Linear Regression

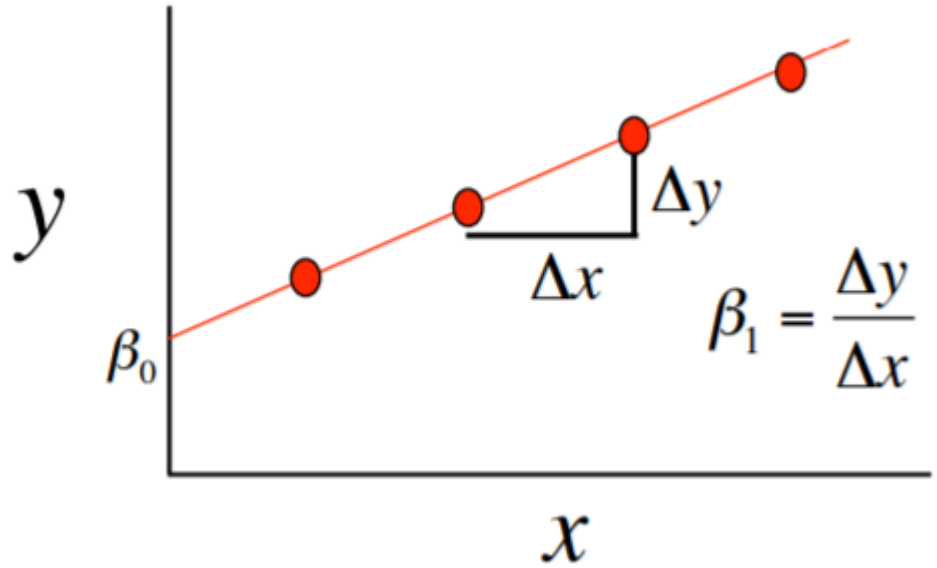
Linear Regression

- Regression is the analysis of the relation between an independent and one or more dependent variable assuming a linear relation.
- In Linear Regression the idea is to fit a line that is the best fit to data.
- The technique to fit the best line is called **Least Square Method**.
- Least square method is a process to fit a curve to set of data by reducing the distance between actual value and predicted value.
- The distance b/w actual value & predicted value is called **Error**.

Linear Regression

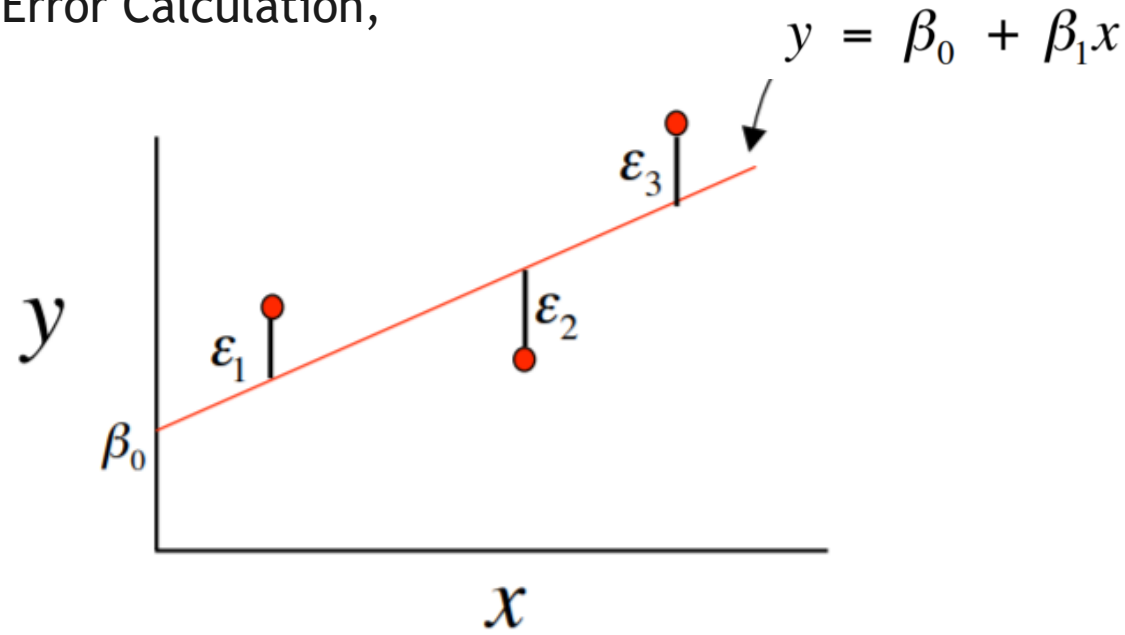
- Mathematically,

$$y = \beta_0 + \beta_1 x$$



Linear Regression

- Error Calculation,



Loss Function

- The loss is the error in the predicted value of **m** and **c** in the eq. **y=mx+c**
- The loss is calculated using Mean Squared Error (MSE)

$$E = \frac{1}{n}\sum_{i=0}^{n}(y_i - \bar{y}_i)^2$$

n	Total number of Samples
y_i	**i**th instance of true value
$\bar{y}_i$	**i**th instance of predicted value

We have defined the loss function, In order to get the best fit line we need to minimize it.

How to reduce the cost function?

- To minimize the loss function we need to get the optimum value of **m** and **c** such that the error is reduced.
- Initially let us start **m** and **c** at 0.
- we update their value iteratively such that the loss function is minimized.

Once the cost function is minimized the m and c values will be optimum and the prediction made using those m and c value will give the best fit line.

The Gradient Descent

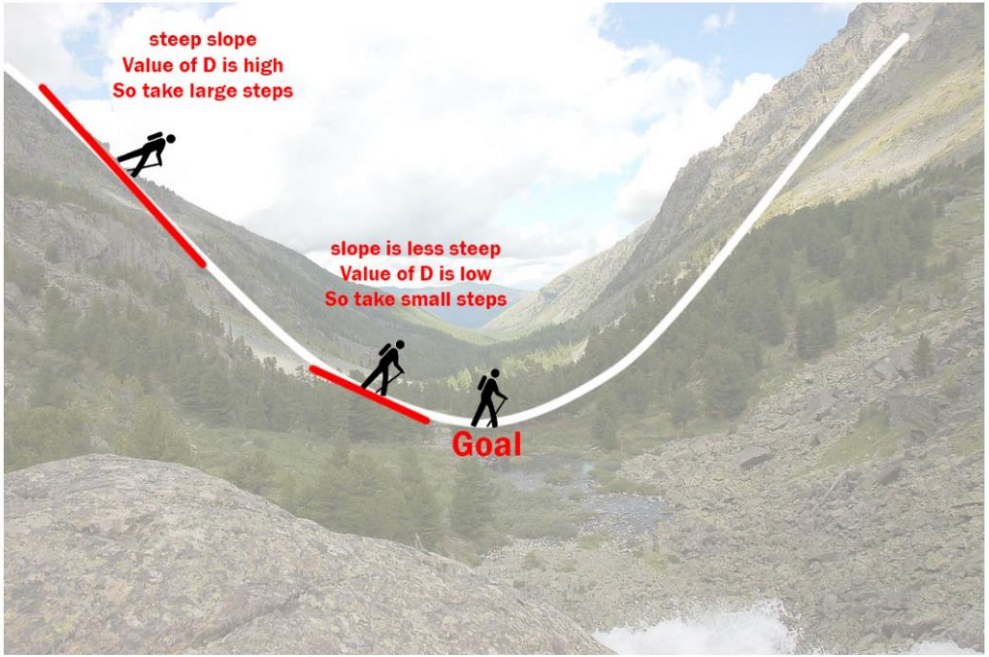


Illustration of how the gradient descent algorithm works

The gradient descent is an optimization algorithm which is used to train the models.

The Gradient Descent

Let L = Learning rate, determines how much m changes

Now we will calculate the derivative of Loss function wrt m and c.

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

$$D_c = \frac{-2}{n}\sum_{i=0}^{n}(y_i - \bar{y}_i)$$

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

$$D_m = \frac{-2}{n}\sum_{i=0}^{n} x_i(y_i - \bar{y}_i)$$

$$m = m - L \times D_m$$

$$c = c - L \times D_c$$

The Gradient Descent

Let L = Learning rate, determines how much m changes

Now we will calculate the derivative of Loss function wrt m and c.

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

$$D_c = \frac{-2}{n}\sum_{i=0}^{n}(y_i - \bar{y}_i)$$

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

$$D_m = \frac{-2}{n}\sum_{i=0}^{n} x_i(y_i - \bar{y}_i)$$

$$m = m - L \times D_m$$

$$c = c - L \times D_c$$