q6

October 11, 2019

0.1 Question 6: Is gradient descent with Armijo line-search faster than gradient descent with simple line-search in terms of running time? Is gradient descent with Armijo line-search faster than gradient descent with simple line-search in terms of number of required iterations? Explain any performance differences between the two approaches. Marks: 10

Question breakdown (Please note that all the experimental numbers are reported on my machine and might differ from any other machine, but the main conclusion should still hold):

Is gradient descent with Armijo line-search faster than gradient descent with simple line-search in terms of running time? Yes. GD with simple line search takes 15.6 seconds in total, where GD with Armijo line search with $\gamma = 0.13$ takes 14 seconds in total. Both reaching the same tolernace as specified by the parameter epsilon.

Is gradient descent with Armijo line-search faster than gradient descent with simple line-search in terms of number of required iterations? Yes. GD with simple line search takes 91 iterations, where GD with Armijo line search with $\gamma = 0.13$ takes 77 iterations in total. Both reaching the same tolernace as specified by the parameter epsilon.

Explain any performance differences between the two approaches. The main difference between the two methods is that the Armijo line search is more "agressive" in finding the α parameter than the simple line search. The check is done such that the objective function decreased sufficiently (by at least $\alpha\gamma||\nabla f(x)||_2^2$), not just decreased as in simple line search. This may lead to more line search iterations (and hence a longer time per gradient descent step), but it will also lead to less total number of gradient descent steps, since the convex function is forced to be decreased more every iteration.

[ ]: