Translational AI Center (TrAC) Journal Club
Spring 2022

Xuping Tian

April 15 at 3PM (US Central Time)

https://iastate.zoom.us/j/93596996347?pwd=c2xlZEJoSlNpcWJ6V2lKZmNuRHZCUT09

Adaptive gradient methods with energy and momentum

Reading material:

1. https://arxiv.org/pdf/2010.05109.pdf
2. https://arxiv.org/pdf/2203.12199.pdf
3. https://arxiv.org/pdf/2203.12191.pdf

Abstract

We introduce AEGD, a first-order gradient-based algorithm for general optimization problems, based on a dynamically updated 'energy' variable, and its variants with both energy and momentum. Such energy-adaptive gradient algorithms are shown to be unconditionally energy stable, irrespective of the base step size. An energy-dependent convergence rate in the general nonconvex stochastic setting and a regret bound in the online convex setting are provided. We also study the dynamic behavior of the proposed algorithms through analysis of a high-resolution ODE system. Experimental results demonstrate that the energy-adaptive gradient algorithms show better generalization performance than SGD with momentum in training some deep neural networks.