MARK SCHMIDT, UBC

Global-Local Smoothness: Line Search can really help! Really!

Iteration complexities for first-order optimization algorithms are typically stated in terms of a global Lipschitz constant of the gradient, and optimal results can typically be achieved using fixed step sizes. But many objective functions that arise in practice have regions with small Lipschitz constants where larger step sizes can be used. Many local Lipschitz assumptions have thus been proposed, which lead to results showing that adaptive step sizes and/or line searches yield improved convergence rates over fixed step sizes. However, these faster rates tend to depend on the path taken by the algorithm, which makes it difficult to compare the iteration complexities of different methods.

We consider a simple characterization of global and local smoothness that only depends on properties of the function. This allows explicit lower and upper bounds on iteration complexities in terms of problem-dependent constants, which allows us to compare iteration complexities between algorithms. Under this assumption it is straightforward to show the advantages of line searches over fixed step sizes, of line searches with unbounded step sizes over those that bound the maximum step size, and even that in some settings gradient descent with line search has a better iteration complexity than accelerated gradient methods with fixed step sizes.