Cost Function:

Where,

are the predicted points

are the true output points

By subtracting the two, we get the error value

We are squaring it because we are using the cost function called Mean Squared Error

There are different cost functions as well such as mean absolute error, root mean squared error.

is the intercept

is the slope

By changing them we can minimize our cost function

Final aim: Minimize cost function

Equation of a straight line

For understanding purposes, making = 0, for line passing through origin

So

Dataset:

|  |  |
| --- | --- |
| x | y |
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |

First we take as 1, second time as 0.5 and third time as 0

We also plot a graph between J() and . The entire curve obtained is called Gradient descent.

At the same time, putting random values for is not optimum, hence converge algorithm is applied.

