**GRADIENT DESCENT**

Gradient Descent is the Optimization Algorithm,Optimization is the big part of machine learning.

Almost every machine learning algorithm have optimization algorithm

Graident Descent is the Optimization Problem used to find out the parameters of function which minimize the cost function

Gradient descent is best when parameters cant be calculate analytical and must be searched by optimization algo.

For example: liner regression we look for best fit line and lets see how best fit line is choosen , there is method called sse in which we determine sum of (actual - predict) sq and whicch line get least error is choosen .hence we need a idea by which we can get best fit line or optimize our problme using less iteration there we use gradient desscent

Equation of line : y=mx+c

For gradient descent u should know about partial derivatives and chin rule

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**.

### **Example by hand :**

**Question** : Find the local minima of the function y=(x+5)² starting from the point x=3

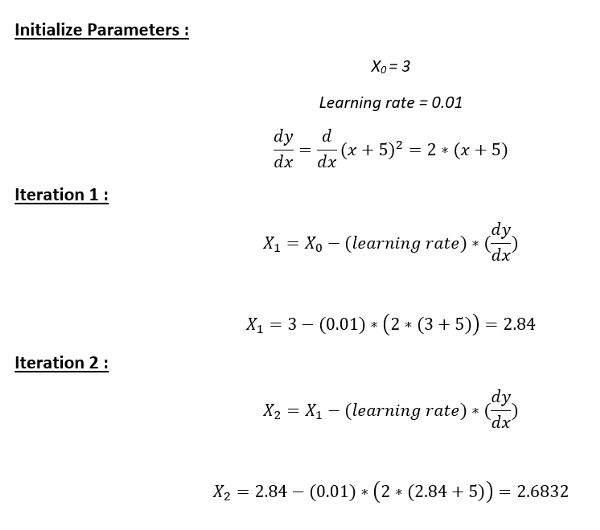
**Solution :** We know the answer just by looking at the graph. y = (x+5)² reaches it’s minimum value when x = -5 (i.e when x=-5, y=0). Hence x=-5 is the local and global minima of the function.

Now, let’s see how to obtain the same numerically using gradient descent.

**Step 1** : Initialize x =3. Then, find the gradient of the function, dy/dx = 2\*(x+5).

**Step 2** : Move in the direction of the negative of the gradient ([Why?](https://www.khanacademy.org/math/multivariable-calculus/multivariable-derivatives/gradient-and-directional-derivatives/v/why-the-gradient-is-the-direction-of-steepest-ascent)). But wait, how much to move? For that, we require a learning rate. Let us assume the **learning rate → 0.01**

**Step 3** : Let’s perform 2 iterations of gradient descent



**Step 4** : We can observe that the X value is slowly decreasing and should converge to -5 (the local minima). However, how many iterations should we perform?

Let us set a precision variable in our algorithm which calculates the difference between two consecutive “x” values . If the difference between x values from 2 consecutive iterations is lesser than the precision we set, stop the algorithm !