Scale Invariant Stochastic Gradient Method with Momentum

FILIP NIKOLOVSKI*, IRENA STOJKOVSKA†

*Ss. Cyril and Methodius University, Faculty of Mechanical Engineering, Skopje
†Ss. Cyril and Methodius University, Faculty of Natural Sciences and Mathematics, Skopje

Optimization in noisy environments arises frequently in applications. Solving this problem quickly, efficiently, and accurately is therefore of great importance. The Stochastic Gradient Descent (SGD) method has proven to be a basic and an effective tool which is flexible enough to allow modifications which improve its convergence properties. In this paper we propose a new algorithm which combines the SGD with a modified momentum term using a two-point step size estimation in the Barzilai-Borwein framework. We analyze the convergence properties of the algorithm. Numerical experiments are performed on a standard set of test-functions in a noisy environment. We demonstrate that the proposed method is scale invariant and that it outperforms the "vanilla" SGD with momentum.