Network Lasso: Clustering and Optimization in Large Graphs

D. Hallac, J. Leskovec, and S. Boyd

Proceedings SIGKDD, pages 387-396, 2015.

- network_lasso.pdf

- SnapVX software

Convex optimization is an essential tool for modern data analysis, as it provides a framework to formulate and solve many problems in machine learning and data mining. However, general convex optimization solvers do not scale well, and scalable solvers are often specialized to only work on a narrow class of problems. Therefore, there is a need for simple, scalable algorithms that can solve many common optimization problems. In this paper, we introduce the network lasso, a generalization of the group lasso to a network setting that allows for simultaneous clustering and optimization on graphs. We develop an algorithm based on the Alternating Direction Method of Multipliers (ADMM) to solve this problem in a distributed and scalable manner, which allows for guaranteed global convergence even on large graphs. We also examine a non-convex extension of this approach. We then demonstrate that many types of problems can be expressed in our framework. We focus on three in particular —- binary classification, predicting housing prices, and event detection in time series data —- comparing the network lasso to baseline approaches and showing that it is both a fast and accurate method of solving large optimization problems.

Page generated 2018-11-24 09:00:14 PST, by jemdoc.