Stillfjord, Tony; Williamson, Måns
SRKCD: a stabilized Runge-Kutta method for stochastic optimization. (English)

Summary: We introduce a family of stochastic optimization methods based on the Runge-Kutta-Chebyshev (RKC) schemes. The RKC methods are explicit methods originally designed for solving stiff ordinary differential equations by ensuring that their stability regions are of maximal size. In the optimization context, this allows for larger step sizes (learning rates) and better robustness compared to e.g. the popular stochastic gradient descent method. Our main contribution is a convergence proof for essentially all stochastic Runge-Kutta optimization methods. This shows convergence in expectation with an optimal sublinear rate under standard assumptions of strong convexity and Lipschitz-continuous gradients. For non-convex objectives, we get convergence to zero in expectation of the gradients. The proof requires certain natural conditions on the Runge-Kutta coefficients, and we further demonstrate that the RKC schemes satisfy these. Finally, we illustrate the improved stability properties of the methods in practice by performing numerical experiments on both a small-scale test example and on a problem arising from an image classification application in machine learning.

MSC:
90C15 Stochastic programming
65K05 Numerical mathematical programming methods
65L06 Multistep, Runge-Kutta and extrapolation methods for ordinary differential equations
65L20 Stability and convergence of numerical methods for ordinary differential equations

Keywords: stochastic optimization; convergence analysis; Runge-Kutta-Chebyshev; stability

Full Text: DOI arXiv

References: