Convex and Non-Smooth Optimization
Spring, 2018
Project Proposal
Peter Will

Overview

For my project I would like to examine the Linearized Alternating Direction Method with Adaptive Penalty (LADMAP) algorithm developed by Lin et al in 2011 to solve the Low Rank Recovery (LRR) problem. Briefly, LRR assumes that one has a matrix with elements that have been corrupted by a small amount of Gaussian noise. The objective is to recover a matrix which is both low-rank and of minimum distance from the original matrix. The authors make use of a convex relaxation of LRR that relies upon minimization of the nuclear norm as an approximation to rank minimization. The LADMAP method proposed is a variant of more general Alternating Direction Methods (ADM).

For the original work portion of the project I want to see how LADMAP performs on a modified version of the convex optimization problem formulated in Weinberger and Saul 2004 which is solved to find a lower-dimensional manifold embedding of higher dimensional data. Weinberger and Saul formulate a sdp convex optimization problem as follows:

$$\begin{aligned} \text{maximize} \quad & \mathrm{tr}(K) \\ \text{subject to} \quad & K \succeq 0 \\ & \Sigma_{ij} K_{ij} = 0 \\ & K_{ii} + K_{jj} - K_{ij} - K_{ji} = G_{ii} + G_{jj} - G_{ij} - G_{ji} \quad \text{for } i, j \in \{i, j | \eta_{ij} = 1\} \end{aligned}$$

The details are in the paper, but here K and G are the Gram matrices, respectively, for the constructed embedding and original input. η is a binary matrix, with elements equal to 1 if points i and j are considered neighbors, and 0 otherwise. The first constraint simply requires K to be SPD. The second constraint requires that the entries of K be centered about 0. The third constraint requires that the "isometry" of the output matches the "irometry" of the input for neighboring points.

The intuition here is that maximizing the trace of the output's Gram matrix corresponds to maximizing the squared pairwise distance between all the points in the embedded output. Essentially, we're "pulling apart" or "unrolling" the input data. The problem is made bounded by the third constraint. This requires that for each point and it's k-closest neighbors, pairwise distances and angles are preserved. Thus, the overall goal is to discover a manifold in the input dataset, that when projected to lower dimensions preserves local geometry.

I would like to explore how LADMAP performs on a "tweaked" version of this problem:

$$\begin{aligned} \text{maximize} \quad & ||Z||_* - \lambda E \\ \text{subject to} \quad & K = Z + E \\ & K \succeq 0 \\ & \Sigma_{ij} K_{ij} = 0 \\ & K_{ii} + K_{jj} - K_{ij} - K_{ji} = G_{ii} + G_{jj} - G_{ij} - G_{ji} \quad \text{for } i, j \in \{i, j | \eta_{ij} = 1\} \end{aligned}$$

My motivation is that decomposing the output's Gram matrix into two matrices, admits the use of an "error" margin when performing the optimization. That is, local isometries will only be approximately preserved. Additionally, I'm curious to see the effect of the nuclear norm instead of the trace operator on the Gram matrix.

Sources

The two primary sources for my project have so far been:

- Lin, Zhouchen, Risheng Liu, and Zhixun Su. "Linearized alternating direction method with adaptive penalty for low-rank representation." In Advances in neural information processing systems, pp. 612-620. 2011.
- Weinberger, Kilian Q., and Lawrence K. Saul. "Unsupervised learning of image manifolds by semidefinite programming." International journal of computer vision 70, no. 1 (2006): 77-90.

I've also looked at the following papers for background and motivation

- Cands, Emmanuel J., and Benjamin Recht. "Exact matrix completion via convex optimization." Foundations of Computational mathematics 9, no. 6 (2009): 717.
- Wright, John, Arvind Ganesh, Shankar Rao, Yigang Peng, and Yi Ma. "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization." In Advances in neural information processing systems, pp. 2080-2088. 2009.