Les mathématiques de l'apprentissage machine
« Atelier sur les défis modernes en théorie d'apprentissage »
23 au 26 avril 2018

Mathematics of Machine Learning
"Workshop on Modern Challenges of Learning Theory"
April 23-26, 2018

Representation, optimization and generalization properties of deep neural networks

Peter L. Bartlett *

bartlett@cs.berkeley.edu

We describe some recent results on deep neural networks in three directions. First, we investigate the impact of depth on representational properties of deep residual networks, which compute near-identity maps at each layer, showing how their representational power improves with depth and that the functional optimization landscape has the desirable property that stationary points are optimal. Second, we study implications for optimization in deep linear networks, showing how the success of a family of gradient descent algorithms that regularize towards the identity function depends on a positivity condition of the regression function. Third, we consider how the performance of deep networks on training data compares to their predictive accuracy, we demonstrate deviation bounds that scale with a certain "spectral complexity", and we compare the behavior of these bounds with the observed performance of these networks in practical problems.

This is joint with Steve Evans, Dylan Foster, Dave Helmbold, Phil Long and Matus Telgarsky.

*Department of Statistics, University of California, Berkeley, 367 Evans Hall, #3860, Berkeley, CA 94720-3860, USA