ABSTRACT

The stochastic gradient method plays a central role in large-scale statistical learning, where vast amounts of data and high dimensional models are employed. This method is, however, difficult to tune and several variance reduction methods have been recently proposed to address this problem. After motivating these methods, I will propose a new variant that enjoys improved learning properties. Numerical results on problems arising in text classification, speech and image recognition will be presented.
SPEAKER BIO

Jorge Nocedal is the David and Karen Sachs Professor of Industrial Engineering and Management Sciences at Northwestern University. He obtained a B.S. degree in physics from UNAM and a PhD in mathematical sciences from Rice University. His research interests are in optimization and its application in machine learning. He has served as the Editor in Chief of the SIAM Journal on Optimization, is a SIAM Fellow, and was awarded the 2012 George B. Danzig Prize.