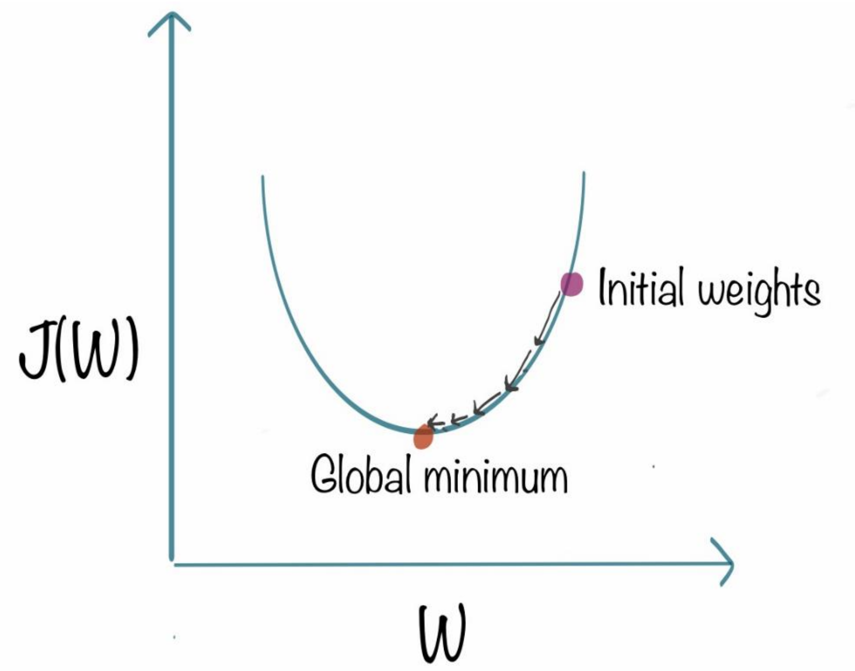
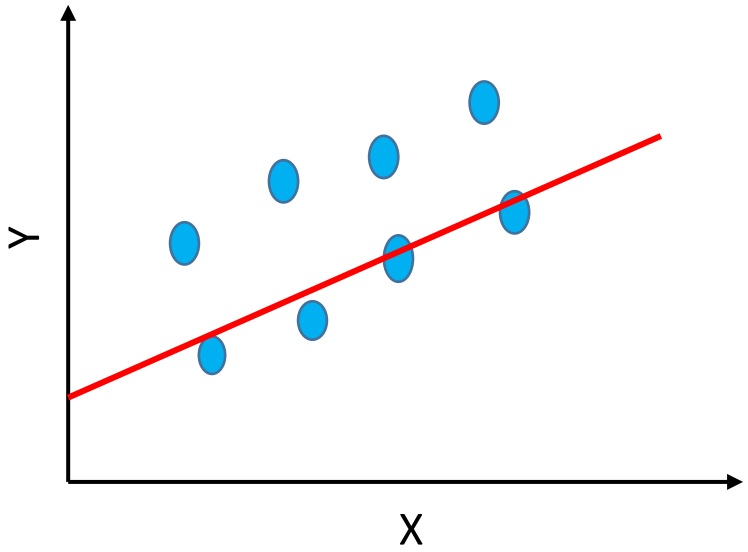
Siddhardhan

Gradient Descent for Linear Regression



Model Optimization

Optimization refers to determining best parameters for a model, such that the loss function of the model decreases, as a result of which the model can predict more accurately.

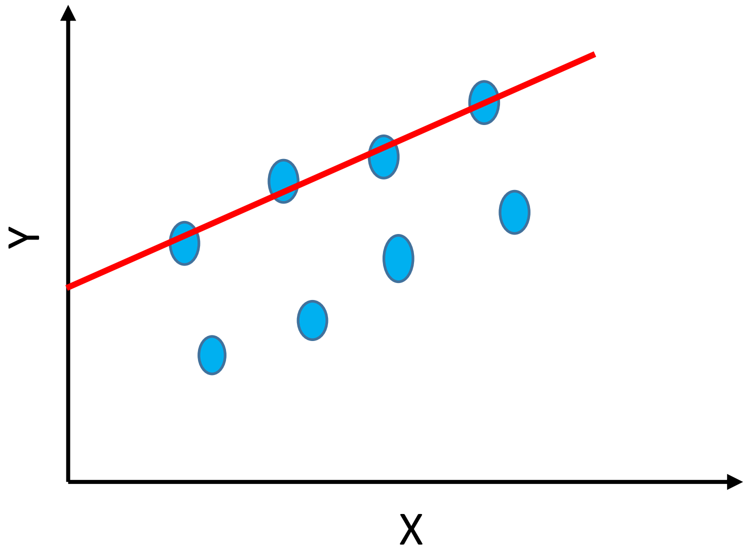


$Y = m_1X + C_1$

(m_1 & C_1 are the parameters of the line)

Model Optimization

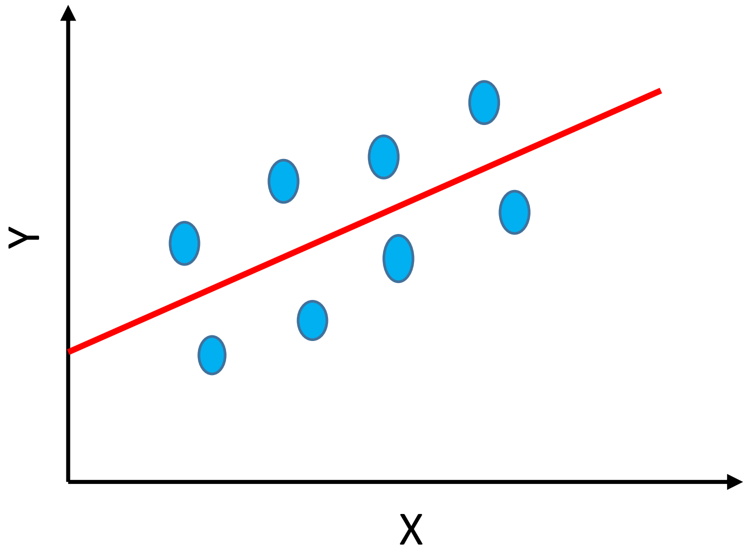
Optimization refers to determining best parameters for a model, such that the loss function of the model decreases, as a result of which the model can predict more accurately.



$$Y = m_2X + C_2$$

Model Optimization

Optimization refers to determining best parameters for a model, such that the loss function of the model decreases, as a result of which the model can predict more accurately.



$Y = m_3X + C_3$

Hence, m_3 & C_3 are the best parameters

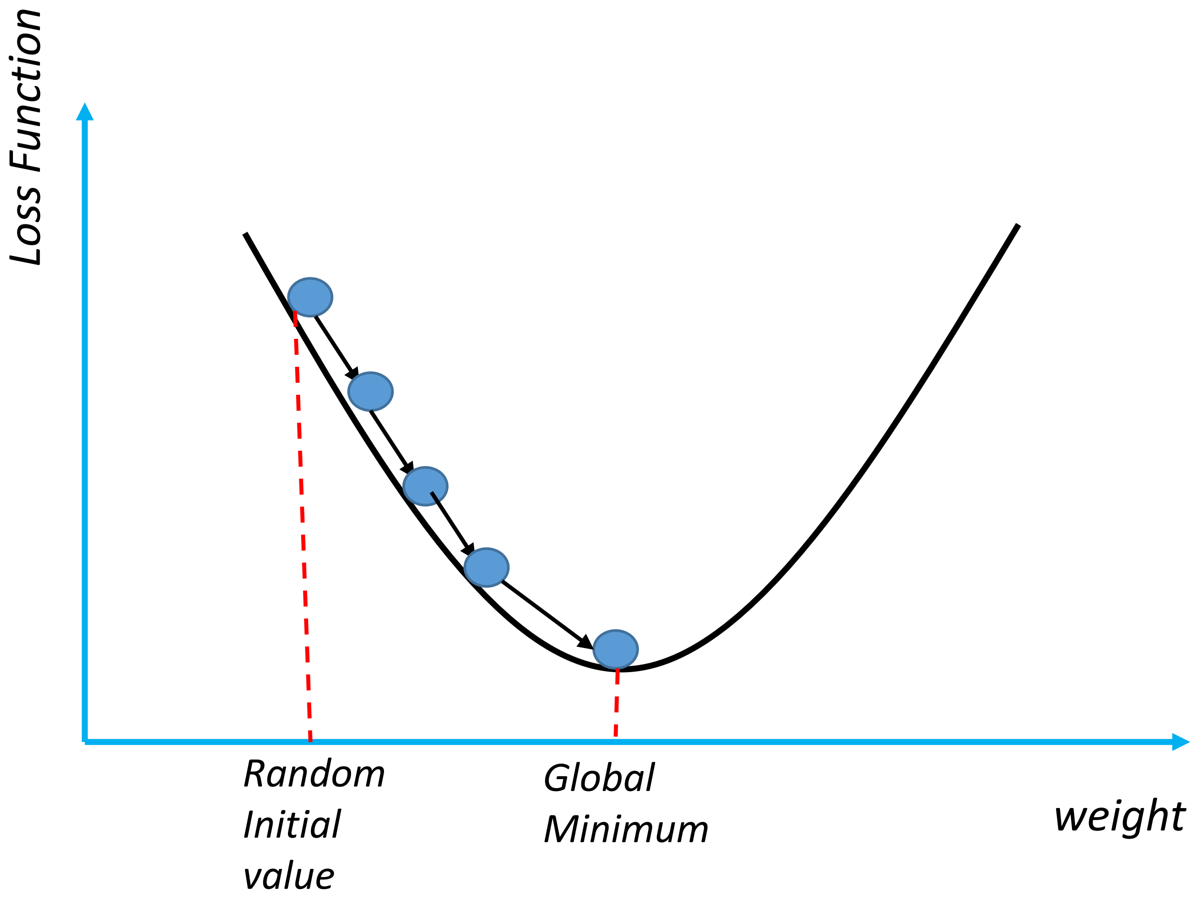
Loss Function

Loss function measures how far an estimated value is from its true value.

It is helpful to determine which model performs better & which parameters are better.

$$\textit{\textbf{Loss}} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

Gradient Descent



Gradient Descent

Gradient Descent is an optimization algorithm used for minimizing the loss function in various machine learning algorithms. It is used for updating the parameters of the learning model.

$$m = m - LD_m$$

$$c = c - LDc$$

m --> slope
c --> intercept
L --> Learning Rate

D_m --> Partial Derivative of loss function with respect to m

D_c --> Partial Derivative of loss function with respect to c

Gradient Descent

$$D_m = \frac{\partial(Cost\ Function)}{\partial m} = \frac{\partial}{\partial m}\left(\frac{1}{n}\sum_{i=0}^{n}\left(y_i - y_{i\,pred}\right)^2\right)$$

$$= \frac{1}{n}\frac{\partial}{\partial m}\left(\sum_{i=0}^{n}\left(y_i - (mx_i + c)\right)^2\right)$$

$$= \frac{1}{n}\frac{\partial}{\partial m}\left(\sum_{i=0}^{n}\left(y_i^2 + m^2x_i^2 + c^2 + 2mx_ic - 2y_imx_i - 2y_ic\right)\right)$$

$$= \frac{-2}{n}\sum_{i=0}^{n}x_i\left(y_i - (mx_i + c)\right)$$

$$= \frac{-2}{n}\sum_{i=0}^{n}x_i\left(y_i - y_{i\,pred}\right)$$

$$D_c = \frac{\partial(Cost\ Function)}{\partial c} = \frac{\partial}{\partial c}\left(\frac{1}{n}\sum_{i=0}^{n}\left(y_i - y_{i\,pred}\right)^2\right)$$

$$= \frac{1}{n}\frac{\partial}{\partial c}\left(\sum_{i=0}^{n}\left(y_i - (mx_i + c)\right)^2\right)$$

$$= \frac{1}{n}\frac{\partial}{\partial c}\left(\sum_{i=0}^{n}\left(y_i^2 + m^2x_i^2 + c^2 + 2mx_ic - 2y_imx_i - 2y_ic\right)\right)$$

$$= \frac{-2}{n}\sum_{i=0}^{n}\left(y_i - (mx_i + c)\right)$$

$$\frac{-2}{n}\sum_{i=0}^{n}\left(y_i - y_{i\,pred}\right)$$