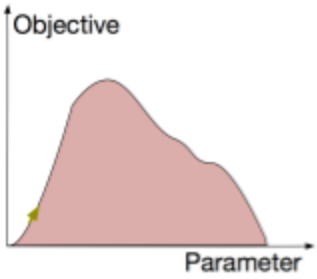
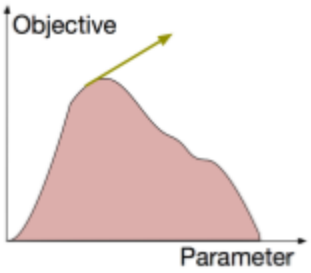
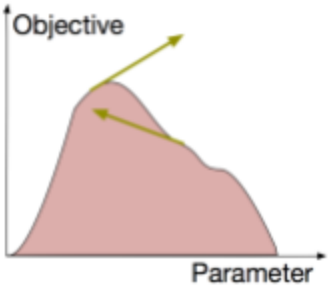
#Used the 4 late day advantage in this assignment since I took 1 late day in last assignment.

1. What is the role of the learning rate? [Give examples/graphs]

Learning rate helps you to approach your optimal beta value. The role of the learning rate is to help improve objective function toward the optimal weight. For example here in the picture figure 1 have a small learning rate so it will take few steps but eventually it will be its goal. Whereas figure 2 and figure 3 have larger amount of learning rate so it will over shoot and will never get to be in the goal or will take forever to be in. It means assumed step size.







2. How many passes over the data do you need to complete?

We only need one pass over the data. Our beta value will converge after one pass. Therefore, there is no use of updating this pass value since it does not update beta value.

3. What words are the best predictors of each class? How (mathematically) did you find them?

B*est for hockey - hockey, playoffs, golchowy, goals, pick, next, ice, biggest, names, playoff, goal, points, coverage, round, finals.*

Best for baseball are swing, anyone, usa, hitter, run, bat, stance, rickert, book, ball, catcher, pitching, hit, baseball, runs.

Mathematically, I argsorted the beta value array so it will give the beta value sorted by the indices from lower to higher. Then I looped over the first and last 15 index, first 15 will be the negative, which is hockey and last 15 will be mostly positive which is baseball.

4. What words are the poorest predictors of classes? How (mathematically) did you find them?

riel, broad, racist, vintage, blasted, intermissions, memoriam, deceased, rode, hesitate, pitiful, tone, everywhere, wrestling, silence.

Mathematically, I argsored the beta value but additionally I took the absolute value of the beta. The first 15 indices will be the poorest predictor because the beta value will be 0 or close to 0.