Tree-Projected Gradient Descent for Estimating Gradient-Sparse Parameters on Graphs

Zhou Fan

Yale University

Abstract:

We study estimation of a gradient-sparse parameter vector theta in dimension p, having strong gradient-sparsity s on an underlying graph G. Given n observations and a smooth convex loss function L for which theta minimizes the population risk, we propose to estimate theta by a projected gradient descent algorithm that iteratively and approximately projects gradient steps onto spaces of vectors having small gradient-sparsity over low-degree spanning trees of G. We show that, under suitable restricted strong convexity and smoothness assumptions for the loss, the resulting estimator achieves the squared-error risk (s/n) log (1+p/s) up to a multiplicative constant that is independent of G. In contrast, previous polynomial-time algorithms have only been shown to achieve this guarantee under additional assumptions for G and/or the sparsity pattern of theta. As applications of our general framework, we apply our results to the examples of linear models and generalized linear models with random design.