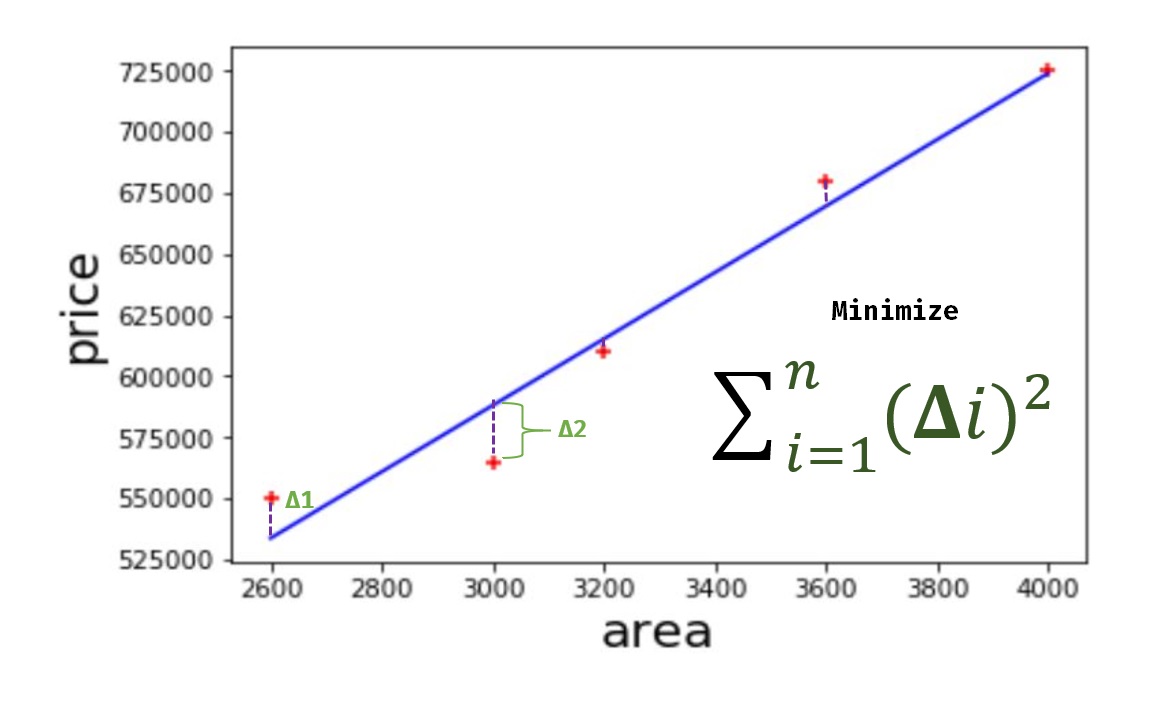
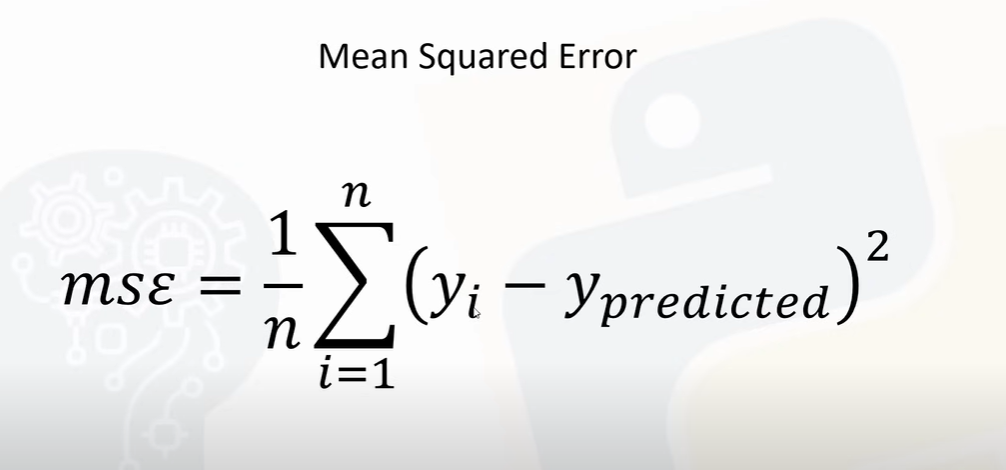


We can draw multiple line for data points but how can we identify that which of this line is perfect fit. so, for that one way is:





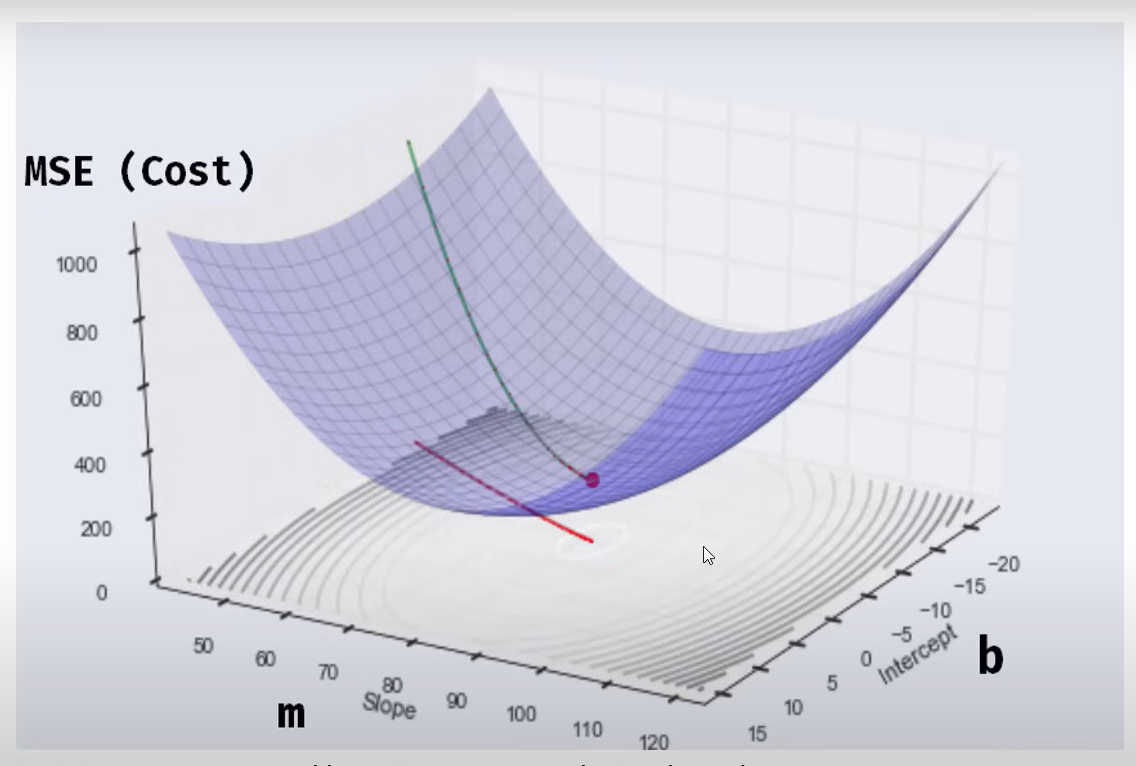


Ypredicated = mx + b

Line which has minimum mean squared error is a best fit line

So, another way is **gradient descent** is an algorithm that finds best fit line for given training data set.

Working of gradient descent algorithm:

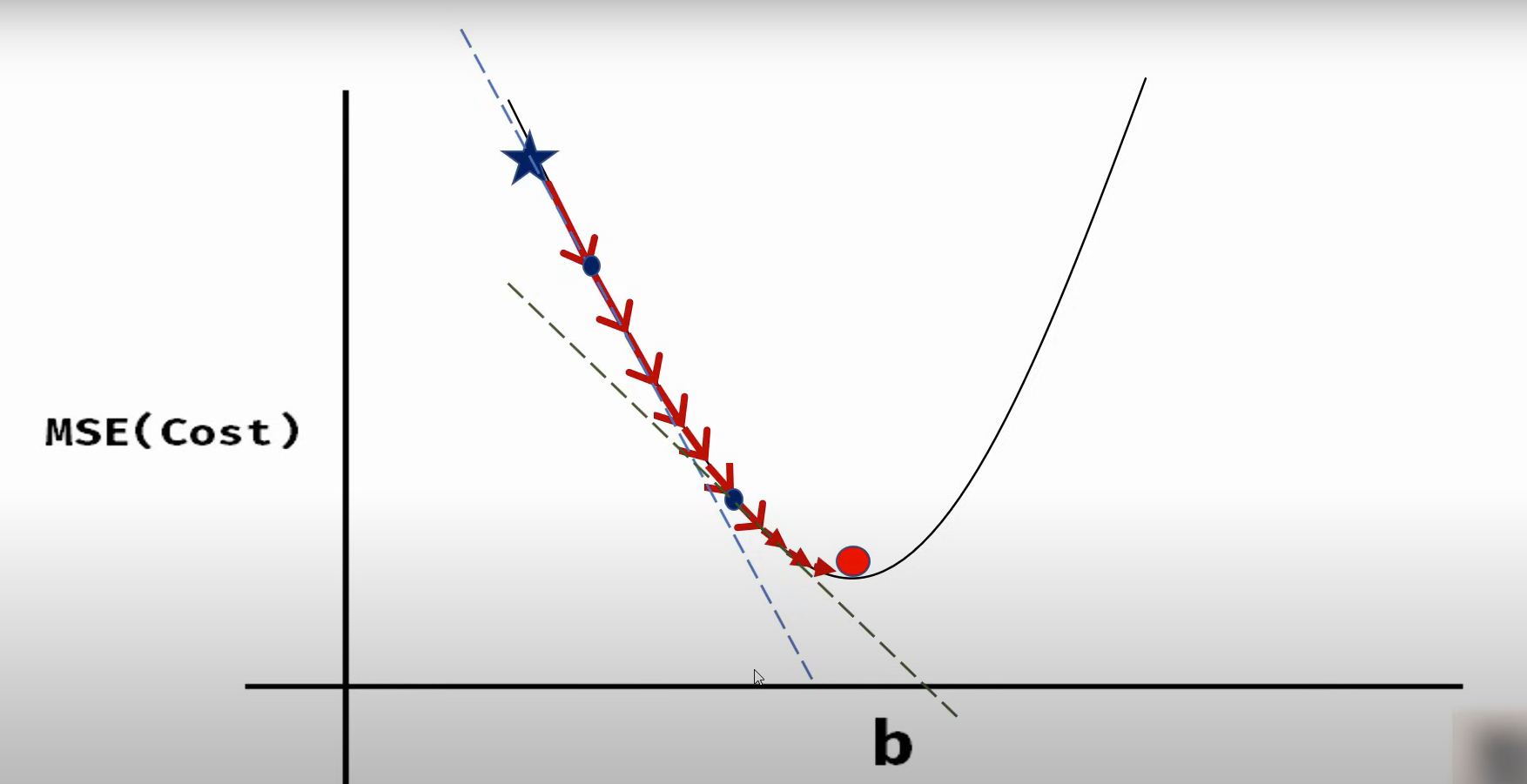


Above graph gives MSE for different value of m (slope/coefficient) and for b (intercept)

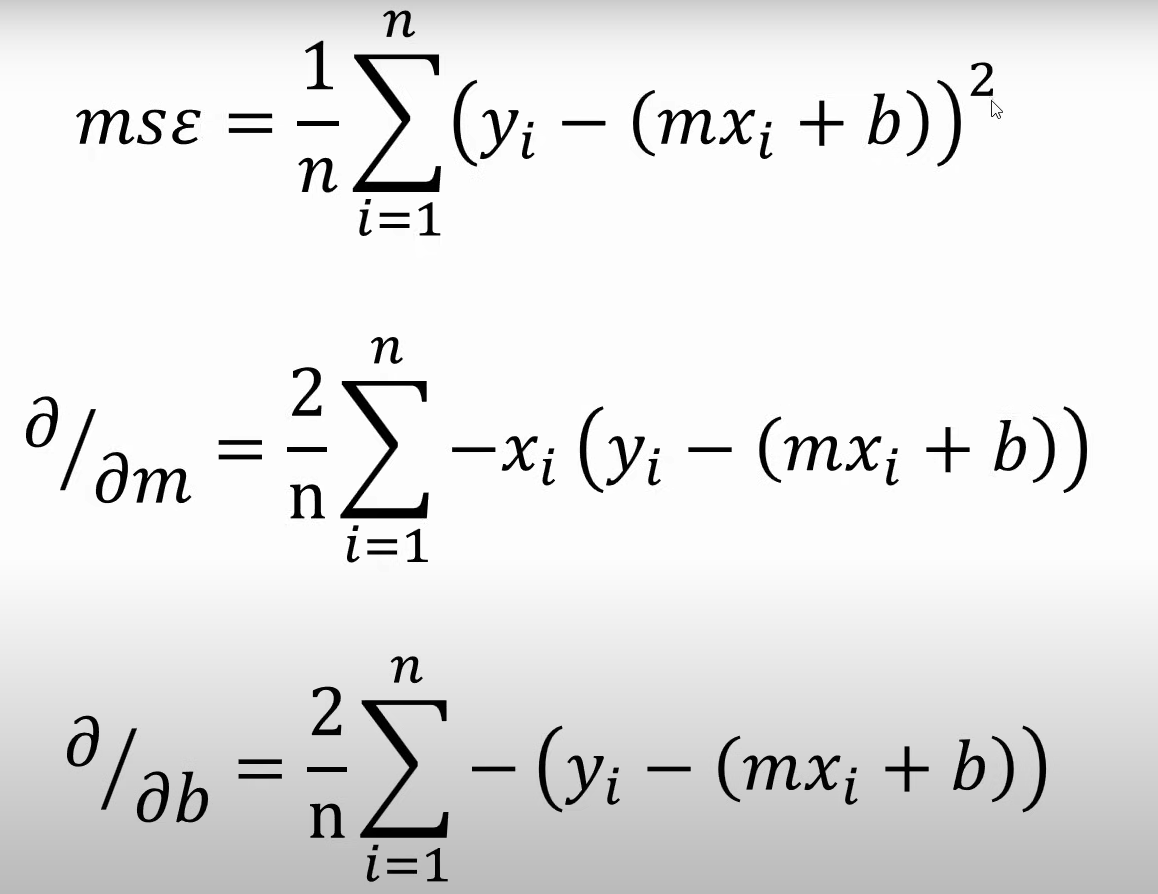
We have to find value of m and b for which cost will be minimum.

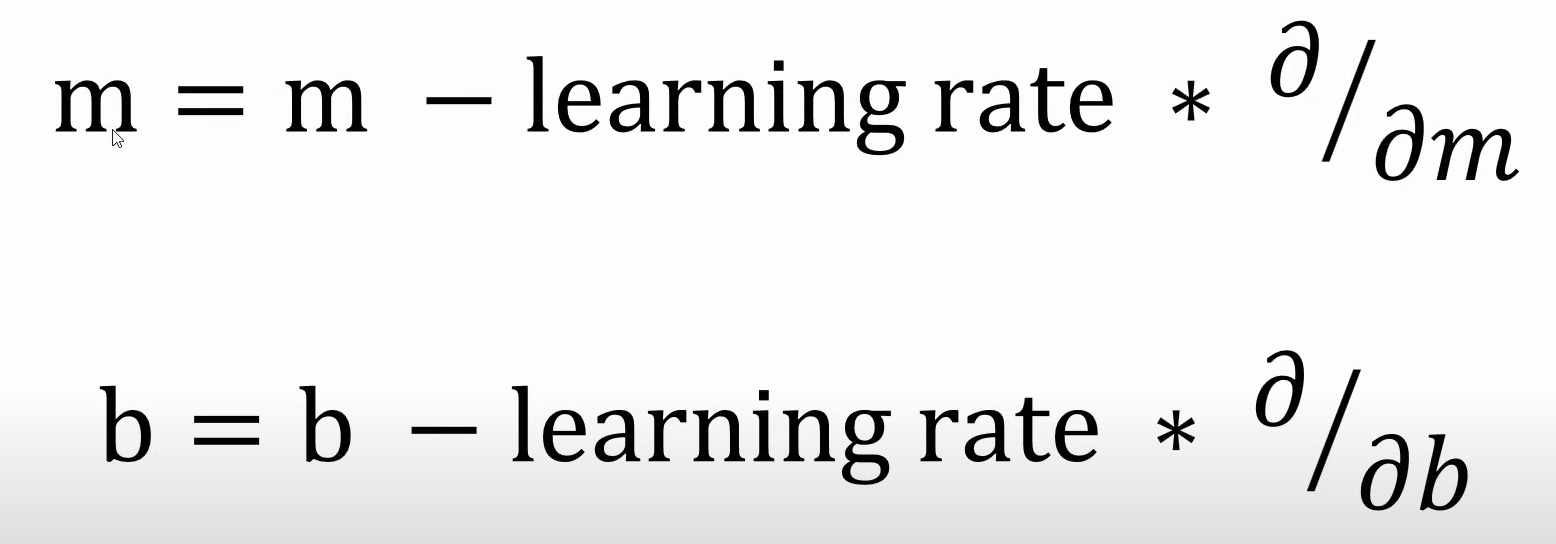
1) In above graph for 0 m and 0 b value of MSE is 1000 (highest point of green line). By taking mini steps (on green line, downward) we can see that value of MSE is decreased.

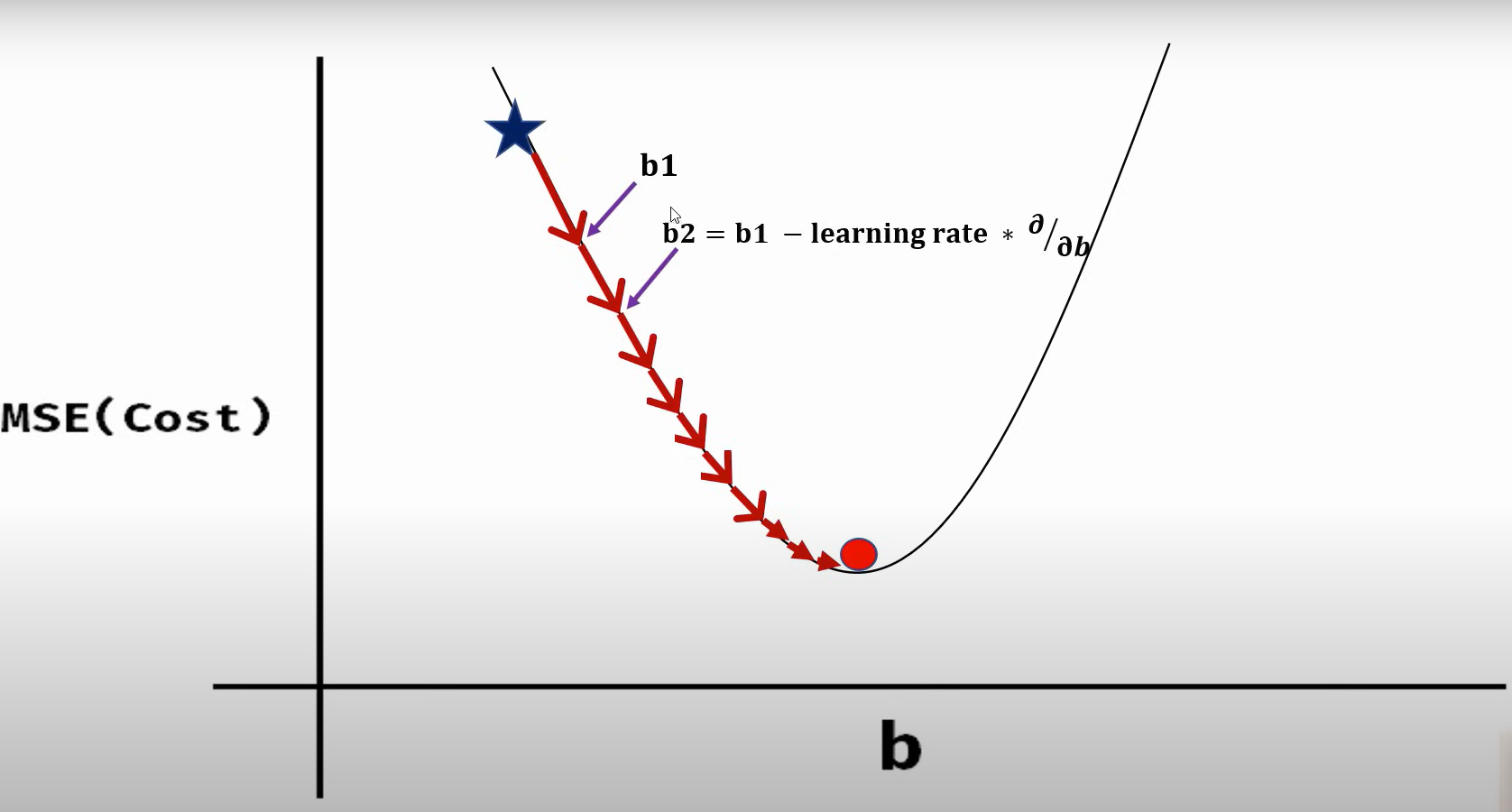
2) we keep on taking this step until we reach to minimum value.



For this we have to calculate slope at each point which give us direction to move forward and learning rate will help to us to reach on next point. For slope we have to find partial derivatives (derivation of MSE with respect to b, derivation of MSE with respect to a).





****