Department of Advanced Design and Systems Engineering
香港城市大學
City University of Hong Kong

Distinguished ADSE Seminar Series

New Theory and Improved Computational Performance of the Frank-Wolfe Method for Data Science Applications

Professor Robert M. Freund

Theresa Seley Professor in Management Science
Sloan School of Management
Massachusetts Institute of Technology, USA

1 DEC 2021 (Wed)
10:30 am

Seminar Link: **https://cityu.zoom.us/j/99513678654**
Meeting ID: **995 1367 8654**

Enquiry: 3442 8422
All are welcome

Abstract

The Frank-Wolfe method has emerged in the last decade as a very attractive algorithm for optimization in data science applications because the algorithm itself encourages desirable solution structure (such as sparsity or low-rank) without sacrificing computational guarantees. This talk will report on two ongoing research projects related to the Frank-Wolfe method. The first project is focused on the Frank-Wolfe method for empirical risk minimization in the linearly-structured setting (the most often occurring in practice) in the regime where the number of samples n is huge and computing an exact gradient is unattractive. In this setting the most efficient stochastic Frank-Wolfe methods depend linearly on n, even though their proximal counterparts (such as stochastic gradient descent (SGD)) scale independent of n. Here we present a new deterministic approximate-gradient Frank-Wolfe method that substantially reduces the dependence on the sample size n, with improved computational complexity bounds and demonstrably improved computational performance in practice.

The second project is a new generalized Frank-Wolfe method for the class of composite optimization problems (P): minimize_x f(*A*x) + h(x), where f is a θ-logarithmically-homogeneous self-concordant barrier, *A* is a linear operator, and the convex function h has bounded domain but is possibly non-smooth. Problem instances of (P) include the D-optimal design problem in experimental design, the minimum volume covering ellipsoid problems in convex geometry, Poisson image deblurring problems with TV (Total Variation) regularization, and PET (positron emission tomography) image reconstruction. The lack of smoothness of the objective function of (P) has been a bottleneck to the development of both theory and practical computational performance for the Frank-Wolfe method as well as other first-order methods that would otherwise be attractive when the problem dimensions are huge. We develop a new theory that yields attractive computational complexity for this class of problems. Our theoretical results rely on just a few problem-instance measures that arise naturally for this problem class, and that point to certain intrinsic connections between θ-logarithmically homogeneous barriers and the Frank-Wolfe method. Last of all, we present computational experiments that point to the potential usefulness of our generalized Frank-Wolfe method on Poisson image de-blurring problems with TV regularization, and on simulated PET problem instances.

This research is joint with students Renbo Zhao and Zikai Xiong of the MIT Operations Research Center.

Read more: www.cityu.edu.hk/adse/seminar_2021-22_5.htm