2021 年"优化与应用"学术研讨会

会 议 日 程

主办方：中国科学院数学与系统研究院优化与应用研究中心

时　　间：2021 年 12 月 4 日

腾讯会议：744 487 261

日期	时间	报告信息
12 月 4 日（周六）上午	08:50－09:00	开幕式： 优化与应用研究中心名誉主任 袁亚湘 致辞 中国科学院计算数学与科学计算工程研究所所长 周爱辉 致辞
	09:00－10:00	主持人：郭田德
	09:00－09:30	**Zhijun Wu（Iowa State University, USA）** Social Distancing as a Social Dilemma Game and its Optimal Strategies
	09:30－10:00	印卧涛（阿里巴巴达摩院） Closing the Gap: Alternating SGD Methods Work Well for Bilevel Problems
	10:00－11:30	主持人：刘歆
	10:00－10:30	**Feng Ruan（University of California, Berkeley）** Adapting Maximum Likelihood Theory to Modern Applications
	10:30－11:00	李肖（香港中文大学（深圳）） Convergence of Random Reshuffling Under The Kurdyka-Lojasiewicz Inequality
	11:00－11:30	包承龙（清华大学） Anderson Acceleration for Training Neural Networks
	11:30－14:00	午休
	14:00－16:00	主持人：马俊杰
	14:00－14:30	凌舒扬（上海纽约大学） Generalized Power Method for Generalized Orthogonal Procrustes Problem

12月4日（周六）下午	14:30－15:00	史斌（中国科学院数学与系统科学研究院） Understanding the Acceleration Phenomenon via High-Resolution Differential Equations	
	15:00－15:30	张国涵（北京邮电大学） Recent advances in Extended Second-order cones	
	15:30－16:00	顾然（南开大学） A Random Active Set Method For Strictly Convex Quadratic Problem With Simple Bounds	
	16:00－17:30	主持人：刘亚锋	
	16:00－16:30	阴佳腾（北京交通大学） Benders 分支切割法及其在结构化整数规划模型中的应用	
	16:30－17:00	陈亮（中国科学院数学与系统科学研究院） An Exact Separation Algorithm for Unsplittable Flow Capacitated Network Design Arc-set Polyhedron	
	17:00-17:30	**Houduo Qi（University of Southampton）** Proximal Distance Algorithms for Euclidean Distance Matrix Optimization	
	17:30－17:40	闭幕式： 优化与应用研究中心学术委员会副主任 杨新民 致辞 优化与应用研究中心主任 戴彧虹 致辞	

附：报告题目与摘要

09:00－09:30
报告人：**Zhijun Wu（Iowa State University, USA）**
报告题目：**Social Distancing as a Social Dilemma Game and its Optimal Strategies**
报告摘要：Since the outbreak of the global COVID-19 pandemic, social distancing has been known to everyone and recommended almost everywhere every day. Social distancing has been and will be one of the most effective measures and sometimes, the only available one for fighting epidemics and saving lives. However, it has not been so clear how social distancing should be practiced or managed, especially when it comes to regulating everyone's otherwise normal social activities. The debate on how to implement social distancing often leads to a heated political argument, while research on the subject is lacking. In this talk, I will discuss a theoretical framework for the understanding of the scientific nature of social distancing by considering social distancing as a social dilemma game played by every individual against his/her population. From this perspective, every individual needs to make decision on how to engage in social distancing, or risk being trapped into a dilemma either exposing to deadly diseases or getting no access to necessary social activities. As the players of the game, the individual's decisions depend on the population's actions and vice versa, and an optimal strategy can be found when the game reaches an equilibrium. I will show how an optimal strategy can be determined for a population with either closely related or completely separated social activities and with either single or multiple social groups, and how the collective behaviors of social distancing can be simulated by following every individual's actions as the distancing game progresses. The simulation results for populations of varying sizes and complexities will be presented, which not only justify the choices of the strategies based on the game theoretic analysis, but also demonstrate the convergence of the individual actions to an optimal distancing strategy in silico and possibly in natura as well, if every individual makes rational distancing decisions.

09:30－10:00
报告人：印卧涛（阿里巴巴达摩院）
报告题目：**Closing the Gap: Alternating SGD Methods Work Well for Bilevel Problems**
报告摘要：This talk briefly overviews a few stochastic nested optimization problems, including stochastic bilevel, min-max, and compositional optimization, which quickly gain popularity in machine learning (meta learning, reinforcement learning, etc.). They share the same nested structure, but existing works treat them separately, developing problem-specific algorithms and analyses. Although one may think SGD is slower on nested problems than on single-level problems, it is not the case. We develop a method called ALternating Stochastic gradient dEscenT (ALSET), which leverages hidden smoothness and achieves the same complexity of $O(eps^{-2})$. Under certain regularity conditions, applying our results to stochastic compositional,

min-max, and reinforcement learning problems either improves upon or matches the best-known sample complexity in the respective cases.

This is joint work with Prof. Tianyi Chen (RPI) and Dr. Yuejiao Sun (recently graduated from UCLA).

10:00－10:30
报告人：**Feng Ruan（University of California, Berkeley）**
报告题目：**Adapting Maximum Likelihood Theory to Modern Applications**
报告摘要：Maximum likelihood estimation (MLE) is influential because it can be easily applied to generate optimal, statistically efficient procedures for broad classes of estimation problems. Nonetheless, the theory does not apply to modern settings—such as problems with computational, communication, or privacy considerations—where our estimators have resource constraints. In this talk, I will introduce a modern maximum likelihood theory that addresses these issues, focusing specifically on procedures that must be computationally efficient or privacy-preserving. To do so, I first derive analogues of Fisher information for these applications, which allows a precise characterization of tradeoffs between statistical efficiency, privacy, and computation. To complete the development, I also describe a recipe that generates optimal statistical procedures (analogues of the MLE) in the new settings, showing how to achieve the new Fisher information lower bounds.

10: 30－11:00
报告人：**李肖（香港中文大学（深圳））**
报告题目：**Convergence of Random Reshuffling Under The Kurdyka-Lojasiewicz Inequality**
报告摘要：We study the random reshuffling (RR) method for smooth nonconvex optimization problems with a finite-sum structure. Though this method is widely utilized in practice, e.g., in the training of neural networks, its convergence behavior is only understood in several limited settings. In this paper, under the well-known Kurdyka-Lojasiewicz (KL) inequality, we establish strong limit-point convergence results for RR with appropriate diminishing step sizes, namely, the whole sequence of iterates generated by RR is convergent and converges to a single stationary point in an almost sure sense. In addition, we derive the corresponding rate of convergence, depending on the KL exponent and suitably selected diminishing step sizes. When the KL exponent lies in [0,1/2], the convergence is at a rate of $O(t^{-1})$ with t counting the iteration number. When the KL exponent belongs to (1/2,1), our derived convergence rate is of the form $O(t^{-q})$ with q in (0,1) depending on the KL exponent. The standard KL inequality-based convergence analysis framework only applies to algorithms with a certain descent property. We conduct a novel convergence analysis for the non-descent RR method with diminishing step sizes based on the KL inequality, which generalizes the standard KL framework. We summarize our main steps and core ideas in an informal analysis framework, which is of independent interest. As a direct application of this framework, we also establish

similar strong limit-point convergence results for the shuffled proximal point method.

11:00－11:30
报告人：包承龙（清华大学）
报告题目：**Anderson Acceleration for Training Neural Networks**
报告摘要：Anderson mixing is a powerful method in scientific computing, but its convergence analysis has not been fully explored. In this talk, we add the adaptive regularization and damped regularization to the original Anderson mixing, and provide the detailed convergence analysis and iteration complexity results. Moreover, to reduce the memory burden, we propose a short-term version that can mimic long-term memory. Various experiments on deep neural networks are demonstrated for the advantages of this method.

14:00－14:30
报告人：凌舒扬（上海纽约大学）
报告题目：**Generalized Power Method for Generalized Orthogonal Procrustes Problem**
报告摘要：Given a set of multiple point clouds, how to find the rigid transformations (rotation, reflection, and shifting) such that these point clouds are well aligned? This problem, known as the generalized orthogonal Procrustes problem (GOPP), plays a fundamental role in several scientific disciplines including statistics, imaging science and computer vision. Despite its tremendous practical importance, it is still a challenging computational problem due to the inherent nonconvexity. In this paper, we study the semidefinite programming (SDP) relaxation of the generalized orthogonal Procrustes problems and prove that the tightness of the SDP relaxation holds, i.e., the SDP estimator exactly equals the least squares estimator, if the signal-to-noise ratio (SNR) is relatively large. We also prove that an efficient generalized power method with a proper initialization enjoys global linear convergence to the least squares estimator. In addition, we analyze the Burer-Monteiro factorization and show the corresponding optimization landscape is free of spurious local optima if the SNR is large. This explains why first-order Riemannian gradient methods with random initializations usually produce a satisfactory solution despite the nonconvexity. One highlight of our work is that the theoretical guarantees are purely algebraic and do not require any assumptions on the statistical property of the noise. Our results partially resolve one open problem posed in [Bandeira, Khoo, Singer, 2015] on the tightness of the SDP relaxation in solving the generalized orthogonal Procrustes problem. Numerical simulations are provided to complement our theoretical analysis.

14:30－15:00
报告人：史斌（中国科学院数学与系统科学研究院）
报告题目：**Understanding the Acceleration Phenomenon via High-Resolution Differential Equations**
报告摘要：Gradient-based optimization algorithms can be studied from the

perspective of limiting ordinary differential equations (ODEs). Motivated by the fact that existing ODEs do not distinguish between two fundamentally different algorithms---Nesterov's accelerated gradient method for strongly convex functions (NAG-SC) and Polyak's heavy-ball method---we study an alternative limiting process that yields high-resolution ODEs. We show that these ODEs permit a general Lyapunov function framework for the analysis of convergence in both continuous and discrete time. We also show that these ODEs are more accurate surrogates for the underlying algorithms; in particular, they not only distinguish between NAG-SC and Polyak's heavy-ball method, but they allow the identification of a term that we refer to as "gradient correction" that is present in NAG-SC but not in the heavy-ball method and is responsible for the qualitative difference in convergence of the two methods. We also use the high-resolution ODE framework to study Nesterov's accelerated gradient method for (non-strongly) convex functions, uncovering a hitherto unknown result---that NAG-C minimizes the squared gradient norm at an inverse cubic rate. Finally, by modifying the high-resolution ODE of NAG-C, we obtain a family of new optimization methods that are shown to maintain the accelerated convergence rates of NAG-C for smooth convex functions.

15:00－15:30
报告人：张国涵（北京邮电大学）
报告题目：**Recent advances in Extended Second-order cones**
报告摘要：The extended second-order cones(ESOC) are introduced in 2015. This extension seems the most natural extension of second order cones. In this talk, I will first briefly review some basic properties of ESOC. Then I will introduce the monotone extended second order cone (MESOC), which is related to the monotone cone and the second order cone. Finally, I will show some applications of ESOC and MESOC in portfolio optimization.

15:30－16:00
报告人：顾然（南开大学）
报告题目：**A Random Active Set Method For Strictly Convex Quadratic Problem With Simple Bounds**
报告摘要：Active set method aims to find the correct active set of the optimal solution and is a powerful method for solving strictly convex quadratic problem with bound constraints. To guarantee finite step convergence, the existing active set methods all need strict conditions or some additional strategies, which greatly affects the efficiency of the algorithm. In this paper, we propose a random active set method which introduces randomness in the update of active set. We prove that it can converge in finite steps with probability one without any conditions on the problem or any additional strategies. Numerical results show that the algorithm obtains the correct active set within a few iterations, and compared with the existing methods, it has good robustness and efficiency.

16:00－16:30

报告人：阴佳腾（北京交通大学）

报告题目：**Benders 分支切割法及其在结构化整数规划模型中的应用**

报告摘要：Benders 分解是一种常用于求解大规模结构化线性规划（LP）问题的经典方法。然而，对于子问题带有离散整数变量的优化模型，由于其子问题对应的 Recourse function 不再具有分段线性特征，因此无法使用传统 Benders 分解方法求解。为求解带有离散变量的结构化整数规划模型，本文提出一种基于 Benders 分支切割（Branch-and-cut）框架的近似算法，并推导了基于线性松弛影子价格与整数可行解的"Benders 对偶整数割"。通过迭代求解子问题的 LP 松弛模型以及 IP 可行解（或最优解），根据主问题产生的松弛下界与补偿函数值的大小动态加入 Benders 对偶整数割，直至得到高质量可行解。最后，基于北京地铁实际问题：车辆选址方案与运行图一体优化模型进行了数值实验，通过与现有 MIP 商用求解器 CPLEX 的计算效果对比，验证了所提出算法在计算时间、优化精度、可计算模型规模等方面的优势。

16:30－17:00

报告人：陈亮（中国科学院数学与系统科学研究院）

报告题目：**An Exact Separation Algorithm for Unsplittable Flow Capacitated Network Design Arc-set Polyhedron**

报告摘要：In this talk, we concentrate on generating cutting planes for the unsplittable capacitated network design problem. We use the unsplittable flow arc-set polyhedron of the considered problem as a substructure and generate cutting planes by solving the separation problem over it. To relieve the computational burden, we show that, in some special cases, a closed form of the separation problem can be derived. For the general case, a brute-force algorithm, called exact separation algorithm, is employed in solving the separation problem of the considered polyhedron such that the constructed inequality guarantees to be facet-defining. Furthermore, a new technique is presented to accelerate the exact separation algorithm, which significantly decreases the number of iterations in the algorithm. Finally, a comprehensive computational study on the unsplittable capacitated network design problem is presented to demonstrate the effectiveness of the proposed algorithm.

17:00－17:30

报告人：Houduo Qi（University of Southampton）

报告题目：**Proximal Distance Algorithms for Euclidean Distance Matrix Optimization**

报告摘要：Proximal distance algorithms combine the classical quadratic penalty method of constrained minimization with distance majorization. Those algorithms have been independently developed in various areas. For convex differentiable problems, they reduce to proximal gradient algorithms and therefore enjoy well understood convergence properties. However, for nonconvex problems, it is challenging to establish satisfactory convergence results other than a decreasing objective sequence being generated. For the Euclidean Distance Matrix optimization, which is nonconvex and nondifferentiable, strong convergence properties can be

established for one such proximal distance algorithm by taking advantage of the structural properties of the problem. We demonstrate its efficiency for the examples from supervised maximum variance unfolding.