Metric representations: Algorithms and Geometry

Anna Gilbert, University of Michigan

Given a set of distances amongst points, determining what metric representation is most "consistent" with the input distances or the metric that best captures the relevant geometric features of the data is a key step in many machine learning algorithms. In this talk, we focus on 3 specific metric constrained problems, a class of optimization problems with metric constraints: metric nearness (Brickell et al. (2008)), weighted correlation clustering on general graphs (Bansal et al. (2004)), and metric learning (Bellet et al. (2013); Davis et al. (2007)).

Because of the large number of constraints in these problems, however, these and other researchers have been forced to restrict either the kinds of metrics learned or the size of the problem that can be solved. We provide an algorithm, PROJECT AND FORGET, that uses Bregman projections with cutting planes, to solve metric constrained problems with many (possibly exponentially) inequality constraints. We also prove that our algorithm converges to the global optimal solution. Additionally, we show that the optimality error decays asymptotically at an exponential rate. We show that using our method we can solve large problem instances of three types of metric constrained problems, out-performing all state of the art methods with respect to CPU times and problem sizes.

Finally, we discuss the adaptation of PROJECT AND FORGET to specific types of metric constraints, namely tree and hyperbolic metrics.