Mini-Course

Methods for finite sum minimization

by Prof. Nataša Krejić
University of Novi Sad, Serbia

Abstract

Large finite sum minimization problems, where the objective function is the sample mean of a finite family of possibly nonconvex functions, arise in data analysis and machine learning problems, as well as in many other applications. The key difficulty is computational cost due to the expensive evaluation of the objective function and its derivatives. Thus, inexact approximations are used, subsampling being one of the popular approaches. In this mini-course (or sequence of seminars) we consider methods of first and second order based on subsampled directions. The Inexact Restoration framework, as a tool for subsample scheduling, is suitably coupled with an optimization method and yields a natural way of sample-size dynamical adjustment within the optimization method. The sampling strategy is embedded into either a line search or a trust region methodology. We discuss local and global properties for finding approximate first- and second- order optimal points and function evaluation complexity results. The convergence, in almost-sure sense, is proved for stochastic first order and second order directions. The proposed algorithms are validated on binary classification problems (both convex and nonconvex). An important property of this approach is that a burdensome tuning of the parameters involved is not required.

In the second part we focus on line search strategies. The existence of non-martingal errors prevents the direct application of Armijo-like step-size selection in the case of stochastic (subsampled) problems, even in the case of independent identically distributed (i.i.d.) samples. Thus, we propose a method with additional sampling at each iteration which resolves this difficulty with a small cost and combines efficiently the line search step sizes with suitably decaying step sizes. Stationarity of limit points is proved, in the almost-sure sense, while almost-sure convergence of the sequence of approximations to the solution holds with the additional hypothesis that the functions are strongly convex. Numerical experiments, including comparisons, with state-of-the art stochastic optimization methods, show the efficiency of the proposed approach.