3.GradientDescent

Tuesday, August 2, 2022 12:06 PM

Min-Max Scaler

10 — 100

20 60 30 90 80 100 50 10

max = 100
min = 0

min = 0
max = 1

0 - 1

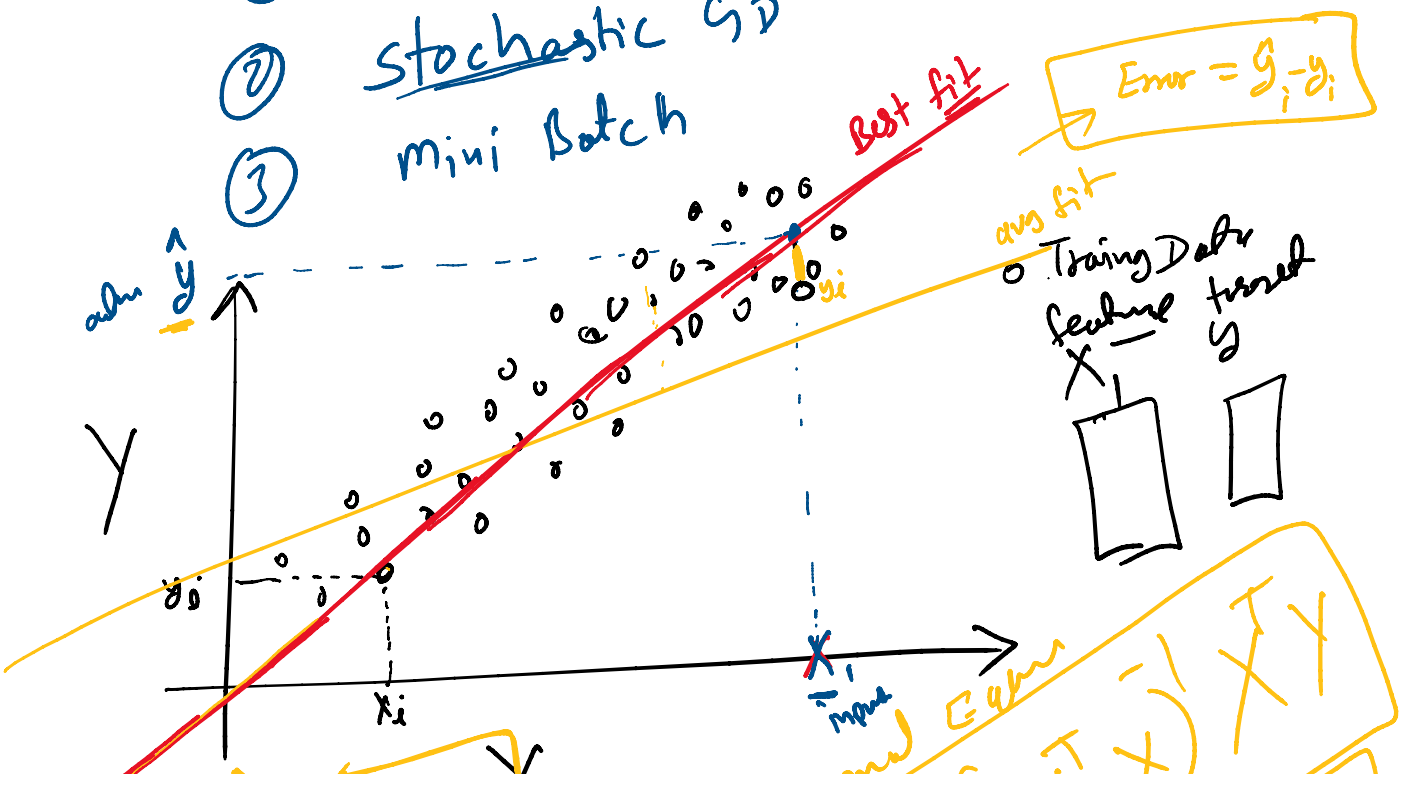
$$X_{scale} = \frac{X_i - min(X)}{max(X) - min(X)}$$

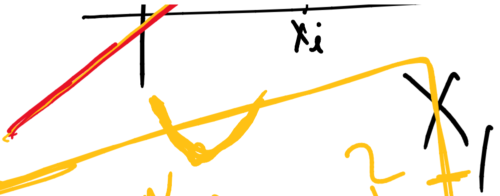
0.2 0.6 0.3 0.9 0.8 1 0.5 0 0.1

Stochastic → Random

Gradient Descent

1. Batch GD / Full GD
2. Stochastic GD
3. Mini Batch





$$MSE = \frac{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2}{N}$$

$$\hat{y}_i = \theta_1 x_1 + \theta_0$$

Normal Eqn

$$\theta = (X^T \cdot X)^{-1} \ldots$$

$$MSE(\theta) = \frac{\sum_{i=1}^{N} (x_i^T\theta - y_i)^2}{N}$$

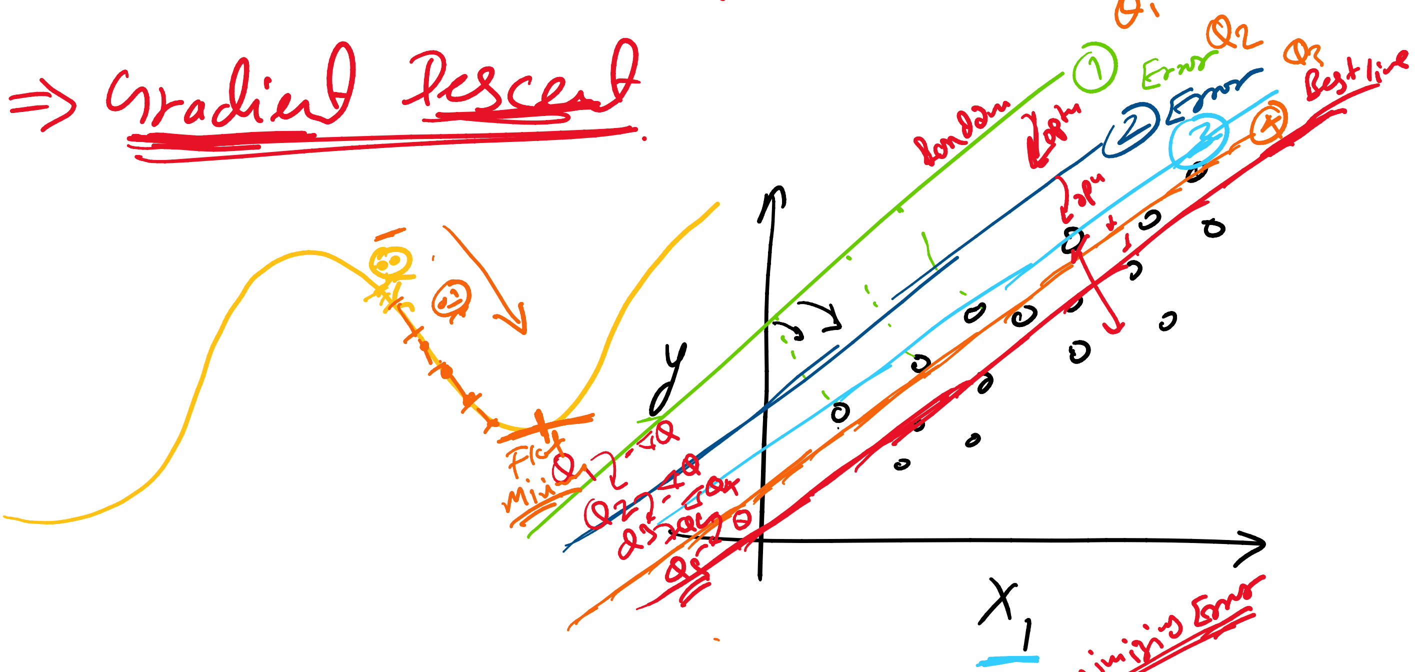
Best optimum

Best line

partial der

cost

⇒ **Gradient Descent**



(1)	Error	= 100
(2)	Error	= 500
(3)	Error	= 200
(4)	Error	= 20 → accept

① prediction function

$$\hat{y} = \theta_0 + \theta_1 \cdot x_1$$

$\hat{y}$ = weight (Predict)

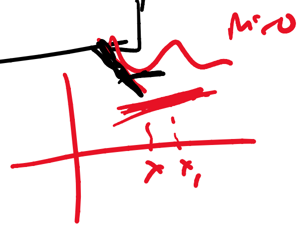
x_1 = heoine

θ_0, θ_1 = Posame

② Cost function $J(\theta$ ...

② Cost function $C\theta$

$$C\theta = \frac{\sum_{i=1}^{N} \left(\hat{\theta}_0 + \hat{\theta}_1 x_1^{(i)2} - y^{(i)}\right)^2}{N}$$

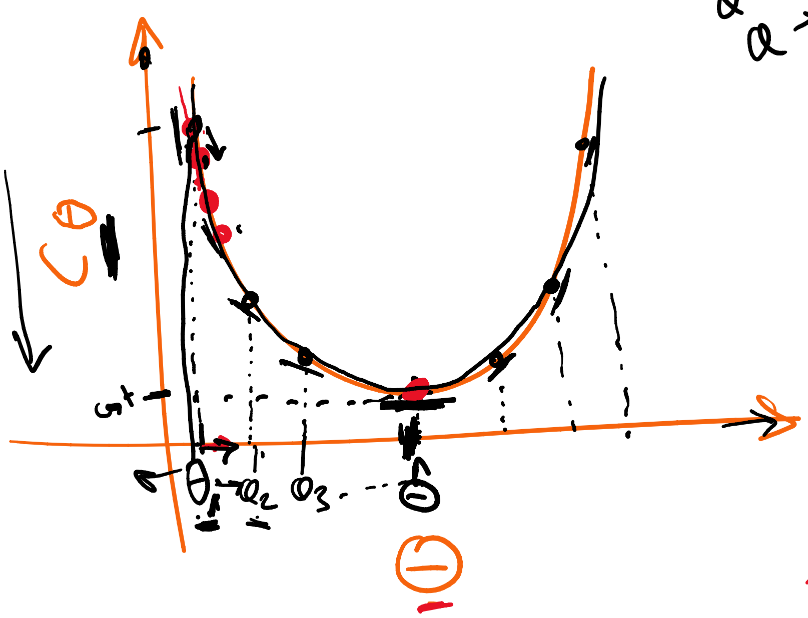


③ Performance Matrix

R2-Score, RMSE, MAE

$\nabla\theta = 0$

$\theta \rightarrow$ increase



optimisation Algo

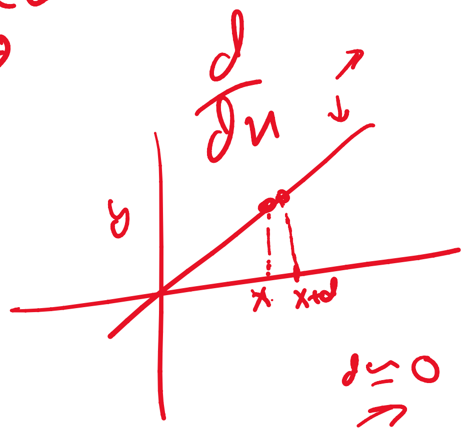
derivates → rate of change given small step on a fn

① initialise θ randomly

② take partial derivative to cost func $C\theta$ to find slop or steepness of $C\theta$ wrt θ

θ_i

$$\nabla_\theta C\theta = \frac{\partial}{\partial\theta} C\theta$$



③ make a decent step to minimise the $C\theta$

$d = 0$

$$\theta_{i+1} = \theta_i - \nabla_\theta C\theta$$

[illegible] reach to

④ Repeat this until algorithm reach to global minima

Hyper-Parameter

eta

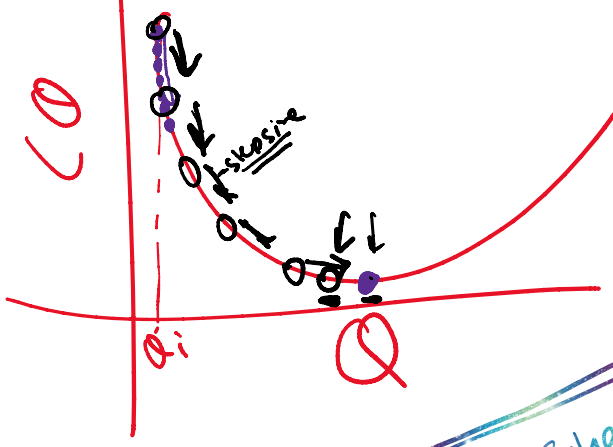
Batch

① Learning Rate → (η)

② no of iterations

$$\theta_{i+1} = \theta_i - \eta \cdot \nabla_\theta(\theta)$$

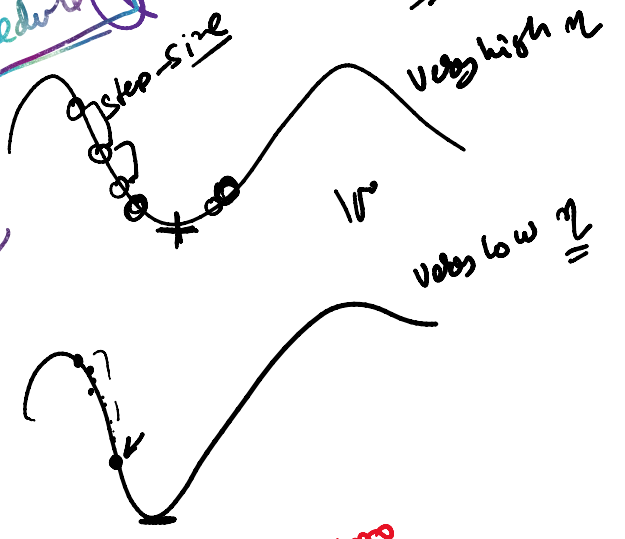
Training Data



Best fit → specific generalization

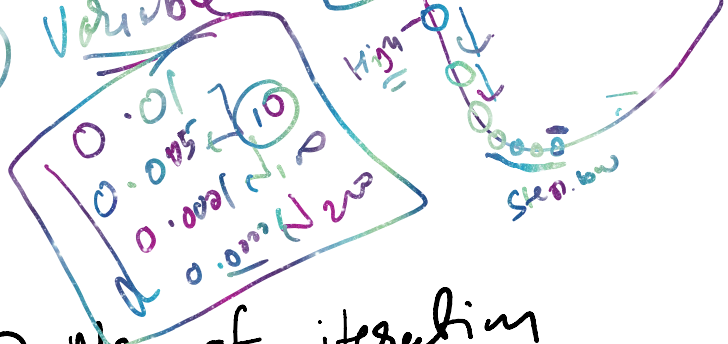
Good fit → small $\nabla\theta$

$\eta \nabla\theta$ → High



① Fixed →

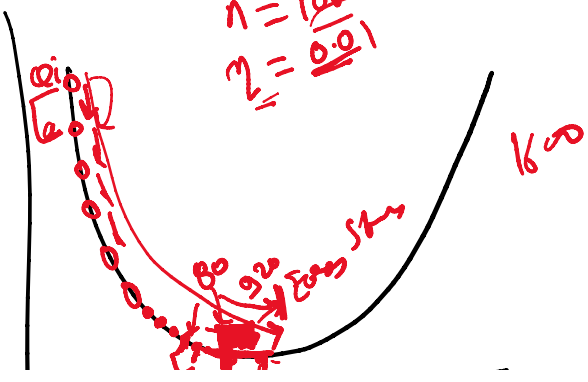
② Variable →



① No of iterations

② Learning Rate

③ Tolerance [Early stopping]



③ Tolerance

tolrance = $\theta_i - \theta_{i+1}$

0.01

$$MSE = \frac{1}{2m} \sum_{i=1}^{m} (mx_i + c - y_i)^2$$

(i) derive wrt to m

$$\frac{\partial}{\partial m} MSE = \frac{1}{2m} \frac{\partial}{\partial m} \sum_{i=1}^{m} (mx_i + c - y_i)^2$$

$$\nabla_m MSE = \frac{1}{N} \sum_{i=1}^{N} (mx_i + c - y_i) \cdot x_i$$

$$\nabla_c MSE = \frac{1}{N} \sum_{i=1}^{N} (mx_i + c - y_i)$$

$m_{i+1} = m_i - \nabla_m$

$c_{i+1} = c_i - \nabla_c$

$\theta - \eta \cdot \nabla \theta$