**LINEAR REGRESSION**

The representation is a linear equation that combines a specific set of input values (x) the solution to which is the predicted output for that set of input values (y). As such, both the input values (x) and the output value are numeric.

Example algo using linear regression:

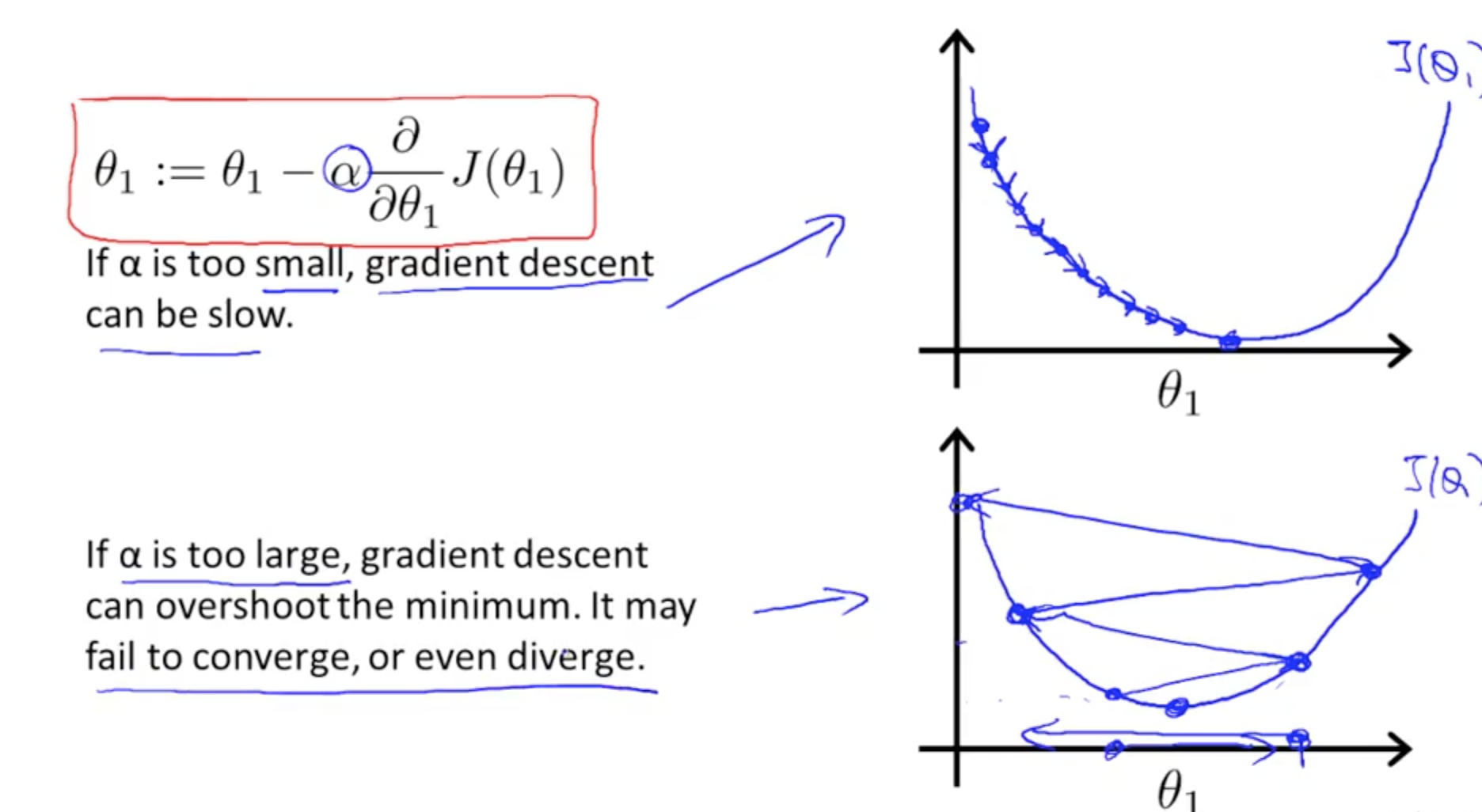
**Gradient Descent:**

Gradient descent is a first-order iterative optimization algorithm for finding a local minimum of a differentiable function. The idea is to take repeated steps in the opposite direction of the gradient of the function at the current point, because this is the direction of steepest descent.

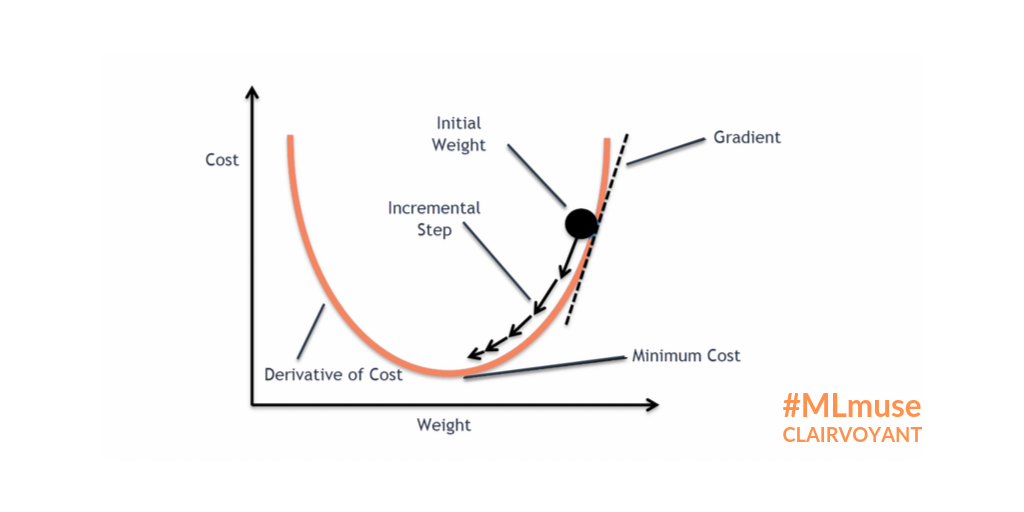
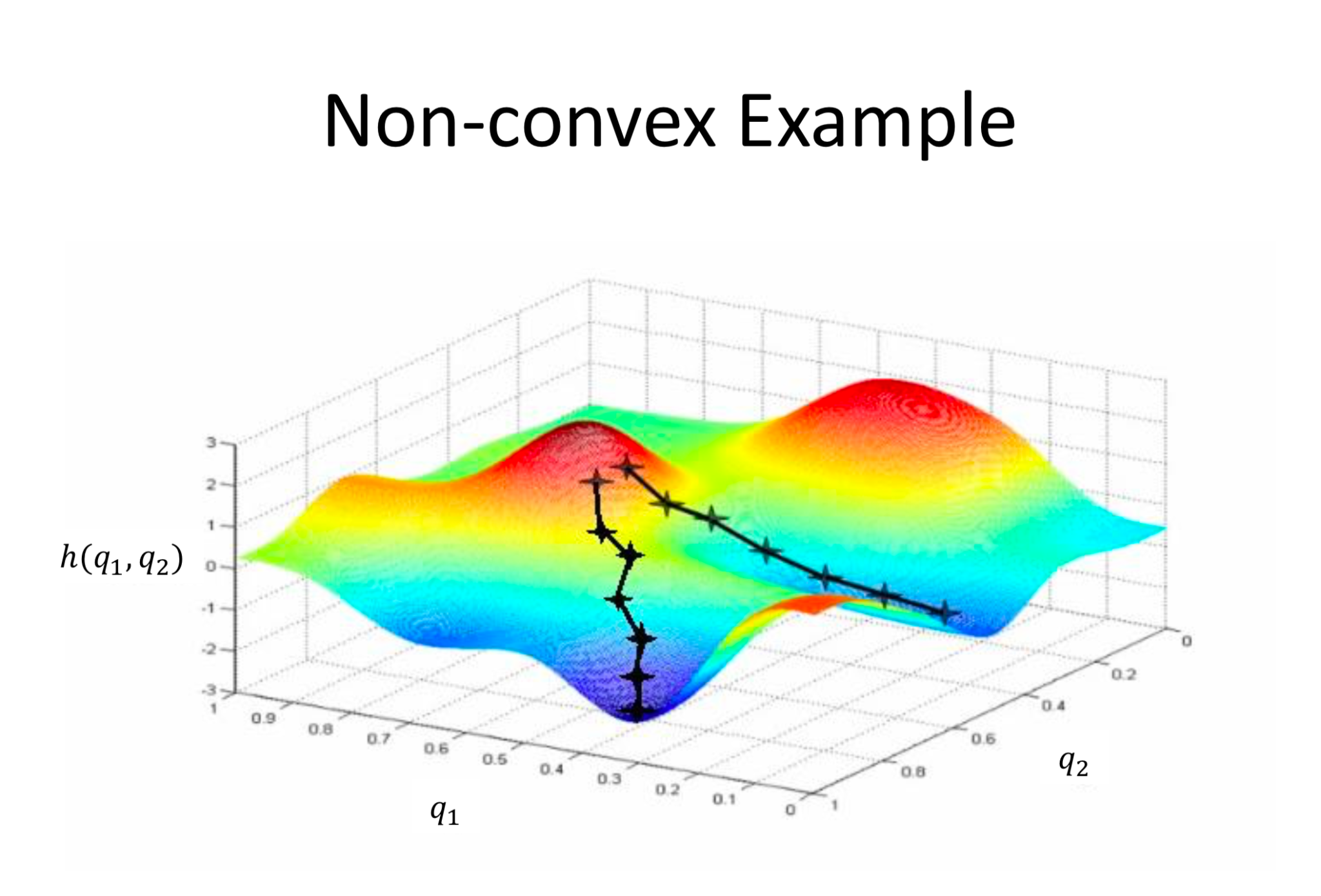
The gradient descent algorithm is repeat until convergence:

j=0,1 represents the feature index number.

represent the rate of descent, if it's too small GD can be slow but if it’s too large GD might fail to converge or even diverge

.

Types :

1. Convex - 
2. NonConvex - 

Depending on where one starts on the graph, one could end up at different points. The image above shows us two different starting points that end up in two different places.