Second order methods with approximate derivatives with application to finite sum problems arising in machine learning

Stefania Bellavia[1]

[1] University of Florence

In the recent years interest kept on steadily increasing around second order methods with inexact Hessian and possibly inexact gradient information. This interest is motivated by problems where the derivative information about f is computationally expensive, such as large-scale optimization problems arising in machine learning and data analysis modeled as finite-sum optimization problems.

In this talk we will discuss some of these recent approaches considering both convex and non-convex problems. In particular we will focus on methods where the Hessian approximation is dynamically chosen with the aim to obtain an overall saving in terms of computational time. Typically, these methods employ loose Hessian approximation at the beginning of the iterative process and increasingly accurate approximations as a stationary point is approached. We will discuss both Inexact Newton methods for convex problems as well Adaptive Regularization with Cubic (ARC) approaches for handling the non-convex case. In this latter case, the strategy for dynamically choosing the Hessian accuracy is designed so that to maintain the optimal worst-case function complexity of the classical ARC methods. Variant allowing for inexact function and gradient have been proposed and will be described, too.

We will also analyse an application of these recent second order methods to large-scale finite-sum minimization where the gradient/Hessian approximation is computed by subsampling. Numerical results on finite-sum optimization problems arising in machine learning will be given to provide evidences of effectiveness.