2022 PKU Workshop on
Operations Research and Machine Learning

Nov 21-22, 2022
BEIJNG, CHINA

http://conference.bicmr.pku.edu.cn/meeting/index?id=102

2022 PKU Workshop on
Operations Research and Machine Learning

Nov 21-22, 2022

BEIJING, CHINA

http://conference.bicmr.pku.edu.cn/meeting/index?id=102

————————————————————

Information for Participants

Sponsors

Committees

Conference Schedule

Abstracts

Contact Information

If you need any help, please feel free to contact

<div align="center">

Ms. Ran Liu: cmlr@pku.edu.cn

Ms. Xiaoni Tan: xntan@bicmr.pku.edu.cn

</div>

Sponsors

Committees

Scientific Committee

Weinan E, Peking Univeristy

Yaxiang Yuan, Chinese Academy of Sciences

Pingwen Zhang, Peking Univeristy

Organizing Committee

Bin Dong, Peking University

Jie Song, Peking University

Zaiwen Wen, Peking University

Kun Yuan, Peking Univeristy

Conference Schedule

日 期	时 间	报告信息
11 月 21 日 （周一） 腾讯会议： 527-322-790	08:40-09:00	嘉宾致辞 鄂维南（北京大学） 赵桂萍（基金委数理学部） 霍　红（基金委管理学部） 袁亚湘（中科院数学与系统科学研究院）
	09:00-10:10	主持人: 袁亚湘（中科院数学与系统科学研究院）
	09:00-09:35	戴彧虹（中科院数学与系统科学研究院） Optimality Conditions and Numerical Algorithms for A Class of Linearly Constrained Minimax Optimization
	09:35-10:10	王彦飞（中科院地质与地球物理研究所） Model-driven and Data-driven Inverse Problems with AI Analysis
	10:10-10:20	休息
	10:20-12:05	主持人：宋洁（北京大学）
	10:20-10:55	李志泽（美国卡内基梅隆大学） Efficient Algorithms for Optimization and Federated Learning
	10:55-11:30	唐聿劼（北京大学） 零阶优化及其在多智能体系统优化与强化学习中的应用
	11:30-12:05	王剑晓（北京大学） Toward Tractable Lithiumion Battery Modeling: Multiphysics Analytics and Grid Integration
	12:05-13:45	午休
	13:45-15:30	主持人：丁超（中科学数学与系统科学研究院）
	13:45-14:20	刘亚锋（中科院数学与系统科学研究院） Uplink-Downlink Duality in Wireless Communications: Where Lagrange Meets Shannon

	14:20-14:55	王勇（中科院数学与系统科学研究院） 基因调控网络推断的非凸优化
	14:55-15:30	高斌（中科院数学与系统科学研究院） New Riemannian Preconditioned Algorithms for Tensor Completion via Polyadic Decomposition
	15:30-15:50	休息
	15:50-17:00	主持人：徐大川（北京工业大学）
	15:50-16:25	韩丛英（中国科学院大学） 求解旅行商问题的高泛化性强化学习算法
	16:25-17:00	何冠楠（北京大学） Time Aggregation based on adaptive clustering for low carbon energy system optimization

日期	时间	报告信息
11 月 22 日 （周二） 腾讯会议： 527-322-790	09:00-10:10	主持人：郭田德（中国科学院大学）
	09:00-09:35	韩德仁（北京航空航天大学） Stochastic alternating structure-adapted proximal gradient descent methodwith variance reduction for nonconvex nonsmooth optimization
	09:35-10:10	袁坤（北京大学） DecentLaM: Decentralized Momentum SGD for Large-batch Deep Training
	10:10-10:20	休息
	10:20-12:05	主持人：张玺（北京大学）
	10:20-10:55	陈旭瑾（中科院数学与系统科学研究院） Condorcet Stable Multiwinner Set: Optimizing Decision-making in Facility Location
	10:55-11:30	吴凌云（中科院数学与系统科学研究院） 差异网络分析与机器学习

	11:30-12:05	刘俊驿(清华大学) Nonconvex and Nondifferentiable Compound Stochastic Programming for Risk-based Statistical Learning
	12:05-13:45	午休
	13:45-15:30	主持人：吴建国(北京大学)
	13:45-14:20	尤鹏程(北京大学) Saddle Flow Dynamics: Convergence, Algorithms, and Application to Constrained Reinforcement Learning
	14:20-14:55	刘歆(中科院数学与系统科学研究院) Decentralized Optimization Over the Stiefel Manifold by an Approximate Augmented Lagrangian Function
	14:55-15:30	孔令臣(北京交通大学) Algorithmic Generalization Ability of PALM for Double Sparse Regularized Regression
	15:30-15:50	休息
	15:50-17:00	主持人：林宙辰(北京大学)
	15:50-16:25	姜海(清华大学) Joint Assortment-Price-Position Optimization Problem under the Exponomial Choice Model
	16:25-17:00	夏勇(北京航空航天大学) On Globally Solving the Nonconvex Trust Region Subproblem via the Projected Gradient Method

Talks and Abstracts

09:00-09:35, Nov. 21, Monday

报告人：戴彧虹(中科院数学与系统科学研究院)

报告题目：Optimality Conditions and Numerical Algorithms for A Class of Linearly Constrained Minimax Optimization

报告摘要：It is well known that there have been many numerical algorithms for solving nonsmooth minimax problems, numerical algorithms for nonsmooth minimax problems with joint linear constraints are very rare. This paper aims to discuss optimality conditions and develop practical numerical algorithms for minimax problems with joint linear constraints. First of all, we use the properties of proximal mapping and KKT system to establish optimality conditions. Secondly, we propose a framework of alternating coordinate algorithm for the minimax problem and analyze its convergence properties. Thirdly, we develop a proximal gradient multi-step ascent decent method (PGmsAD) as a numerical algorithm and provide the iteration complexity result for the algorithm. Finally, we apply PGmsAD to generalized absolute value equations, generalized linear projection equations and linear regression problems and report the efficiency of PGmsAD on large-scale optimization. This is a joint wok with Jiani Wang and Liwei Zhang.

09:35-10:10, Nov. 21, Monday

报告人：王彦飞(中科院地质与地球物理研究所)

报告题目：Model-driven and Data-driven Inverse Problems with AI Analysis

报告摘要：Inverse problems exist in different kinds of science and engineering. The study of inverse problems becomes an exciting area of study in recent decades.Inverse problems are usually ill-posed in the sense that three items about the solution are difficulty to be satisfied simultaneously: existence, uniqueness and stability. In addition, even a solution existed, its uncertainty still needs to be quantified. In this talk, I will address some methodologies related to inverse problem in geophysics, in particular, model-driven and data-driven inverse problems

will be investigated and artificial intelligence techniques will be discussed. Practical examples will be given.

10:20-10:55, Nov. 21, Monday

报告人：李志泽(美国卡内基梅隆大学)

报告题目：Efficient Algorithms for Optimization and Federated Learning

报告摘要：In this talk, I will first review some classical optimization algorithms/results and then introduce our contributions of designing efficient algorithms for convex and nonconvex optimization. Then I will talk about the impact of our optimization algorithms when they apply to machine learning applications, especially for communication-efficient and private federated learning. Finally, I will discuss some future research directions.

10:55-11:30, Nov. 21, Monday

报告人：唐聿劼(北京大学)

报告题目：零阶优化及其在多智能体系统优化与强化学习中的应用

报告摘要：近年来，大规模多智能体系统在各种工程应用中不断涌现，相应的分布式优化、控制与强化学习算法也得到了快速发展。为了应对大规模系统中数学模型难以建立或辨识的困难，我们将零阶优化方法与分布式计算方法相融合，提出了不依靠模型梯度信息的多智能体系统优化与强化学习方法，包括针对多智能体博弈福利优化问题的零阶反馈优化算法，以及针对去中心化线性二次型控制的分布式强化学习算法，并建立了算法的收敛性以及采样复杂度理论。报告的最后，我们将简要介绍相关的延伸工作，包括零阶优化方法逃脱鞍点的技术，以及线性二次型高斯控制的优化图景的研究。

11:30-12:05, Nov. 21, Monday

报告人：王剑晓(北京大学)

报告题目：Toward Tractable Lithiumion Battery Modeling: Multiphysics Analytics and Grid Integration

报告摘要：储能技术被广泛认为是平抑风光新能源波动、纾困电力系统多时空尺度能量平衡的核心关键。受制于电化学储能内部复杂的多物理场机理，现有电力系统运行控制难以内嵌表征电化学动态时变的约束与耐久特性，高比例新能源强烈的随机波动性已超出传统实验室环境下电池的测试激励。以锂电池为例，本报告从储能 P2D 模型的多物理场方程出发，结合仿真数据驱动建模，构造内嵌多场参量分布特性的可解析端口模型，保留精度要求的同时极大程度模型降维，进而打通底层机理与上层应用壁垒，探究该可解析模型在能源交通网络中的优化应用，最后介绍部分正在开展的延续性工作。

13:45-14:20, Nov. 21, Monday

报告人：刘亚锋(中科院数学与系统科学研究院)

报告题目：Uplink-Downlink Duality in Wireless Communications: Where Lagrange Meets Shannon

报告摘要：Many problems arising from communication system design can be formulated as optimization problems. In practice, one is often interested in not only the numerical solution to the problems but also the special structure of their optimal solution. In this talk, we shall use some examples from wireless communications and information theory to show that exploring the Lagrangian dual of these (convex) problems often reveal the structure of their optimal solution and the structure of the optimal solution will further lead to better algorithms for solving the corresponding problems.

14:20-14:55, Nov. 21, Monday

报告人：王勇(中科院数学与系统科学研究院)

报告题目：基因调控网络推断的非凸优化

报告摘要： 系统生物学（Systems biology）提出从复杂系统的观点来看待生命系统，强调以基因、蛋白质和生化反应代谢物为结点形成相互连接、具有动态特征的网络来探求和诠释各种复杂生命现象。本报告将从理工科的视角出发，以网络和数据整合为关键词，介绍我们在数据驱动的基因调控网络推断的非凸优化的建模和算法求解。

14:55-15:30, Nov. 21, Monday

报告人： 高斌(中科院数学与系统科学研究院)

报告题目： New Riemannian Preconditioned Algorithms for Tensor Completion via Polyadic Decomposition

报告摘要： We propose new Riemannian preconditioned algorithms for low-rank tensor completion via the polyadic decomposition of a tensor. These algorithms exploit a non-Euclidean metric on the product space of the factor matrices of the low-rank tensor in the polyadic decomposition form. This new metric is designed using an approximation of the diagonal blocks of the Hessian of the tensor completion cost function, thus has a preconditioning effect on these algorithms. We prove that the proposed Riemannian gradient descent algorithm globally converges to a stationary point of the tensor completion problem, with convergence rate estimates using the Lojasiewicz property. Numerical results on synthetic and real-world data suggest that the proposed algorithms are more efficient in memory and time compared to state-of-the-art algorithms. Moreover, the proposed algorithms display a greater tolerance for overestimated rank parameters in terms of the tensor recovery performance, thus enable a flexible choice of the rank parameter.

15:50-16:25, Nov.21, Monday

报告人： 韩丛英(中国科学院大学)

报告题目： 求解旅行商问题的高泛化性强化学习算法

报告摘要：在用深度学习求解组合优化问题时，组合优化实例的分布差异会显著影响深度学习求解方法的泛化性能，因此提升数据泛化性能的关键在于使模型具备处理不同分布数据的能力。基于此，我们提出了用于提升深度学习求解算法泛化性能的博弈框架，该博弈框架适用于任何组合优化问题和任何深度学习求解算法，该框架基于策略空间响应的 Oracle 构造学习算法和数据生成器之间的二人零和元博弈，经过多轮博弈后可得到一组不同分布上具有不同泛化能力的学习算法，通过模型融合的方法将各个算法进行合并，从而得到一个具有最强泛化能力的模型。在提升模型泛化能力的同时，我们发现随着模型性能的提升，该博弈框架得到的策略渐接近纳什均衡，从而体现出该博弈框架的合理性。将该框架应用于旅行商问题的求解，实验结果表明，对于同一个学习算法，在使用此框架训练时，其泛化能力有明显的提升。

16:25-17:00, Nov.21, Monday

报告人：何冠楠(北京大学)

报告题目：Time Aggregation Based on Adaptive Clustering for Low Carbon Energy System Optimization

报告摘要：Intermittent renewable energy resources like wind and solar introduce uncertainty across multiple time scales, from minutes to years, on the design and operation of power systems. Energy system optimization models have been developed to find the least-cost solution that manages the multi-timescale variability using an optimal portfolio of flexible resources. However, input data that capture such multi-time-scale uncertainty are characterized with a long time horizon and high resolution, which brings great difficulty to solving the optimization model. Here we propose a model-adaptive time aggregation method based on clustering to alleviate the computational complexity, in which the energy system is solved over selected representative time periods instead of the full time horizon. The proposed clustering method is adaptive to various energy system optimization models or settings, because it extracts features from the optimization models to inform the clustering process. Results show that the proposed adaptive method can significantly lower the error in approximating the solution of the optimization model with the full time horizon, compared to traditional time aggregation methods.

09:00-09:35, Nov. 22, Tuesday

报告人：韩德仁(北京航空航天大学)

报告题目：Stochastic Alternating Structure-Adapted Proximal Gradient Descent Method with Variance Reduction for Nonconvex Nonsmooth Optimization

报告摘要：We develop a stochastic alternating structure-adapted proximal (s-ASAP) gradient descent method for solving the block optimization problems. By deploying some state-of-the-art variance reduced gradient estimators (rather than full gradient) in stochastic optimization, the s-ASAP method is applicable to nonconvex consensus optimization problems whose objectives are the sum of a finite number of Lipschitz continuous functions. The sublinear convergence rate of s-ASAP method is built upon the proximal point theory. Furthermore, the linear convergence rate of s-ASAP method can be attainable under some mild conditions on objectives, e.g., the error bound and the Kurdyka-Lojasiewicz (KL) property. Preliminary numerical simulations on some applications in image processing demonstrate the compelling performance of the proposed method.

09:35-10:10, Nov. 22, Tuesday

报告人：袁坤(北京大学)

报告题目：DecentLaM: Decentralized Momentum SGD for Large-batch Deep Training

报告摘要：Decentralized optimization algorithms save remarkable communication overheads in distributed deep learning since each node averages locally with neighbors. While decentralized momentum SGD (DmSGD) has been widely used in real applications, it is observed to suffer severe performance degradations when taking a large batch of data samples per iteration. This talk identifies the root reason behind this phenomenon as the amplifed inconsistency bias caused by the momentum term. When batch-size grows, this bias becomes more evident and hence results in worse performances. We next propose DecentLaM, a novel decentralized large-batch momentum SGD to remove the momentum-incurred bias completely.

The convergence rates for both non-convex and strongly-convex scenarios are established. Our theoretical and empirical results justify the superiority of DecentLaM to DmSGD especially in the large-batch scenarios.

10:20-10:55, Nov. 22, Tuesday

报告人：陈旭瑾(中科院数学与系统科学研究院)

报告题目：Condorcet Stable Multiwinner Set: Optimizing Decision-making in Facility Location

报告摘要：In a facility location problem, a set of facilities is to be built to serve spatially distributed customers. But who decide(s) the solution – locations to choose for building the facilities? Different approaches place the decision in the hands of different groups or individuals. In a democratic approach, the customers collectively make the decision. In a market approach, the facilities play a game for selecting their own locations and return a Nash equilibrium. In this talk, we propose a new solution concept for the democratic approach – Condorcet stability solution, which requires that no unselected candidate location is more popular than any selected location. For the setting with customers continuously distributed on a network, we provide a characterization for Condorcet stability, leading to an efficient algorithm that finds a Condorcet stable solution when the number of facilities is large enough. We measure the efficiency of Condorcet stable solutions w.r.t. the minimum total cost of all customers, using the standard terms of Price of Anarchy and Price of Stability. Compared with the market approach, our democratic approach is shown to be more likely to achieve higher efficiency. (Joint work with Changjun Wang and Chenhao Wang)

10:55-11:30, Nov. 22, Tuesday

报告人：吴凌云(中科院数学与系统科学研究院)

报告题目：差异网络分析与机器学习

报告摘要：差异网络分析是生物信息学中一个非常重要的工具，通过分析不同状态下的生物分子网络之间的差异，可以揭示导致生物体不同状态发生发展的生物学机理和分子机制，对于理解生命规律和指导医学研究具有极其重要的意义。差异网络分析有两种典型技术路线。第一种是分别构建两个状态下的生物网络，然后再对两个网络进行比较分析。这种方法受限于网络构建方法的精确度，以及两个网络的差异比较方法。第二种则是同时推断两个状态下的生物网络及它们之间的差异。这种方法由于可以根据先验信息和生物学及医学知识，对差异网络施加各种正则化约束，因此可以更准确地推断出生物网络及其差异，已经成为目前差异网络分析研究的主要方向。本报告将简要报告差异网络分析的一些进展，以及机器学习在差异网络分析中的应用。

11:30-12:05, Nov. 22, Tuesday

报告人：刘俊驿(清华大学)

报告题目：Nonconvex and Nondifferentiable Compound Stochastic Programming for Risk-based Statistical Learning

报告摘要：Unlike the common scheme of empirical risk minimization in statistical learning, risk-based statistical learning imposes the asymmetric weights to estimation errors via appropriate risk measures to incorporate the attitudes towards errors. In the first part of the talk, we introduce a multi-class classification model employing the cost-sensitive error criteria based on buffered probability-of-exceedance and show that the corresponding parameter estimation problem is a compound stochastic program involving multiple expectations coupled by nonconvex and nondifferentiable functions. For solving such an optimization problem, we present a stochastic majorization-minimization algorithm with the almost sure convergence to the stationary solutions. In the second part of the talk, we introduce an interval CVaR-based robust statistical learning model. By representing the objective function of the parameter estimation problem as the difference of two convex value functions, we develop a stochastic difference-of-convex algorithm with the almost sure convergence to a critical solution. Numerical experiments show that the risk robust statical learning model is significantly more

robust to the contaminated data compared with the ordinary least square estimator and classical robust estimators.

13:45-14:20, Nov. 22, Tuesday

报告人：尤鹏程(北京大学)

报告题目：Saddle Flow Dynamics: Convergence, Algorithms, and Application to Constrained Reinforcement Learning

报告摘要：We study the conditions that underlie the asymptotic and exponential convergence of saddle flow dynamics of convex-concave functions to a saddle point. First, we propose a certificate rooted in observability for asymptotic convergence of saddle flows. It generalizes conventional conditions for convergence, e.g., strict convexity-concavity, and leads to a novel regularization method that is separable and makes minimal requirements on convexity-concavity for asymptotic convergence. Second, we show that saddle flows' global exponential stability is a direct consequence of strong convexity-concavity, which provides a lower-bound estimate of the convergence rate. This insight explains some existing algorithms' convergence properties for equality constrained convex optimization, e.g., proximal gradient. It is further exploited to design a novel alternative regularized algorithm that achieves exponential convergence at a rate depending only on the strong convexity of the objective. Our results generalize to saddle flow dynamics with projections on the vector field and have immediate applications in inequality constrained convex optimization, in particular, distributed linear programs and constrained reinforcement learning.

14:20-14:55, Nov. 22, Tuesday

报告人：刘歆(中科院数学与系统科学研究院)

报告题目：Decentralized Optimization Over the Stiefel Manifold by an Approximate Augmented Lagrangian Function

报告摘要：We study the decentralized optimization problem over the Stiefel manifold, which is defined on a connected network of d agents. The objective is an average of d local functions, and each function is privately held by an agent and encodes its data. The agents can only communicate with their neighbors in a collaborative effort to solve this problem. In existing methods, multiple rounds of communications are required to guarantee the convergence, giving rise to high communication costs. In contrast, this paper proposes a decentralized algorithm, called DESTINY, which only invokes a single round of communications per iteration. DESTINY combines gradient tracking techniques with a novel approximate augmented Lagrangian function. The global convergence to stationary points is rigorously established. Comprehensive numerical experiments demonstrate that DESTINY has a strong potential to deliver a cutting-edge performance in solving a variety of testing problems.

14:55-15:30, Nov. 22, Tuesday

报告人：孔令臣(北京交通大学)

报告题目：Algorithmic Generalization Ability of PALM for Double Sparse Regularized Regression

报告摘要：While the traditional statistical methodology is usually assumed that the training samples are drawn from a homogeneous population, practice belies this statement. The population heterogeneity may lead to bias and misleading conclusions as well as heavily degrade the performance of learning algorithms, which motivates the development of the new methodology to counter the heterogeneity effect becoming a fundamental task. The key challenge is that we usually do not have prior knowledge of the heterogeneity and its source. To address this issue, we propose a novel double sparse regularized estimation procedure for the Generalized Linear Model, in which the heterogeneity can be automatically identified and simultaneously estimated by the heterogeneous-specific intercept. The developed algorithm in our estimation procedure combines the linearization technique and Alternating Minimization and allows for possible non-convex regularization terms. Specifically, we study the algorithmic generalization ability, i.e., approximation error and asymptotic consistency with respect to

iterative sequences and indicate its convergence with high probability. The power of predictive validation is verified by several simulations and real data applications. Experimental results show superior performance of our method in comparison with several state-of-the-art methods.

15:50-16:25, Nov. 22, Tuesday

报告人：姜海(清华大学)

报告题目：Joint Assortment-Price-Position Optimization Problem under the Exponomial Choice Model

报告摘要：We study the joint assortment-price-position optimization problem under the Exponomial Choice Model (ECM). The goal is to determine the revenue-maximizing subset of products with their corresponding selling prices and display positions. We formulate this problem as a non-linear mixed integer program. We identify structural properties of the optimal solution and develop an exact decomposition-based algorithm to obtain the optimal solution.

16:25-17:00, Nov. 22, Tuesday

报告人：夏勇(北京航空航天大学)

报告题目：On Globally Solving the Nonconvex Trust Region Subproblem via the Projected Gradient Method

报告摘要：The trust region subproblem (TRS) is to minimize a possibly nonconvex quadratic function over a Euclidean ball. There are typically two classes for (TRS), the so-called ``easy'' and ``hard'' cases. It may occur even in the ``easy case'' that the sequence generated by the projected gradient method (PG) starting from any initial point in a nonzero measure feasible set converges locally sublinearly to a saddle point. To our surprise, when applying (PG) to solve a cheap and possibly nonconvex reformulation of (TRS), the generated sequence initialized with a uniformly and randomly generated feasible point converges to the global minimizer of (TRS) with probability one. The local convergence rate is at least linear for the ``easy case'', without

assuming that we have to possess the information that the ``easy case'' occurs. We also consider how to use (PG) to globally solve equality-constrained (TRS).