

STOCHASTIC ASYNCHRONOUS PARALLEL METHODS IN OPTIMIZATION

Stephen Wright

University of Wisconsin-Madison, USA

swright@cs.wisc.edu

The ubiquity of multicore computer architectures and clusters and the advent of new applications of optimization in machine learning, data analysis and other areas has prompted a reevaluation of elementary optimization methods that had long been out of fashion. Stochastic synchronous parallel variants of such methods as stochastic gradient, coordinate descent, and successive projections have been a particular focus of interest. We survey such methods here, presenting computational results for several of them. We then focus on two such methods - coordinate descent for convex optimization and the Kaczmarz method for linear systems - and introduce a model of multicore computation that is both close to reality and amenable to analysis of convergence behavior. We show in particular that there is a threshold number of cores below which near-linear speedup of the algorithm (as a function of the number of cores) can be expected.

Joint work with Ji Liu (University of Rochester).