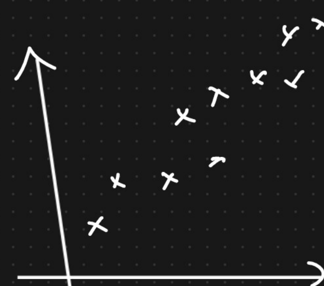
$$= 1 - \left[\frac{(0.09)\ (99)}{96}\right] = 0.9071$$

Adjusted R^2 & $\boxed{R^2}$ $\Rightarrow$

$\Downarrow$

Difference higher





Linear Regression

① There is a linear Relationship with x & y

② Independent feature should have Normal Distribution

③ Always take care of multicollinearity.

③ Homoscedasticity =. Same variance

$\boxed{f_1} \Leftrightarrow f_2$

④ Feature Scaling Required ?? ⟹ Yes

⑤ { Heteroscedasticity ?? → Assignment ??

Gradient = RMSE vs MSE }