DISSERTATION PROPOSAL

Lingqing Shen

"Scalable Algorithms and Performance Guarantees for Convex and Nonconvex Optimization"

Friday, November 17, 2023
10:00am
Tepper 5219

This thesis focuses on theoretical and algorithmic developments for convex optimization and nonconvex optimization. It consists of three topics, addressing each of the following problems: convex bilevel optimization, general principle submatrix selection, and strategic classification.

In the first two chapters, we study the convex bilevel optimization problem. In this problem, the goal is to optimize an outer-level convex objective function over a feasible set, defined as the set of minimizers for an inner-level convex objective function. This is of interest in regularization in statistics, sparsity oriented statistics, and fairness in artificial intelligence and machine learning applications. In the first chapter, inspired by the primal-dual method, we propose a general framework of first-order algorithms for solving the convex bilevel optimization problem. We develop and analyze two algorithms within this framework and establish their theoretical convergence guarantees under various structural assumptions, such as nonsmoothness, smoothness, and strong convexity of the functions involved and support our theoretical findings with two numerical studies as well.

In the second chapter, as part of our ongoing work, we are investigating alternative approaches to exploit the smoothness structure in the context of convex bilevel optimization. In addition, we are also interested in developing adaptive stepsize schemes that have minimal dependence on problem parameters to reduce the impact of worst-case parameter estimation on the algorithmic performance.

Our third and fourth chapters focus on the general principle submatrix selection (GPSS) problem that aims to select a principal submatrix of a given positive semidefinite matrix subject to a variety of criteria such as maximizing the determinant and minimizing the trace of the inverse of the submatrix. GPSS problem generalizes the well-known maximum entropy sampling problem (MESP) that arises in statistics and aims to select a small number of random observations from a possibly large set of candidates to maximize the information obtained. The MESP is widely applied in healthcare, power systems, manufacturing, and data science, among others. GPSS problem admits a formulation as minimizing a general convex objective function applied to the eigenvalues of the selected principle submatrix. Typically, GPSS problems, as well as MESP, are NP-hard due to their natural combinatorial selection aspect. Moreover, these problems are highly nonlinear as they study the eigenvalues of submatrices. In the third chapter, we design convex relaxations for GPSS problem through the lens of majorization. Our results generalize two well-known relaxations for MESP, but apply more generally to principle submatrix selection. We illustrate exactness of our relaxations in special cases and preliminary numerical results indicate high quality at certain parameter regimes. In general, these relaxations are incomparable in terms of their quality, neither of them dominates each other uniformly. We also introduce a reformulation-linearization technique to strengthen these relaxations.

In the fourth chapter, we propose to study the algorithmic aspects related to GPSS problem and its relaxations. In particular, we propose to study the quality of the strengthened relaxations both theoretically and numerically, investigate possible connections between them, and seek further strengthening if possible. We also propose to develop efficient algorithms to solve these relaxations at large-scale and convert fractional solutions from these relaxations to high quality solutions to the original nonconvex GPSS problem. To this end, we will develop tools to handle domains involving majorization constraints and approximation and greedy local search algorithms.

In the fifth chapter, we consider the strategic classification problem which aims to take into account the effect of strategic behavior of the agents whose features are modeled as data points that need to be classified. In particular, in this setting, in contrast to the classical binary classification, data points are assumed to arrive in an online fashion and subject to manipulation by an agent, aiming at getting a positive prediction label at a budgeted cost for manipulation. The learner, with only access to the manipulated data, seeks to predict the agent's true labels. In our model, the cost of manipulation is given as the distance between true and manipulated data, where the distance function is a general norm. We first study the case of linear classification and analyze a strategic perceptron algorithm. We provide a necessary and sufficient condition for the convergence of this algorithm to the best classifier and also establish a finite mistake bound for it. Unfortunately, the strategic perceptron algorithm does not offer any margin guarantee. To address this, we propose a new margin maximization algorithm and show that it converges to the optimal classifier in the stochastic setting. Despite its strong theoretical guarantee, our margin maximization algorithm is expensive as it requires to resolve a classification problem with a growing dataset in each iteration. To mitigate this, we further propose and study an efficient variant of this algorithm with a significantly lower per-iteration cost motivated by the joint estimation-optimization approach. We propose to numerically test the performance of these approaches. In addition, we will investigate the effect of strategic behavior in nonlinear classification settings and as well as settings in which the agent's features, i.e., the data points, are discrete.