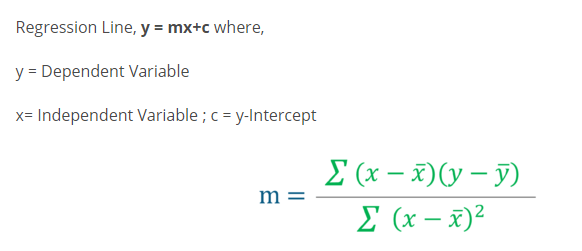
**Linear Regression**

**Impo:** <https://towardsdatascience.com/linear-regression-using-python-b136c91bf0a2>

The objective of a linear regression model is to find a relationship between one or more features(independent variables) and a continuous target variable(dependent variable). When there is only feature it is called Uni-variate Linear Regression and if there are multiple features, it is called Multiple Linear Regression.

## **Least Square Method – Finding the best fit line**

Least squares is a statistical method used to determine the best fit line or the regression line by minimizing the sum of squares created by a mathematical function. The “square” here refers to squaring the distance between a data point and the regression line. The line with the minimum value of the sum of the square is the best-fit regression line.



**Find the value of m and c using above formula:**

**# Mean X and Y**

**mean\_x = np.mean(X)**

**mean\_y = np.mean(Y)**

**# Total number of values**

**n = len(X)**

**# Using the formula to calculate m and c**

**numer = 0**

**denom = 0**

**for i in range(n):**

**numer += (X[i] - mean\_x) \* (Y[i] - mean\_y)**

**denom += (X[i] - mean\_x) \*\* 2**

**m = numer / denom**

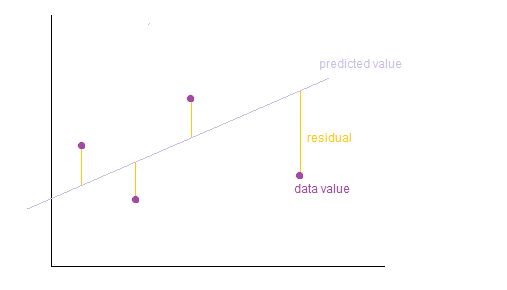
**c = mean\_y - (m \* mean\_x)**

**# Print coefficients**

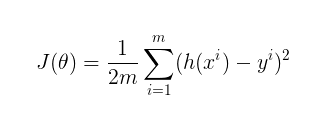
**print(m, c)**

* **How do we determine the best fit line?**

The line for which the error between the predicted values and the observed values is minimum is called the best fit line or the regression line. These errors are also called as residuals. The residuals can be visualized by the vertical lines from the observed data value to the regression line.



* **Cost Function :**

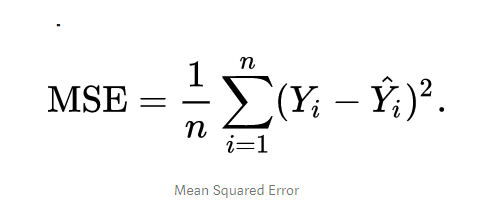
Our objective is to find the model parameters so that the cost function is minimum. We will use Gradient Descent to find this.

* **Cost (Loss) functions :**

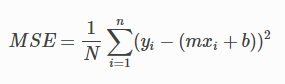
In the case of gradient descent, the objective is to find a line of best fit for some given inputs, or X values, and any number of Y values, or outputs. A cost function is defined as:

…a function that maps an event or values of one or more variables onto a real number intuitively representing some “cost” associated with the event.

1. [**mean squared error**](https://en.wikipedia.org/wiki/Mean_squared_error) **:**

****

**OR**

****

**N**

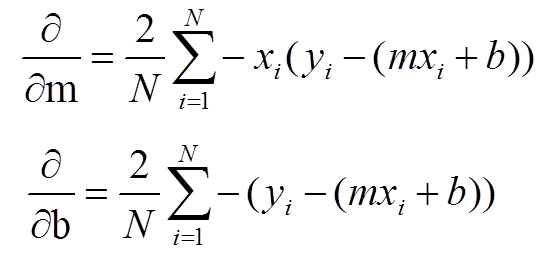
N is the total number of observations (data points)

## **Least Square Method:**

* **Gradient descent**

Gradient descent is a generic optimization algorithm used in many machine learning algorithms. It iteratively tweaks the parameters of the model in order to minimize the cost function.

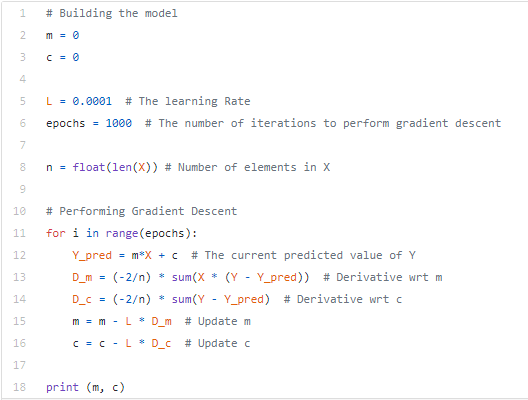
It is an optimization algorithm to find the minimum of a function. We start with a random point on the function and move in the **negative direction** of the **gradient of the function** to reach the **local/global minima**.



What we’re doing here is applying partial derivatives with respect to both ***m*** and **b** to the cost function to point us to the lowest point.

To find the values of **m** and **b**.

Note: **Gradient descent is one of the simplest and widely used algorithms in machine learning, mainly because it can be applied to any function to optimize it.**



[**https://towardsdatascience.com/linear-regression-using-gradient-descent-97a6c8700931**](https://towardsdatascience.com/linear-regression-using-gradient-descent-97a6c8700931)