Stochastic Approximation Method with Second Order Search Directions

Nataša Krejić[1], Kataša Krklec Jerinkić[1], and Zoran Ovcin[2]

[1] Department of Mathematics and Informatics, Faculty of Science, University of Novi Sad, Trg Dositeja Obradovia 4, 21000 Novi Sad, Serbia
natasak@uns.ac.rs, natasa.krklec@dmi.uns.ac.rs
[2] Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia
zovcin@uns.ac.rs

Abstract

This paper presents a discussion on application of second-order- like search directions in the Stochastic Approximation methods together with convergence conditions and some results on numerical implementation. We consider strictly convex problems in noisy environment and assume that only noisy values for the objective function and the gradient are available, as well as some approximate Hessian value. Under the zero mean assumption on noise a convergence analysis is presented for methods that use some approximate second-order direction. We prove that there exists a level of inexactness, governed by the usual gain sequence in SA methods, that does not interfere with the convergence and hence derive the set of convergence conditions that are applicable to a number of search directions. These directions include the so called mini-batch subsampled Hessian in statistical learning and similar directions. A set of numerical tests is presented in order to demonstrate efficiency and implementation issues of the proposed methods.

Keywords: unconstrained optimization, stochastic approximation, subsampled Hessian, inexact methods, Newton-like methods

References

1. R. Bollapragada, R. Byrd, J. Nocedal, Exact and Inexact Subsampled Newton Methods for Optimization *arXiv:1609.08502 [math.OC]*
2. Dembo R, Eisenstat S.C. and Steihaug T., Inexact Newton Methods, *SIAM J. Numer. Anal., 19(2), 400408.*
3. R. H. Byrd, G. M. Chin, W. Neveitt, J. Nocedal, On the Use of Stochastic Hessian Information in Optimization Methods for Machine Learning, *SIAM J. Optim., 21 (3), (2011) pp. 977-995.*
4. R. H. Byrd, G. M. Chin, J. Nocedal, Y. Wu, Sample size selection in optimization methods for machine learning, *Mathematical Programming, 134(1), (2012) pp. 127-155.*

5. N. Krejić, Z. Lužanin, I. Stojkovska, Z. Ovcin Descent direction method with line search for unconstrained optimization in noisy environment, *Optimization Methods and Software* 30,6, (2015), 1164-1184.
6. R. H. Byrd, S. L. Hansen, J. Nocedal, Y. Singer, A Stochastic Quasi-Newton Method for Large-Scale Optimization, *Technical report, arXiv:1401.7020 [math.OC].*
7. M. P. Friedlander, M. Schmidt, Hybrid deterministic-stochastic methods for data fitting, *SIAM J. Scientific Computing 34 No. 3 (2012), pp. 1380-1405.*
8. J. C. Spall, Introduction to Stochastic Search and Optimization, *Wiley-Interscience serises in discrete mathematics, New Jersey, 2003.*