2024 Tenth Indian Control Conference (ICC)
9 - 11 December, 2024
IISER Bhopal, India

Convergence Rates for Stochastic Approximation: Biased Noise with Unbounded Variance, and Applications

Rajeeva L. Karandikar and M. Vidyasagar

Abstract— **In this paper, we study the convergence properties of the Stochastic Gradient Descent (SGD) method for finding a stationary point of a given objective function $J(\cdot)$. The objective function is not required to be convex. Rather, our results apply to a class of "invex" functions, which have the property that every stationary point is also a global minimizer. First, it is assumed that $J(\cdot)$ satisfies a property that is slightly weaker than the Kurdyka-Lojasiewicz (KL) condition, denoted here as (KL'). It is shown that the iterations $J(\boldsymbol{\theta}_t)$ converge almost surely to the global minimum of $J(\cdot)$. Next, the hypothesis on $J(\cdot)$ is strengthened from (KL') to the Polyak-Lojasiewicz (PL) condition. With this stronger hypothesis, we derive estimates on the rate of convergence of $J(\boldsymbol{\theta}_t)$ to its limit. Using these results, we show that for functions satisfying the PL property, the convergence rate of both the objective function and the norm of the gradient with SGD is the same as the best-possible rate for convex functions. While some results along these lines have been published in the past, our contributions contain two distinct improvements. First, the assumptions on the stochastic gradient are more general than elsewhere, and second, our convergence is almost sure, and not in expectation. We also study SGD when only function evaluations are permitted. In this setting, we determine the "optimal" increments or the size of the perturbations. Using the same set of ideas, we establish the global convergence of the Stochastic Approximation (SA) algorithm under more general assumptions on the measurement error, compared to the existing literature. We also derive bounds on the rate of convergence of the SA algorithm under appropriate assumptions.**

Summary:

This paper has been accepted for publication in *Journal of Optimization Theory and Applications*. The final accepted version can be found under:

https://arxiv.org/pdf/2312.02828.pdf

RLK is Professor Emeritus at the Chennai Mathematical Institute; Email: rlk@cmi.ac.in. MV is National Science Chair at Indian Institute of Technology Hyderabad; Email: m.vidyasagar@iith.ac.in. The research of MV was supported by the Science and Engineering Research Board, India.

979-8-3315-1721-2/24/$31.00 ©2024 IEEE