Stability and Optimization Error of SGD Algorithms for Pairwise Learning

Yiming Ying
Department of Mathematics and Statistics, State University of New York at Albany, USA
Email: yying@albany.edu

Pairwise learning refers to a learning task which involves a loss function depending on pairs of instances among which notable examples are bipartite ranking, metric learning, AUC maximization and minimum error entropy (MEE) principle. Stochastic gradient descent (SGD) algorithms have been widely used in pairwise learning. Understanding their statistical and computational properties is of fundamentally importance in theory and practice.

In this talk I present our work on the stability and its trade-off with optimization (computational) error of SGD in the pairwise learning setting. Firstly, we will show stability results for SGD in the convex, strongly convex and non-convex settings, from which generalization errors can be naturally derived. Secondly, we will present trade-off results between stability and optimization error of SGD for pairwise learning. This is achieved by lower-bounding the sum of stability and optimization error by the minimax statistical error over a prescribed class of pairwise loss functions. From this fundamental trade-off, we can derive lower bounds for optimization error (convergence rate) of SGD for pairwise learning.