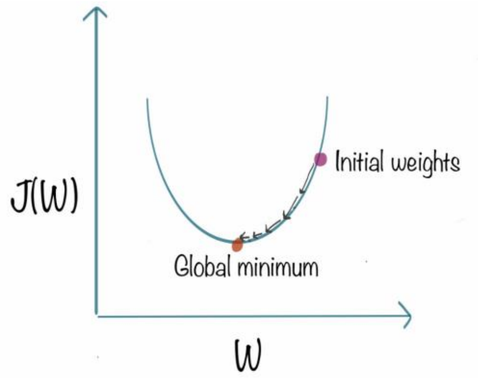
Siddhardhan

Gradient Descent for Logistic Regression



Logistic Regression

About Logistic Regression:

1. Supervised Learning Model
2. Classification model
3. Best for Binary Classification Problem
4. Uses Sigmoid function
5. Binary Cross Entropy Loss Function (or) Log Loss

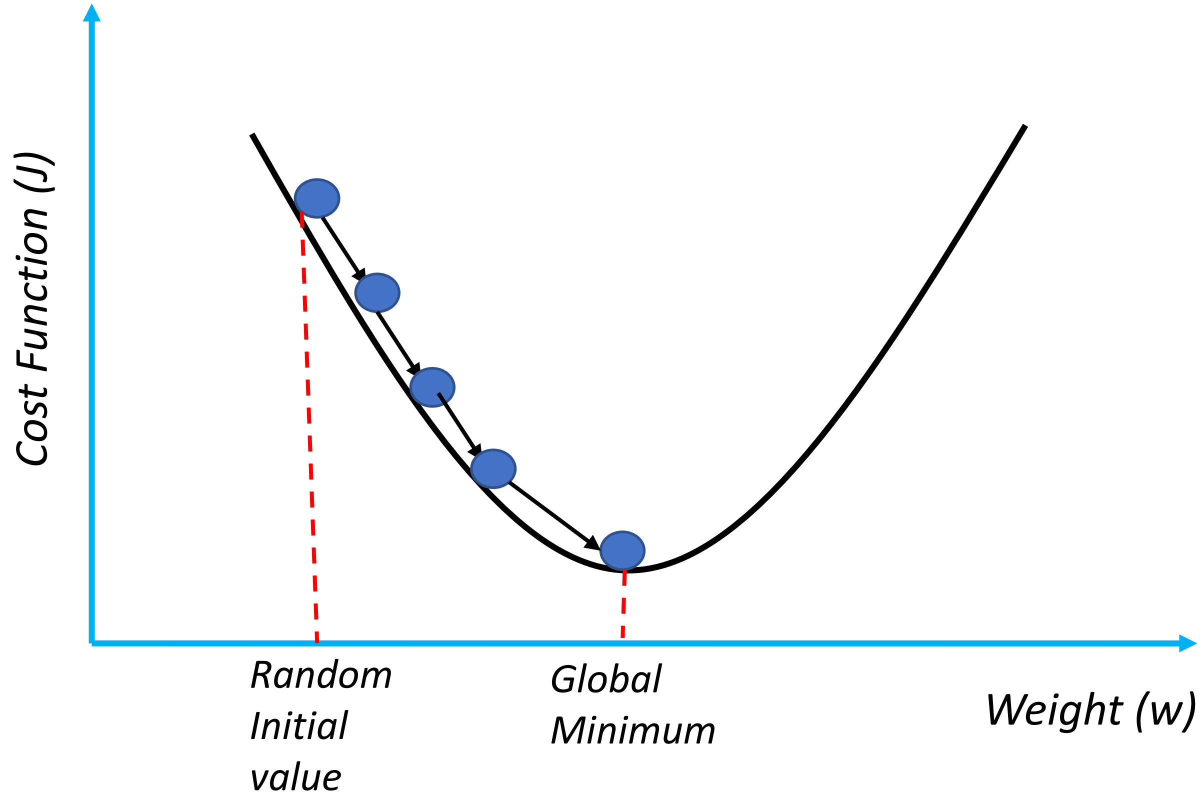
$$\hat{Y} = \frac{1}{1 + e^{-Z}}$$

Sigmoid Function

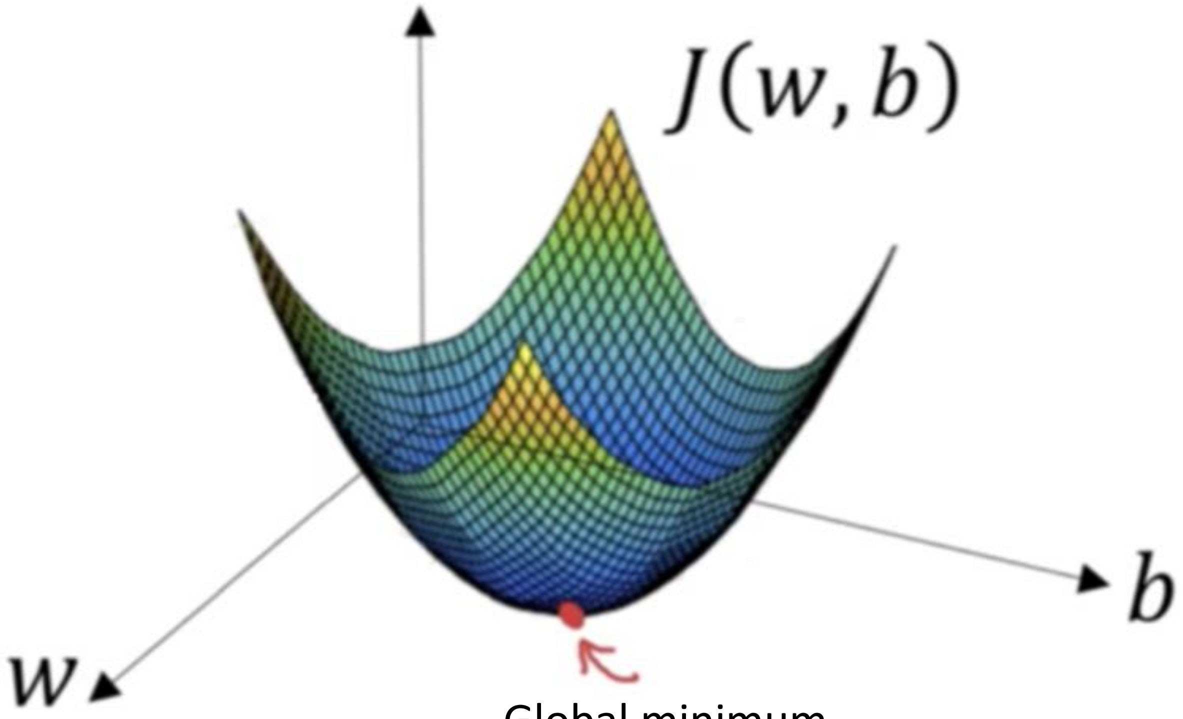
$$Z = w.X + b$$

$$J(w, b) = \frac{1}{m} \Sigma(L(y^{(i)}, \hat{y}^{(i)})) = - \frac{1}{m} \Sigma(y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log (1 - \hat{y}^{(i)}))$$

Gradient Descent



Gradient Descent in 3 Dimension



Global minimum

Gradient Descent

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

$$w_2 = w_1 - L*dw$$

$$b_2 = b_1 - L*db$$

w --> weight
b --> bias
L --> Learning Rate

dw --> Partial Derivative of cost function with respect to w

db --> Partial Derivative of cost function with respect to b

$$dw = \frac{1}{m} * (\hat{Y} - Y).X$$

$$db = \frac{1}{m} * (\hat{Y} - Y)$$

Logistic Regression

Logistic Regression model:

- Sigmoid Function

$$\hat{Y} = \frac{1}{1 + e^{-Z}} \qquad Z = w.X + b$$

- Updating weights through Gradient Descent

$$w_2 = w_1 - L*dw$$

$$b_2 = b_1 - L*db$$

- Derivatives

$$dw = \frac{1}{m} * (\hat{Y} - Y).X$$

$$db = \frac{1}{m} * (\hat{Y} - Y)$$