GRADIENT DESCENT

SWETHA C S

What is Gradient Descent

- Gradient descent is an optimization algorithm.
- It aims at minimizing the loss or cost function by iteratively updating the model's parameters in the opposite direction of gradient.

Cost function of Linear Regression

$$(1/n)\sum_{i=1}^{n}(yi - \hat{y})^{2}$$

Algorithm

STEP 1 : Take the partial derivative of loss function for each parameter m and b.

$$L = \sum_{i=1}^{n} (y_i - (mx_i + b))^2$$

Partial derivate of L wrt m and b

$$\partial L/\partial b = -2\sum(y_i - mx_i - b)$$

$$\partial L/\partial m = -2\sum(y_i - mx_i - b)(x_i)$$

- Step 2 : Assume random values for parameters initially

  m=1 and b=0
- Step 3 : Plug the parameter values into the gradient.
- Step 4 : Calculate the step size

$$step\_size = \eta * slope$$

- Step 5 : Calculate new parameters

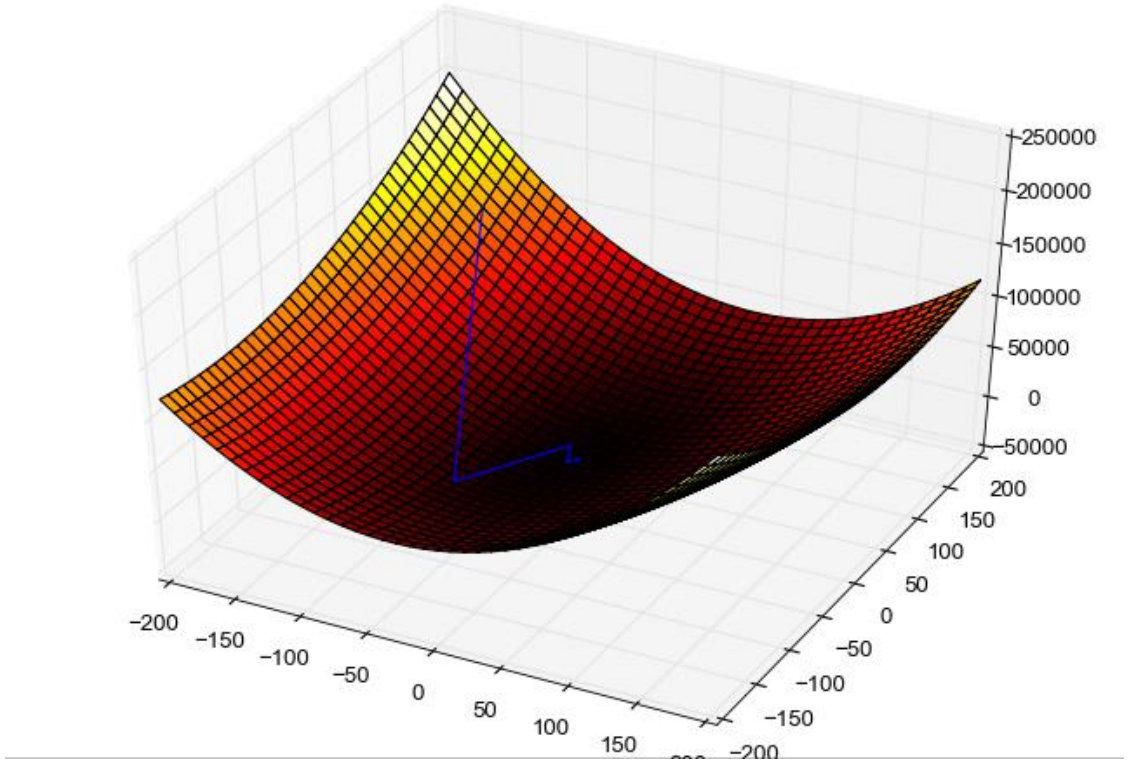
New_parameter = old_parameter - step_size.

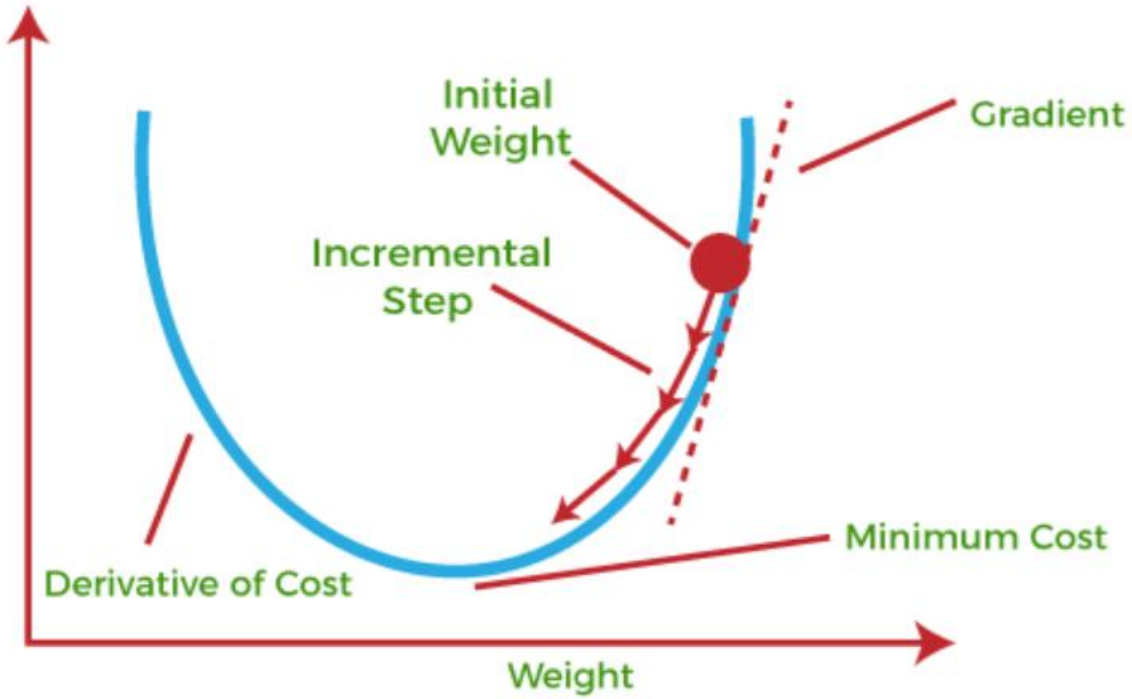
$$b_{new} = b_{old} - step\_size$$

$$m_{new} = m_{old} - step\_size$$

- Step 6 : Go to step 3 and repeat until step size is very small or you reach max epochs.

Global minima




Initial
Weight
Gradient
Incremental
Step
Minimum Cost
Derivative of Cost
Weight

Types of Gradient Descent

- Stochastic Gradient Descent
- Batch Gradient Descent
- Mini-batch Gradient Descent

QA