

Efficient Second-Order Optimization Methods for Machine Learning

P. Xu¹, F. Roosta², and M. Mahoney³

Abstract: Contrary to the scientific computing community which has, wholeheartedly, embraced the second-order optimization algorithms, the machine learning community has long nurtured a distaste for such methods, in favour of first-order alternatives. In this talk, we argue that such reluctance to employ curvature information can indeed hinder the training procedure in a variety of ways. Specifically, in the context of convex and non-convex machine learning problems, we demonstrate the theoretical properties as well as empirical performance of a variety of efficient Newton-type algorithms. In the process, we highlight the serious disadvantages of first-order methods and, in their light, showcase the practical advantages offered by such second-order methods.

¹ Institute for Computational and Mathematical Engineering
Stanford University *pengxu@stanford.edu*

² School of Mathematics and Physics
University of Queensland, Brisbane, Australia,
and
International Computer Science Institute, Berkeley, USA
fred.roosta@uq.edu.au

³ International Computer Science Institute
and Department of Statistics
University of California at Berkeley
mmahoney@stat.berkeley.edu