The Second Greater Bay Area Workshop on Computational Optimization

Department of Applied Mathematics (AMA)
The Hong Kong Polytechnic University (PolyU)

10--11 December 2021

In hybrid mode: onsite and on Zoom

Sponsors: AMSS-PolyU Joint Lab and AMA

中国科学院数学与系统科学研究院
Academy of Mathematics & Systems Science, CAS

DEPARTMENT OF APPLIED MATHEMATICS
應 用 數 學 系

The purpose of this workshop is to provide a platform for researchers and practitioners from the Greater Bay Area and beyond to exchange ideas and share their views on solving challenging optimization problems in data science and machine learning.

Zoom link: https://polyu.hk/VSzJM

Friday, 10 December

Time	Speaker	Title	Chair
8:50–9:00	Christopher Chao VPRI of PolyU (Onsite)	Opening Speech	Defeng Sun
9:00–9:40	Bo Bai Theory Lab, Huawei	10 Challenging Problems in Post Shannon Times	Xiaojun Chen
9:40–10:20	Kim-Chuan Toh National University of Singapore	An Inexact Projected Gradient Method with Rounding and Lifting by Nonlinear Programming for Solving Rank-One Semidefinite Relaxation of Polynomial Optimization	
10:20–10:30	Discussion Break		
10:30–11:10	Michael Friedlander University of British Columbia	Communication and Fairness in Federated Learning	Xiaoqi Yang
11:10–11:50	Huifu Xu The Chinese University of Hong Kong (Onsite)	Preference Robust Optimization	
11:50–12:30	Xiyin Zheng Yunnan University	Well-Posed Solvability of Convex Optimization Problems on Differentiable Closed Convex Sets	
12:30–14:30	Lunch		
14:30–15:10	Xin Liu Chinese Academy of Sciences	A Global Optimization Approach for Multi-Marginal Optimal Transport Problems with Coulomb Cost	Zaikun Zhang
15:10–15:50	Anthony Man-Cho So The Chinese University of Hong Kong (Onsite)	Non-Convex Exact Community Recovery in Stochastic Block Model	
15:50–16:00	Discussion Break		
16:00–16:40	Michael Zhang Theory Lab, Huawei (Onsite)	Practical Experiences and Case Studies of Solving Large-Scale LP/ILP Problems in the ICT Industrial Area	Defeng Sun
16:40–17:20	Rui Zhang Theory Lab, Huawei (Onsite)	Optimization in Compressed Sensing with Applications	

Saturday, 11 December

Time	Speaker	Title	Chair
9:00–9:40	Zhi-Quan Luo The Chinese University of Hong Kong, Shenzhen	Learning to Maximize a Convex Quadratic Function with Application to Intelligent Reflection Surface for Wireless Communication	Xiaojun Chen
9:40–10:20	Yin Zhang The Chinese University of Hong Kong, Shenzhen	Distributed (Sparse) PCA: Seeking Consensus on a Subspace	
10:20–10:30	Discussion Break		
10:30 – 11:10	Karthik Natarajan Singapore University of Technology and Design	Robustness to Dependency in Influence Maximization	Ting Kei Pong
11:10–11:50	Sunyoung Kim Ewha W. University	Doubly Nonnegative Relaxations and Completely Positive Reformulations of Quadratic Optimization Problems with Block-Clique Graph Structures	
11:50–12:30	Shaohua Pan South China University of Technology	Column $l_{2,0}$-Norm Regularized Factorization Model of Low-Rank Matrix Recovery and its Computation	
12:30–14:30	Lunch		
14:30–15:10	Chao Ding Chinese Academy of Sciences	A Semismooth Newton Based Augmented Lagrangian Method for Nonsmooth Optimization on Matrix Manifolds	Man-chung Yue
15:10–15:50	Yaohua Hu Shenzhen University	Mix Sparse Optimization: Theory and Applications	
15:50–16:00	Discussion Break		
16:00–16:40	Hou-Duo Qi University of Southampton	Newton Hard-Thresholding Pursuit for Sparse Optimization	Xiaoqi Yang
16:40–17:20	Jacek Gondzio University of Edinburgh	Sparse Approximations with Interior Point Methods	

10 Challenging Problems in Post Shannon Times

Bo Bai
Theory Lab, Huawei

近年来，以 5G，大数据，云计算和人工智能为代表的信息科学与技术飞速发展，给人们的生产，生活和学习方式带来了深刻的变革。这些技术的发展给信息科学的基础理论——信息论带来了全新的挑战。本报告将结合信息技术发展趋势和华为公司的实际研发需求，简要介绍后香农时代的十大挑战问题。期望与各位专家展开研讨合作，共同努力突破信息科学中的挑战问题，打造后香农时代的新信息理论。

白铂博士，2010 年毕业于清华大学电子系，获工学博士学位，清华大学电子系学术新秀。2010 年 10 月至 2012 年 3 月，在香港科技大学做博士后研究，任副研究员。2012 年 7 月至 2017 年 1 月，在清华大学电子系任教，获得清华大学骨干人才计划支持。目前，他担任华为理论研究部主任、信息论高级研究员，研究方向为信息论、6G 无线网络、图信息学等。他是 IEEE 高级会员，在 IEEE/ACM 的主要期刊和会议上发表学术论文 130 余篇，专著章节 2 章，清华大学电子系核心课教材 1 部。

A Semismooth Newton Based Augmented Lagrangian Method for Nonsmooth Optimization on Matrix Manifolds

Chao Ding
Chinese Academy of Sciences

This talk is devoted to studying an augmented Lagrangian method for solving a class of manifold optimization problems, which have nonsmooth objective functions and non-negative constraints. Under the constant positive linear dependence condition on manifolds, we show that the proposed method converges to a stationary point of the nonsmooth manifold optimization problem. Moreover, we propose a globalized semismooth Newton method to solve the augmented Lagrangian subproblem on manifolds efficiently. The local superlinear convergence of the manifold semismooth Newton method is also established under some suitable conditions. Finally, numerical experiments on compressed modes and (constrained) sparse principal component analysis illustrate the advantages of the proposed method.

Communication and Fairness in Federated Learning

Michael P. Friedlander
University of British Columbia

Federated learning is a decentralized machine learning scheme that allows many data owners to train a model collaboratively without sharing private data. We describe a dual decomposition approach that leads to efficient communication among the participants. Because participation is best encouraged by fairly rewarding owners for the quality of their data, we also propose a variation of the Shapley value from cooperative game theory that is computationally efficient.

Sparse Approximations with Interior Point Methods

Jacek Gondzio
University of Edinburgh

I will briefly introduce the key ideas which make interior point methods (IPMs) such a powerful optimization technique, and provide arguments that they can work efficiently with inexact Newton directions computed with iterative (Krylov-subspace) methods. Then, I will interpret IPM as a homotopy type approach and link it with other (relaxed) variants of IPMs including the preconditioned Newton conjugate gradient method. Then I will present some very recent developments which aim at solving a variety of sparse approximation problems using highly specialized IPMs. Those include problems arising from: compressed sensing, multi-period portfolio optimization, classification of data coming from functional Magnetic Resonance Imaging, restoration of images corrupted by Poisson noise, and classification via regularized logistic regression. In all these cases the performance of IPM will be compared against competitive first-order methods. The results reveal that specialized IPM compares favourably and often outperforms the cutting-edge first-order methods.

This is a joint work with: Valentina De Simone, Daniela di Serafino, Spyridon Pougkakiotis, and Marco Viola

For more details, see:

[1] V. De Simone, D. di Serafino, J. Gondzio, S. Pougkakiotis, and M. Viola Sparse approximations with interior point methods, SIAM Review (accepted 24 Nov 2021).

Mix Sparse Optimization: Theory and Applications

Yaohua Hu
Shenzhen University

In this talk, we will consider a mix sparse optimization problem, that is, sparsity structures at intra-group and inter-group levels are considered simultaneously. This mix sparse structure widely appears in applications. For the mix sparse optimization problem, we will discuss the nonconvex regularization method, as well as a first-order iterative algorithm, and present its consistency theory, asymptotic theory and convergence theory. Applications to gene regulatory networks and differential optical absorption spectroscopy will be presented.

Doubly Nonnegative Relaxations and Completely Positive Reformulations of Quadratic Optimization Problems with Block-Clique Graph Structures

Sunyoung Kim
Ewha W. University

We study the equivalence among a nonconvex QOP, its CPP and DNN relaxations under the assumption that the aggregate and correlative sparsity of the data matrices of the CPP relaxation is represented by a block-clique graph G. By exploiting the correlative sparsity, we decompose the CPP relaxation problem into a clique-tree structured family of smaller subproblems. We establish the equivalence between the QOP and its DNN relaxation from the equivalence between the reduced family of subproblems and their DNN relaxations by applying the known results on: (i) CPP and DNN reformulation of a class of QOPs with linear equality, complementarity and binary constraints in 3 nonnegative variables. (ii) DNN reformulation of a class of quadratically constrained convex QOPs with any size. (iii) DNN reformulation of LPs with any size.

A Global Optimization Approach for Multi-Marginal Optimal Transport Problems with Coulomb Cost

Xin Liu
Chinese Academy of Sciences

In this work, we construct a novel numerical method for solving the multi-marginal optimal transport problems with Coulomb cost. This type of optimal transport problems arises in quantum physics and plays an important role in understanding the strongly correlated quantum systems. With a Monge-like ansatz, the orginal high-dimensional problems are transferred into mathematical programmings with generalized complementarity constraints, and thus the curse of dimensionality is surmounted. However, the latter ones are themselves hard to deal with from both theoretical and practical perspective. Moreover in the presence of nonconvexity, brute-force searching for global solutions becomes prohibitive as the problem size grows large. To this end, we propose a global optimization approach for solving the nonconvex optimization problems, by exploiting an efficient proximal block coordinate descent local solver and an initialization subroutine based on hierarchical grid refinements. We provide numerical simulations on some typical physical systems to show the efficiency of our approach. The results match well with both theoretical predictions and physical intuitions, and give the first visualization of optimal transport maps for some two dimensional systems.

Learning to Maximize a Convex Quadratic Function with Application to Intelligent Reflection Surface for Wireless Communication

Zhi-Quan Luo
The Chinese University of Hong Kong, Shenzhen

In this talk we consider learning and optimizing a rank-2 convex quadratic function over N variables, each taking K discrete values on the unit circle. This problem arises from optimal design of a passive beamformer for intelligent reflecting surface (IRS) in order to maximize the overall channel strength. When $K = 2$ and the quadratic function (or channel state information) is known, we propose a linear time algorithm that is capable of reaching a globally optimal solution of the problem. When the quadratic function is unknown (i.e. in the absence of channel state information) we propose a random max sampling (RMS) method and a conditional sample mean (CSM) method to maximize the quadratic function. We show that RMS method can provide a SNR boost that is linear in N (the number of reflective elements in IRS), while the CSM can boost the SNR quadratically (in N), all with polynomial number of samples. Field trial results demonstrate the effectiveness of the proposed CSM method in the commercial 5G communication networks, providing >10dB SNR gain in both typical indoor and outdoor scenarios, and with no need to modify the current communication protocols and design.

This is a joint work with Yaowen Zhang, Shuyi Ren, and Kaiming Shen.

Robustness to Dependency in Influence Maximization

Karthik Natarajan
Singapore University of Technology and Design

In this paper, we introduce a correlation robust model for the influence maximization problem. Unlike the classic independent cascade model, this model's diffusion process is adversarially adapted to the choice of seed set. More precisely, rather than only the independent coupling of known individual edge probabilities, we now evaluate a seed set's expected influence under all possible correlations - specifically, the one that presents the worst-case. We show that any seed set's worst-case expected influence can be efficiently computed, and though optimizing the worst-case (over seed sets) is NP-hard, a $(1 - 1/e)$ approximation algorithm can be obtained.

Column $\ell_{2,0}$-Norm Regularized Factorization Model of Low-Rank Matrix Recovery and its Computation

Shaohua Pan
South China University of Technology

This paper is concerned with the column $\ell_{2,0}$-regularized factorization model of low-rank matrix recovery problems and its computation. The column $\ell_{2,0}$-norm of factor matrices is introduced to promote column sparsity of factors and lower rank solutions. For this nonconvex discontinuous optimization problem, we first demonstrate its superiority from theory, and then develop an alternating majorization-minimization (AMM) method with extrapolation and a hybrid AMM, in which a majorized alternating proximal method is proposed to seek an initial factor pair with less nonzero columns and the AMM with extrapolation is then employed to minimize of a smooth nonconvex loss. We provide the global convergence analysis for the proposed AMM methods and apply them to the matrix completion problem with non-uniform sampling schemes. Numerical experiments are conducted with synthetic and real data examples, and comparison results with the nuclear-norm regularized factorization model and the max-norm regularized convex model show that the column $\ell_{2,0}$-regularized factorization model has an advantage in offering solutions of lower error and rank within less time.

Newton Hard-Thresholding Pursuit for Sparse Optimization

Hou-Duo Qi
University of Southampton

Algorithms based on the hard thresholding principle have been well studied with sounding theoretical guarantees for sparsity-constrained optimization. It is widely observed in existing empirical studies that when a restricted Newton step was used (as the debiasing step), the hard-thresholding algorithms tend to meet halting conditions in a significantly low number of iterations and hence are very efficient. However, the thus obtained Newton hard-thresholding algorithms do not offer any better theoretical guarantees than their simple hard-thresholding counterparts. This annoying discrepancy between theory and empirical studies has been known for some time.

In this talk, we report our attempts in resolving this discrepancy. We provide a theoretical justification for the use of a restricted Newton step in the algorithm of Newton Hard-Thresholding Pursuit (NHTP). Our main result shows that NHTP is quadratically convergent under the standard assumption of restricted strong convexity and smoothness. We also establish its global convergence to a stationary point under a weaker assumption. In the special case of the compressive sensing, NHTP eventually reduces to some existing hard-thresholding algorithms with a Newton step. Consequently, our fast convergence result justifies why those algorithms perform better than without the Newton step. The efficiency of NHTP was demonstrated on both synthetic and real data in compressed sensing and sparse logistic regression (Joint work with Zhou Shenglong and Xiu Naihua).

Non-Convex Exact Community Recovery in Stochastic Block Model

Man-Cho Anthony So
The Chinese University of Hong Kong

In this talk, we consider the problem of exact community recovery in a graph that is generated by the symmetric stochastic block model. Although the maximum likelihood estimation formulation is non-convex and has binary variables, we show that in the logarithmic sparsity regime of the problem, a properly initialized projected gradient-type method can achieve exact recovery of the communities down to the information-theoretic limit in nearly linear time. This demonstrates the potential of non-convex methods in tackling large-scale community recovery problems.

The talk covers joint works with Huikang Liu, Peng Wang, and Zirui Zhou.

An Inexact Projected Gradient Method with Rounding and Lifting by Nonlinear Programming for Solving Rank-One Semidefinite Relaxation of Polynomial Optimization

Kim-Chuan Toh
National University of Singapore

We consider solving high-order semidefinite programming (SDP) relaxations of polynomial optimization problems (POPs) that often admit degenerate rank-one optimal solutions. Instead of solving the SDP alone, we propose a new algorithmic framework that blends local search using the non-convex POP into global descent using the convex SDP. In particular, we first design a globally convergent inexact projected gradient method (iPGM) for solving the SDP that serves as the backbone of our framework. We then accelerate iPGM by taking long, but safeguarded, rank-one steps generated by fast nonlinear programming algorithms. We prove that the new framework is still globally convergent for solving the SDP. To solve the iPGM subproblem of projecting a given point onto the feasible set of the SDP, we design a two-phase algorithm with phase one using a symmetric Gauss-Seidel based accelerated proximal gradient method to generate a good initial point, and phase two using a modified limited-memory BFGS method to obtain an accurate solution. We conduct numerical experiments for solving second-order SDP relaxations arising from a diverse set of POPs. Our framework demonstrates state-of-the-art efficiency, scalability, and robustness in solving degenerate rank-one SDPs to high accuracy, even in the presence of millions of equality constraints.

Preference Robust Optimization

Huifu Xu
The Chinese University of Hong Kong

Preference robust optimization concerns decision making problems where the decision maker's utility/risk preference is ambiguous and the optimal decision is based on the worst case utility function or risk measure from a set of plausible utility functions/risk measures constructed with partially available information. In this talk, we will discuss some PRO models based on the expected utility theory and general choice functions. We will also discuss the PRO models in risk management and some computational schemes.

Practical Experiences and Case Studies of Solving Large-Scale LP/ILP

Michael Zhang
Theory Lab, Huawei

Optimization solvers have been widely used in the industry of information and communication technology (ICT) in the past decades. As the recent development of advanced ICT such as 5G, AI and smart network, industry has been facing very large-scale optimization problems which become the core of realizing the technologies. However, such optimization problems have seldomly been touched or fully exploited before. Most of these problems are sparse LP/ILP and the problem sizes are from millions to billions in terms of the number of variables and constraints. The requirement of solving these large-scale LP/ILP problems differs in the dimensions of feasibility, optimality, and time efficiency. How to strike a balance among these factors is the key to meet the stringent requirement from the practical scenarios so as to really deploy the solver-based solutions in the products. In this talk, we are going to discuss how to fill the gap between the theoretical algorithm design and practical deployment by walking you through our experiences using case studies.

Optimization in Compressed Sensing with Applications

Rui Zhang
Theory Lab, Huawei

In this talk, I will present some applications of compressed sensing algorithms in industry. High performance and low complexity are two goals which we are pursuing, and I will clarify the exact meanings under variant circumstances. The differences of optimization between academia and industry are mainly discussed in the field of compressed sensing.

Distributed (Sparse) PCA: Seeking Consensus on a Subspace

Yin Zhang
The Chinese University of Hong Kong, Shenzhen

We propose and study a distributed algorithm for doing principle component analysis (PCA) in a centralized network in which each node privately holds a subset of the dataset and only exchanges "safe" information with a center server during the analysis process. In the framework of alternating direction methods of multipliers (ADMM), we propose a novel formulation for building consensus by equalizing subspaces spanned by splitting variables instead of equalizing the variables themselves. This technique greatly relaxes feasibility restrictions and significantly reduces the rounds of communication. Moreover, it is practically improbable for any node to uncover private data stored in any other node based on publicly shared quantities in the algorithm, which is not the case in many existing distributed or parallelized algorithms. We will also discuss extending this approach to the sparse PCA problem.

Well-Posed Solvability of Convex Optimization Problems on Differentiable Closed Convex Sets [1]

Xiyin Zheng
Yunnan University

Motivated by Gale and Klee's continuity of a closed convex set (Math. Scand. 7, 370-391, 1959), several kinds of differentiability for a closed convex set are introduced by its support functional. Given a closed convex set A in a Banach space X, using technique of variational analysis, we prove that A is C^p-differentiable (Fréchet differentiable or Gateaux differentiable) if and only if *for every continuous linear (or convex) function* $f : X \to \mathbb{R}$ bounded below on A the constrained convex optimization problem $\inf\limits_{x\in A} f(x)$ is well-solvable in the sense of the corresponding well-posedness. In the reflexive space case, it is proved that A is continuous if and only if for every continuous linear (or convex) function $f : X \to \mathbb{R}$ bounded below on A the corresponding optimization problem $\inf\limits_{x\in A} f(x)$ is weakly well-posed solvable. We also prove that if the conjugate function f^* of a given continuous convex function f on X is Fréchet differentiable (resp. continuous) on $\mathrm{dom}(f^*)$ then *for every closed convex set* K in X with $\inf\limits_{x\in K} f(x) > -\infty$ the corresponding optimization problem with objective f and constraint set K is well-posed (resp. weakly well-posed) solvable. In the framework of finite-dimensional spaces, several sharper results are established.

[1] This research was supported by the National Natural Science Foundation of P. R. China (Grant No. 11771384).