

ELE 538: Large-Scale Optimization Project Proposal

Zachary Hervieux-Moore

Sunday 18th March, 2018

Abstract

For the course project, I wish to do a survey on the ADMM literature with an application to neural network training. The goal is to understand ADMM better in a nonconvex setting so that neural network training may be approached. The main paper of interest is “Convergence Analysis of Alternating Direction Method of Multipliers for a Family of Nonconvex Problems, ” M. Hong, Z.Q. Luo, and M. Razaviyayn, 2016. After this paper, I want to review the literature of using ADMM specifically applied to the neural network setting such as “Training Neural Networks Without Gradients: A Scalable ADMM Approach, ” G. Taylor et al., 2016. Finally, I would like to do my own analysis of a few toy neural network architectures as well as do some experiments to see if there are any improvements to training with ADMM in certain architectures.