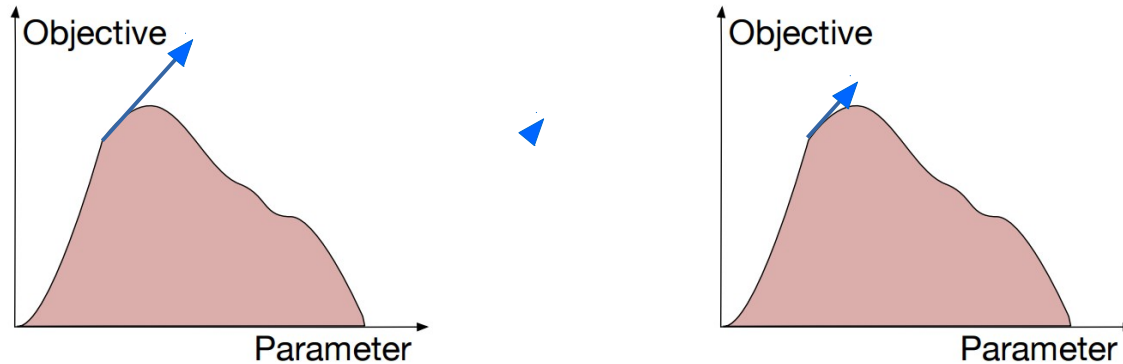


The learning rate is used to determine how large of a step is taken while converging on the extrema



The above images are altered from an image taken from the 10c Lecture slides (http://www.cs.colorado.edu/~jbg/teaching/CSCI_3022/10c.pdf). The left represents a learning rate of 1 while the right has a learning rate of 0.5

The number of passes over the data is equal to the number of examples, which for the full dataset is 164.

Some of the words that are the best predictors for the baseball class are:

ball, catcher, pitching, hit, baseball, run

Some of the words that are the best predictors for the hockey class are:

hockey, playoffs, golchowy, goals, pick, next

The method of finding these is by using the words with the lowest and highest coefficients respectively

Some of the words that are the worst predictors of both classes are:

dirt, handle, domination, rounding, ranked, tandem

The method of finding this is by using the words that have the coefficients that are closest to 0