Preface

Welcome to Workshop on Optimization, Equilibrium and Complementarity, organized by the Department of Applied Mathematics of The Hong Kong Polytechnic University.

There have been very rapid developments in optimization theory and algorithms in data science, engineering and Artificial Intelligence. This workshop is to celebrate a 70 + 50 + 45 milestone and foster the sustained development in optimization, equilibrium and complementarity. It will encourage international collaboration and provide an opportunity for young researchers to learn the current state of the art in optimization, equilibrium and complementarity. We would like to take this opportunity to thank all the speakers and participants for their support to the workshop.

We are grateful to the members of the organizing committee for their hard work. We are also grateful to the staff members of the Department of Applied Mathematics for their time and effort in making this workshop possible.

The workshop gratefully acknowledges the patronage of The Hong Kong Polytechnic University, and the financial support of CAS AMSS-PolyU Joint Laboratory of Applied Mathematics (JLab).

Organizing Committee

Acknowledgments

We are very grateful for the support and contributions to the workshop given by

The Hong Kong Polytechnic University (PolyU)

Faculty of Science, PolyU

Department of Applied Mathematics, PolyU

CAS AMSS-PolyU Joint Laboratory of Applied Mathematics (JLab)

We would also like to thank the staff members in the Department of Applied Mathematics who have worked together in handling the various arrangements related to the workshop. Their assistance has been indispensable.

Organizing Committees

Co-chairs

- Xiaojun Chen (PolyU)
- Defeng Sun (PolyU)

Members

- Kuang Bai (PolyU)
- Xiaoxia Liu (PolyU)
- Ting-kei Pong (PolyU)
- Houduo Qi (PolyU)
- Xindong Tang (PolyU)
- Xiaoqi Yang (PolyU)
- Yancheng Yuan (PolyU)
- Zaikun Zhang (PolyU)

Workshop Highlights

Workshop Venue

The workshop will be held on the PolyU campus. Please refer to the campus map. Please note that smoking is not allowed on the PolyU campus.

Opening Ceremony

The opening ceremony will be held in V322 at 9: 00 in the Thursday morning of 17 August 2023.

Registration & Reception & Information Desk

Registration desk will be set up outside V322 from 17:00-19:00 on 16 August 2023. A notice board will be available nearby the registration desk. Updated information and announcement about the workshop will be posted on the notice board.

On-campus Internet Access

Desktop computers are available for internet access in room M302. The helper can take you to M302.

Wi-Fi access is also provided from 16–19 August 2023. Please use the login name and password listed below.

User ID: gstoec
Password: August2023

Remarks: Only SSID: PolyUWLAN can be used.

Registered participants are entitled to:

- Attendance to all workshop sessions
- Invitation to the welcome reception
- Invitation to the opening ceremony
- Coffee breaks
- A free copy of the Workshop Abstracts and Workshop Programme

Lunch

Lunch tickets are attached to your official workshop name tag. Be sure to bring along your tickets when you attend the lunch. In addition, there are a number of cafeterias available on campus. Please refer to the on-campus cafeteria map.

The nearby Tsim Sha Tsui East area, which is within walking distance from PolyU, provides a variety of teahouse, restaurants and fast food services. Please refer to the map of TST East.

Banquet

The Workshop Banquet will be held on 18 August 2023 at 18:30 at SJ Cuisine (港晶軒), 2/F Harbour Crystal Centre, Tsim Sha Tsui East. A banquet ticket is attached to your official workshop name tag. Be sure to bring along your banquet ticket when you attend the banquet.

Group Photo

Workshop Group Photo will be taken on 17 August 2023 after the Opening Ceremony.

Schedule of Workshop at Room V322

	17 August (Thursday)	18 August (Friday)	19 August (Saturday)
9:00-9:10	**Opening Ceremony** Group Photo		
	Invited talk **Chair:** Defeng Sun	**Invited talk** **Chair:** Xiaojun Chen	**Invited talk** **Chair:** Kim-Chuan Toh
9:10-9:50	Terry Rockafellar	Yinyu Ye	Marc Teboulle
9:50-10:30	Suvrajeet Sen	Kim-Chuan Toh	Shuzhong Zhang
10:30-10:50	Coffee Break [Outside V322]	Coffee Break [Outside V322]	Coffee Break [Outside V322]
	Invited talk **Chair:** Jong-Shi Pang	**Invited talk** **Chair:** Daniel Ralph	**Invited talk** **Chair:** Jie Sun
10:50-11:30	Michael C. Ferris	Jong-Shi Pang	Meisam Razaviyayn
11:30-12:10	Uday V. Shanbhag	Jinglai Shen	Michael Hintermüller
12:10	Lunch	Lunch	Lunch
	Invited talk **Chair:** Uday V. Shanbhag	**Invited talk** **Chair:** Shuzhong Zhang	**Invited talk** **Chair:** Jinglai Shen
14:00-14:40	Jie Sun	Daniel Ralph	Junyi Liu
14:40-15:20	Guoyin Li	Tom Luo	Miju Ahn
15:20-15:40	Coffee Break [Outside V322]	Coffee Break [Outside V322]	Coffee Break [Outside V322]
	Invited talk **Chair:** Guoyin Li	**Invited talk** **Chair:** Zaikun Zhang	**Invited talk** **Chair:** Meisam Razaviyayn
15:40-16:20	Xinyuan Zhao	Wei Bian	Xudong Li
16:20-17:00	Chao Zhang	Chao Ding	Hailin Sun
17:00-17:40	Xiaoqi Yang	Houduo Qi	Ting-kei Pong
17:40-17:50			**Closing Ceremony**
18:30		**Banquet**	

Invited Talks

---

17 August 2023, Thursday

Chair: Defeng Sun **[V322]**

09:10-09:50 **Terry Rockafellar** (University of Washington, USA)
Progressive decoupling and the proximal method of multipliers

09:50-10:30 **Suvrajeet Sen** (University of Southern California, USA)
Compromise policies: Extensions to multi-stage and discrete stochastic optimization

Chair: Jong-Shi Pang **[V322]**

10:50-11:30 **Michael C. Ferris** (University of Wisconsin-Madison, USA)
From complementarity to risk-averse stochastic equilibria: models and algorithms

11:30-12:10 **Uday V. Shanbhag** (Pennsylvania State University, USA)
On the tractable resolution of chance-constrained optimization Problems

Chair: Uday V. Shanbhag **[V322]**

14:00-14:40 **Jie Sun** (Curtin University, Australia & National University of Singapore, Singapore)
Progressive decoupling on conic structure

14:40-15:20 **Guoyin Li** (University of New South Wales, Australia)
Proximal methods for nonsmooth and nonconvex fractional programs: when sparse optimization meets fractional programs

Chair: Guoyin Li **[V322]**

15:40-16:20 **Xinyuan Zhao** (Beijing University of Technology, China)
An accelerated proximal alternating direction method of multipliers for optimal decentralized control of uncertain systems

16:20-17:00 **Chao Zhang** (Beijing Jiaotong University, China)
Stochastic smoothing accelerated gradient method for nonsmooth convex stochastic composite optimization

17:00-17:40 **Xiaoqi Yang** (The Hong Kong Polytechnic University, China)
A globally convergent regularized Newton method for l_q-norm composite optimization problems

Invited Talks

---

18 August 2023, Friday

Chair: Xiaojun Chen **[V322]**

09:10-09:50 **Yinyu Ye** (Stanford University, USA)
An alternative to the Newton step: homogenized second-order descent direction

09:50-10:30 **Kim-Chuan Toh** (National University of Singapore, Singapore)
A feasible method for linearly constrained convex SDP problems

Chair: Daniel Ralph **[V322]**

10:50-11:30 **Jong-Shi Pang** (University of Southern California, USA)
A happy home coming after a fruitful "non"-journey

11:30-12:10 **Jinglai Shen** (University of Maryland(Baltimore County), USA)
Dynamic stochastic variational inequality and its computation

Chair: Shuzhong Zhang **[V322]**

14:00-14:40 **Daniel Ralph** (University of Cambridge, United Kingdom)
POST as an all-purpose algorithm for solving penalised stopping problems: from function spaces to finite dimensions

14:40-15:20 **Tom Luo** (The Chinese University of Hong Kong (Shenzhen), China)
Bridging distributional and risk-sensitive reinforcement learning with provable regret bounds

Chair: Zaikun Zhang **[V322]**

15:40-16:20 **Wei Bian** (Harbin Institute of Technology, China)
Local saddle points of convex-concave minimax problems with cardinality penalty

16:20-17:00 **Chao Ding** (Chinese Academy of Sciences, China)
Nonsmooth optimization over Riemannian manifolds

17:00-17:40 **Houduo Qi** (The Hong Kong Polytechnic University, China)
Sparse SVM with hard-margin loss: Newton-augmented Lagrangian method in reduced dimensions

Invited Talks

---

19 August 2023, Saturday

Chair: Kim-Chuan Toh **[V322]**

09:10-09:50 **Marc Teboulle** (Tel Aviv University, Israel)
Lagrangian methods for nonsmooth nonconvex optimization

09:50-10:30 **Shuzhong Zhang** (University of Minnesota, USA)
New algorithms for solving non-monotone variational inequality Problems

Chair: Jie Sun **[V322]**

10:50-11:30 **Meisam Razaviyayn** (University of Southern California, USA)
Tradeoffs between convergence rate and noise amplification of first-order methods

11:30-12:10 **Michael Hintermüller** (Weierstrass Institute for Applied Analysis and Stochastics & Humboldt–Universität zu Berlin, Germany)
Optimal Control of (Quasi)Variational Inequalities: Stationarity, Risk-Aversion, and Numerical Solution

Chair: Jinglai Shen **[V322]**

14:00-14:40 **Junyi Liu** (Tsinghua University, China)
Nonconvex and Nonsmooth Approaches for Affine Chance-Constrained Stochastic Programs

14:40-15:20 **Miju Ahn** (Southern Methodist University, USA)
Tractable continuous approximations for a constraint selection Problem

Chair: Meisam Razaviyayn **[V322]**

15:40-16:20 **Xudong Li** (Fudan University, China)
Data-driven minimax optimization with expectation constraints

16:20-17:00 **Hailin Sun** (Nanjing Normal University, China)
Dynamic stochastic projection method for multistage stochastic variational inequalities

17:00-17:40 **Ting-kei Pong** (The Hong Kong Polytechnic University, China)
Frank-Wolfe-type methods for nonconvex inequality-constrained problems

Progressive decoupling and the proximal method of multipliers

R. Tyrrell Rockafellar

University of Washington

Abstract. The proximal method of multipliers in convex optimization applies the proximal point algorithm to find a saddle point of the convex-concave Lagrangian function associated with the problem. An extension to nonconvex optimization is possible through the local convexity-concavity of the augmented Lagrangian that comes from a variational sufficient condition for local optimality.

Applications of the progressive decoupling algorithm in optimization can be derived as a specialization. A new variable-metric version of the proximal point algorithm with generic linear convergence has a key role in supporting error analysis and understanding how the convergence rate depends on proximal parameters.

Compromise policies: Extensions to multi-stage and discrete stochastic optimization

Suvrajeet Sen

University of Southern California

Abstract. Compromise Decisions for Stochastic Programming were first proposed for Two-stage Stochastic LPs (SLP), and have been shown to provide very good decisions even in problems with very high variability. In this lecture we will discuss extensions of Compromise Decisions of SLP to the case of Multi-stage Stochastic LPs as well as Stochastic MIPs (with binary variables). We will show that despite the differences in the mathematical structures arising in continuous and discrete optimization problems, the same notion of compromise policies provides a common umbrella for variance reduced decisions (or policies). This talk is based on joint work with several former Ph.D. students.

From complementarity to risk-averse stochastic equilibria: models and algorithms

Michael C. Ferris

University of Wisconsin-Madison

Abstract. We present a mechanism for describing and solving collections of optimization problems that are linked by equilibrium conditions. We show how to incorporate stochastic information into these systems and give examples of their use and possible extensions to hierarchical modeling. We outline several algorithms and investigate their computational efficiency.

Stochastic equilibria can be used to model many data driven applications, including many dynamic models of competition. In this model, players solve risk averse optimization problems that are coupled by a shared scenario tree. Players are linked by equilibrium conditions at nodes of that tree. Applications to energy planning and their interactions with environmental concerns will be outlined.

On the tractable resolution of chance-constrained optimization problems

Uday V. Shanbhag

Pennsylvania State University

Abstract. Chance-constrained optimization represents a challenging class of problems in stochastic programming. Such models assume relevance in the modeling of reliability and robustness and have found applicability in energy systems, service systems, communication networks, amongst others. First studied by Charnes and Cooper in the 1950s, this problem class has seen significant recent study by both continuous and discrete optimizers. Important recent results include the development of schemes that have combined smoothing, sampling, and variational analysis. However, tractable algorithms, equipped with rate and complexity guarantees, for computing approximate global minimizers have remained elusive. We consider the resolution of precisely such a gap for a subclass of problems. We first examine the setting of probability maximization where under suitable distributional assumptions, by leveraging properties of Minkowski functionals, the probability of interest can be expressed as the expectation of a Clarke-regular function with respect to an appropriately defined Gaussian density (or its variant). In fact, we may show that a suitably defined equivalent compositional representation is shown to be convex. We then derive rate and complexity guarantees for a regularized variance-reduced framework for computing approximate global maximizers of the original problem. We extend our framework to a chance-constrained regime where distributional assumptions are weakened, where a global minimizer of this chance-constrained problem is equivalently obtained by solving a stochastic inclusion with a compositional monotone operator. Complexity guarantees for resolving this inclusion are provided for an inexact variance-reduced proximal-point scheme.

Progressive decoupling on conic structure

Jie Sun

National University of Singapore and Curtin University

Abstract. This talk focuses on the extension of the progressive decoupling algorithm (PDA) of Rockafellar to stochastic optimization problems with conic structures such as the stochastic second order cone programs and stochastic semidefinite programs. It is shown that a generalized Douglass-Rachford method and the Halpern-Peaceman-Rachford method can be used to solve these problems. Preliminary Numerical results are reported.

This is a joint work with Min Zhang from the Chinese Academy of Sciences.

Proximal methods for nonsmooth and nonconvex fractional programs: when sparse optimization meets fractional programs

Guoyin Li

University of New South Wales

Abstract. Nonsmooth and nonconvex fractional programs are ubiquitous and also highly challenging. It includes the composite optimization problems studied extensively lately, and encompasses many important modern optimization problems arising from diverse areas such as the recent proposed scale invariant sparse signal reconstruction problem in signal processing, the robust Sharpe ratio optimization problems in finance and the sparse generalized eigenvalue problem in discrimination analysis. In this talk, we will introduce extrapolated proximal methods for solving nonsmooth and nonconvex fractional programs and analyse their convergence behaviour. Interestingly, we will show that the proposed algorithm exhibits linear convergence for sparse generalized eigenvalue problem with either cardinality regularization or sparsity constraints. This is achieved by identifying the explicit desingularization function of the Kurdyka-{\L}ojasiewicz inequality for the merit function of the fractional optimization models. Finally, if time permits, we will present some preliminary encouraging numerical results for the proposed methods for sparse signal reconstruction and sparse Fisher discriminant analysis.

An accelerated proximal alternating direction method of multipliers for optimal decentralized control of uncertain systems

Xinyuan Zhao

Beijing University of Technology

Abstract. To ensure the system stability of the H2-guaranteed cost optimal decentralized control problem (ODC), an approximate semidefinite programming (SDP) problem is formulated based on the sparsity of the gain matrix of the decentralized controller. To reduce data storage and improve computational efficiency, the SDP problem is vectorized into a conic programming (CP) problem using the Kronecker product. Then, a proximal alternating direction method of multipliers (PADMM) is proposed to solve the dual of the resulted CP. By linking the (generalized) PADMM with the (relaxed) proximal point algorithm, we are able to accelerate the proposed PADMM via the Halpern fixed-point iterative scheme. This results in a fast convergence rate for the Karush-Kuhn-Tucker (KKT) residual along the sequence generated by the accelerated algorithm. Numerical experiments further demonstrate that the accelerated PADMM outperforms both the well-known CVXOPT and SCS algorithms for solving the large-scale CP problems arising from H2-guaranteed cost ODC problems.

Stochastic smoothing accelerated gradient method for nonsmooth convex stochastic composite optimization

Chao Zhang

Beijing Jiaotong University

Abstract. We propose a novel stochastic smoothing accelerated gradient (SSAG) method for general constrained nonsmooth stochastic composite convex optimization, and analyze the convergence rates. The SSAG method allows various smoothing techniques, and can deal with the nonsmooth term that is not easy to compute its proximal term, or that does not own the linear max structure. To the best of our knowledge, it is the first time to propose a stochastic smoothing accelerated first-order method to solve the maximization of numerous nonlinear convex functions. We prove that the SSAG method achieves the best-known complexity bounds in terms of the stochastic first-order oracle (SFO), using either diminishing smoothing parameters or a fixed smoothing parameter. We give two applications of our results to distributionally robust optimization problems, and provide numerical results that demonstrate the effectiveness and efficiency of the proposed SSAG method.

This is a joint work with Ruyu Wang.

A globally convergent regularized Newton method for l_q-norm composite optimization problems

Xiaoqi Yang

The Hong Kong Polytechnic University

Abstract. In this talk we will discuss the composite optimization problem with a twice continuously differentiable loss function and an l_q-norm regularized term. For this class of nonconvex and nonsmooth composite problems, we will review some existing first-order and second-order algorithms. In this work, we propose a hybrid of proximal gradient method and subspace regularized Newton method. The whole iterate sequence produced by the algorithm is proved to have a finite length and converge to an L-type stationary point under a mild curve-ratio condition and the Kurdyka-{\L}ojasiewicz property of the cost function, which converges linearly if further a Kurdyka-{\L}ojasiewicz property of exponent 1/2 holds. Moreover, a superlinear convergence rate for the iterate sequence is also achieved under an additional local error bound condition. Our convergence results do not require the isolatedness and strict local minimality properties of the L-stationary point. Numerical comparisons with some existing algorithms for the l_q--norm regularized linear and logistic regressions on real data indicate that our algorithm not only requires much less computing time but also yields comparable even better sparsities and objective function values.

An alternative to the Newton step: homogenized second-order descent direction

Yinyu Ye

Stanford University

Abstract. Most second-order optimization methods are based on the Newton step to solve a system of linear equations in each iteration which costs $O(n^3)$ operations. We show how this step can be replaced by finding the eigenvector corresponding to the minimum eigenvalue of a homogenized symmetric matrix that integrates the Hessian and Gradient. The extreme eigenvector can be computed in general by fast (randomized) iterative method with matrix-vector product operations. We show a few theoretical results on how this step can be placed for several popular second-order convex and nonconvex optimization methods, including the trust-region method, the cubic-regularization method, and the homotopic or path-following Newton framework. We further demonstrate its practical performances from various applications in Machine/Deep Learning and Data/Information Sciences.

A feasible method for linearly constrained convex SDP problems

Kim-Chuan Toh

National University of Singapore

Abstract. In this work, we propose a low rank decomposition (SDPLRE) of a linearly constrained convex SDP problem that uses a squared slack variable to transfer inequality constraints into equality constraints. We prove that a random perturbation to appropriate components of the constant vector of the constraints in (SDPLRE) can eliminate all singular points with probability 1. We also propose a rank-adaptive feasible method to solve a general linearly constrained convex SDP problem with convergence guarantee. Compared with previous work, our method does not require the objective function to be twice differentiable and achieves a better iteration complexity. Numerical experiments are conducted to verify the high efficiency and robustness of our method.

A happy home coming after a fruitful "non"-journey

Jong-Shi Pang

University of Southern California

Abstract. The home in the title refers to the origin of optimization, i.e., linear programming (LP) and its extensions to (convex) quadratic programming (QP) and the linear complementarity problem (LCP), the latter being the virgin focus of the speaker's research career. The "non"-journey refers to the speaker's recent exploration in the area of optimization with "non"-features. After wandering many years in this journey, the speaker returns to the basics of the field and will talk about some new, and some not so new results about the solution of some classes of LPs, convex QPs, and LCPs by simplex-type pivoting methods, a topic that is (almost) forgotten amid today's IT hype of ML, AI, chatGPT, etc. The results pertain to bounds on the number of pivots, some of which are linear in the number of variables and thus render the problems strongly polynomially solvable.

A Chinese poem composed by the speaker and its English translation to describe the talk follow:

浪子回头，十年不晚，逆波漂逐，艰苦奋斗，
開疆拓土，建立根基， 受益下代， 造福将来。

The journeyer finds his way home,

Years in the wild, but not too late to rewind,

Defying the currents, against the tides, adrift and never tired,

Enduring hardships, fueled by tenacity,

Creating new paths, building foundation, a legacy born,

Benefitting the youth, future adorned.

Dynamic stochastic variational inequality and its computation

Jinglai Shen

University of Maryland Baltimore County

Abstract. In this talk, we introduce the dynamic stochastic variational inequality (DSVI). The DSVI is an ODE whose right hand side is defined by the natural mapping of a deterministic VI, referred to as the first-stage VI, which is coupled with a stochastic VI, referred to as the second-stage SVI. The DSVI provides a unified modeling framework for various applications involving equilibrium/optimality conditions (VI), dynamics (ODE), and uncertainties (stochasticity). We establish solution existence and uniqueness for two classes of DSVIs: the first class is defined by a strongly monotone SVI in the second stage, and the second class pertains to a box-constrained stochastic linear VI with the P-property in the second stage. Certain properties on switching dynamics of the DSVI are discussed. We then develop sample average approximation (SAA) and time-stepping schemes to compute the DSVI. The uniform/exponential convergence is established for the SAA under suitable conditions. A time-stepping EDIIS (energy direct inversion on the iterative subspace) method is proposed to solve the differential VI arising from the SAA of the DSVI. These results are illustrated by an instantaneous dynamic user equilibrium problem in transportation engineering.

This is a joint work with Xiaojun Chen from the Hong Kong Polytechnic University.

POST as an all-purpose algorithm for solving penalised stopping problems: from function spaces to finite dimensions

Daniel Ralph

University of Cambridge

Abstract. Introduction. Stopping problems involve following a stochastic process with the option to stop, i.e., to take (a function of) the value at your current state as your exit reward at a time of your choosing. The optimal value is a function on the state space, namely the most you can receive on expectation by judiciously stopping. Optimal stopping is ubiquitous for framing decisions under uncertainty in finance (e.g., Black-Scholes-Merton for the American option), economics and the interface between operations and finance (e.g., trading commodities which can be stored).

Function spaces. We consider the well-known penalty version of optimal stopping by observing it is equivalent to a new stopping problem with a restriction on stopping times: you can only stop at one of a sequence of random Poisson arrivals whose intensity is set to be the penalty parameter. We show that the associated Poisson Optional Stopping Time (POST) algorithm provides a dichotomy for all stopping problems, including those whose processes and data functions are unbounded: POST iterates - functions on the state space - either converge pointwise to an optimal solution of the penalty problem, or diverge monotonically to infinity at some point of the state space (and the penalty problem has no solution). This dramatically widens the scope of analysis of optimal stopping compared to the traditional topological approach which has a supremum norm (bounded data) argument at the core.

Finite dimensions. For computation, it is well known that discretising the penalty problem is equivalent to forming a linear complementarity problem (LCP). We show that the discretised POST method is equivalent to a well-known splitting method for LCPs (see Cottle-Pang-Stone's LCP book) with a twist: We start at the origin, $x = 0$, and the vector increases monotonically (componentwise) at each iteration, whereas the original method starts at a feasible point, x, and decreases monotonically.

This is a joint work with Rutger-Jan Lange and Jean-Claude Hessing from the Econometric Institute, Erasmus School of Economics, Rotterdam, Netherlands, and Jan van Casteren from the Department of Mathematics, University of Antwerp, Belgium.

Bridging distributional and risk-sensitive reinforcement learning with provable regret bounds

Tom Luo

The Chinese University of Hong Kong, Shenzhen

Abstract. Risk-sensitive decision-making is crucial in high-stakes applications, such as finance, medical treatment, and operations research. In these scenarios, decision-makers aim to maximize a risk measure of the return distribution, rather than simply maximizing the expected return, a.k.a. Risk-sensitive reinforcement learning (RSRL). As the first risk measure applied to RSRL, the exponential risk measure has been successfully applied in various domains. However, current approaches involve complicated algorithmic designs and regret analysis to handle the non-linearity of the exponential risk measure.

To tackle these challenges, we propose a novel distributional reinforcement learning (DRL) algorithms for RSRL with regret guarantee. Our algorithm does not involve complicated bonus design to guide exploration, and enjoy a simpler and more interpretable regret analysis. We build a risk-sensitive distributional dynamic programming. Furthermore, we provide a regret upper bound of the algorithm via distributional optimism. We also fix and tighten the previous minimax lower bound.

Local saddle points of convex-concave minimax problems with cardinality penalty

Wei Bian

Harbin Institute of Technology

Abstract. In this paper, we focus on a class of convexly constrained convex-concave minimax problems with cardinality functions. Although such nonconvex-nonconcave and discontinuous problems may not have a saddle point, but they have local saddle points and global minimax points. We show that some local saddle points have the lower bound properties for stability of variable selection and we define a class of strong local saddle points based on it. Moreover, we give a framework to construct continuous relaxations of the discontinuous minimax problems based on the convolution, such that they have the same strong local saddle points with the original problems. We also establish the relations between the relaxation problems and the original problems regarding local saddle points, global minimax points, local minimax points and stationary points. Finally, we apply our results to the sparse robust bond portfolio construction problems.

This is a joint work with Xiaojun Chen.

Nonsmooth optimization over Riemannian manifolds

Chao Ding

Academy of Mathematics and Systems Science,

Chinese Academy of Sciences

Abstract. Nonsmooth optimization over Riemannian manifolds has become a growing field of interest and research in recent years because of its applications in various domains, including machine learning and computer vision, among others. This approach considers the optimization of non-smooth functions defined on curved surfaces or manifolds, which is different from its Euclidean counterpart and presents a unique set of theoretical and practical challenges. The theoretical study of nonsmooth optimization over Riemannian manifolds is different from that of Euclidean optimization because of the geometric nature of Riemannian manifolds, and as a result, few results are available. In this talk, we introduce a local Riemannian nonsmooth analysis framework, which provides a theoretical foundation for nonsmooth optimization over Riemannian manifolds. Moreover, we present recent progress on the theoretical study of those Riemannian nonsmooth optimization problems, which holds great promise for advancing our understanding of optimization and related topics in diverse fields.

Sparse SVM with hard-margin loss: Newton-augmented Lagrangian method in reduced dimensions

Houduo Qi

The Hong Kong Polytechnic University

Abstract. The hard margin loss function has been at the core of the support vector machine (SVM) research from the very beginning due to its generalization capability. On the other hand, the cardinality constraint has been widely used for feature selection, leading to sparse solutions. This paper studies the sparse SVM with the hard-margin loss that integrates the virtues of both worlds. However, SSVM-HM is one of the most challenging models to solve. In this paper, we cast the problem as a composite optimization with the cardinality constraint. We characterize its local minimizers in terms of P-stationarity that well captures the combinatorial structure of the problem. We then propose an inexact proximal augmented Lagrangian method (iPAL). The different parts of the inexactness measurements from the P-stationarity are controlled on different scales in a way that the generated sequence converges both globally and at a linear rate. This matches the best convergence theory for composite optimization. To make iPAL practically efficient, we propose a gradient-Newton method in a subspace for the iPAL subproblem. This is accomplished by detecting active samples and features with the help of the proximal operator of the hard margin loss and the projection of cardinality constraint. Extensive numerical results on both simulated and real datasets demonstrate that the proposed method is fast, produces sparse solution of high accuracy, and can lead to effective reduction on active samples and features when compared with several leading solvers.

This is a joint work with Naihua Xiu and Penghe Zhang.

Lagrangian methods for nonsmooth nonconvex optimization

Marc Teboulle

Tel Aviv University

Abstract. Nonconvex and nonsmooth optimization models are prevalent in modern applications, yet they are very hard to solve and pose serious theoretical and computational challenges. This talk addresses some of the theoretical challenges by focusing on the design and analysis of adaptive Lagrangian based methods for a comprehensive class of nonsmooth nonconvex models with nonlinear functional composite structures. We discuss the main obstacles, highlighting the ways they can be avoided and the pillars of the convergence theory.

New algorithms for solving non-monotone variational inequality problems

Shuzhong Zhang

University of Minnesota

Abstract. In this talk we present our recent results on solving some classes of constrained non-monotone variational inequality problems. We shall present two types of algorithms: (1) high-order approximation-based extra-gradient type methods; (2) augmented Lagrangian-based primal-dual methods. Attention is then paid on the conditions under which the algorithms are shown to be convergent. This talk is based on the author's joint work with Lei Zhao and Daoli Zhu, and with Kevin Huang.

Tradeoffs between convergence rate and noise amplification of first-order methods

Meisam Razaviyayn

University of Southern California

Abstract. We study momentum-based first-order optimization algorithms in which the iterations utilize information from the two previous steps and are subject to additive white noise. This class of algorithms includes Polyak's heavy-ball and Nesterov's accelerated methods as special cases and noise accounts for uncertainty in either gradient evaluation or iteration updates. For strongly convex problems, we use the steady-state variance of the error in the optimization variable to quantify noise amplification and identify fundamental stochastic performance tradeoffs. Our approach utilizes the Jury stability criterion to provide a novel geometric characterization of conditions for linear convergence, and it clarifies the relation between the noise amplification and convergence rate as well as their dependence on the condition number and the constant algorithmic parameters. This geometric insight leads to simple alternative proofs of standard convergence results and allows us to establish analytical lower bounds on the product between the settling time and noise amplification that scale quadratically with the condition number. Our analysis also identifies a key difference between the gradient and iterate noise models: while the amplification of gradient noise can be made arbitrarily small by sufficiently decelerating the algorithm, the best achievable variance amplification for the iterate noise model increases linearly with the settling time in decelerating regime. Furthermore, we introduce two parameterized families of algorithms that strike a balance between noise amplification and settling time while preserving order-wise Pareto optimality for both noise models. Finally, by analyzing a class of accelerated gradient flow dynamics, whose suitable discretization yields the two-step momentum algorithm, we establish that stochastic performance tradeoffs also extend to continuous time.

Optimal Control of (Quasi)Variational Inequalities: Stationarity, Risk-Aversion, and Numerical Solution

Michael Hintermüller

Humboldt-Universität zu Berlin

Abstract. Motivated by various applications such as superconductivity, thermoforming, electrostatics or impulse control, in this talk elliptic quasi-variational inequality (QVI) problems are discussed. Due to the QVI constraint this class of optimization problems is non-smooth and non-convex. Moreover, the constraints are typically degenerate, thus challenging analytical as well as numerical investigations. Concerning the state system, i.e., the QVI, and given a control, the existence of minimal and maximal solutions is argued and the directional differentiability of these solutions with respect to the control is shown. The latter is useful for algorithmic purposes, but also for characterizing stationary points of the control problem. Indeed, such stationarity conditions are derived via a limit process involving regularized control problems. The latter represents a class of PDE constrained optimization problems with a non-linear state system (approximating the QVI) which, in contrast to the QVI constraint, is regular. As a consequence, KKT conditions in Banach space can be employed.

Moreover, as problem data is often uncertain, possibly following a specific distribution, also risk-averse formulations for control of VI problems are studied and, again, stationarity conditions are derived. Here the VI can be derived from the QVI by fixing the constraint set for the QVI at a reference point, thus leading to an elliptic VI problem. The renowned obstacle problem is a particular example for the latter. It will be our focus for the further development. Concerning the underlying risk-measure we are particularly interested in CVaR which gives rise to a non-smooth objective. Finally, numerical tests conclude the talk.

Nonconvex and Nonsmooth Approaches for Affine Chance-Constrained Stochastic Programs

Junyi Liu

Tsinghua University

Abstract. Chance-constrained programs (CCPs) constitute a difficult class of stochastic programs due to its possible nondifferentiability and nonconvexity even with simple linear random functionals. Existing approaches for solving the CCPs mainly deal with convex random functionals within the probability function. In the present talk, we consider two generalizations of the class of chance constraints commonly studied in the literature; one generalization involves probabilities of disjunctive nonconvex functional events and the other generalization involves mixed-signed affine combinations of the resulting probabilities; together, we coin the term affine chance constraint (ACC) system for these generalized chance constraints. Our proposed treatment of such an ACC system involves the fusion of several individually known ideas: (a) parameterized upper and lower approximations of the indicator function in the expectation formulation of probability; (b) external (i.e., fixed) versus internal (i.e., sequential) sampling-based approximation of the expectation operator; (c) constraint penalization as relaxations of feasibility; and (d) convexification of nonconvexity and nondifferentiability via surrogation. We will discuss several algorithmic strategies based on the integration of these techniques with various degrees of practicality and computational efforts for solving the nonconvex ACC-SP.

Tractable continuous approximations for a constraint selection problem

Miju Ahn

Southern Methodist University

Abstract. This presentation introduces a constraint selection problem where the decision-maker solves an optimization problem with a set of constraints that are preferred to be satisfied. We formulate the problem as a cardinality minimization problem (CMP) that penalizes the number of unsatisfied such soft constraints using an indicator function. Our approach reformulates the discrete CMP as continuous problems. We present an equivalent formulation of a mathematical program with complementarity constraints and an approximation as a difference-of-convex program. The stationary solutions of the alternative formulations are investigated, emphasizing the recovery of the local solutions of the CMP. Our numerical study results demonstrate our method's effectiveness in enforcing desired conditions on several suitable applications.

Data-driven minimax optimization with expectation constraints

Xudong Li

Fudan University

Abstract. Attention to data-driven optimization approaches has grown significantly over recent decades, but data-driven constraints have rarely been studied. In this talk, we focus on the non-smooth convex-concave stochastic minimax regime and formulate the data-driven constraints as expectation constraints. Then, we propose a class of efficient primal-dual algorithms to tackle the minimax optimization with expectation constraints, and show that our algorithms converge at the optimal rate of $\mathcal O(\frac{1}{\sqrt{N}})$, where N is the number of iterations. We also verify the practical efficiency of our algorithms by conducting numerical experiments on large-scale real-world applications.

Dynamic stochastic projection method for multistage stochastic variational inequalities

Hailin Sun

Nanjing Normal University

Abstract. Stochastic approximation (SA) type methods have been well studied for solving single-stage stochastic variational inequalities (SVIs). This paper proposes a dynamic stochastic projection method (DSPM) for solving multistage SVIs. In particular, we investigate an inexact single-stage SVI and present an inexact stochastic projection method (ISPM) for solving it. Then we give the DSPM to a three-stage SVI by applying the ISPM to each stage. We show that the DSPM can achieve an $\mathcal{O}(\frac{1}{\epsilon^2})$ convergence rate regarding the total number of required scenarios for the three-stage SVI. We also extend the DSPM to the multistage SVI when the number of stages is larger than three. The numerical experiments illustrate the effectiveness and efficiency of the DSPM.

Frank-Wolfe-type methods for nonconvex inequality-constrained problems

Ting Kei Pong

The Hong Kong Polytechnic University

Abstract. The Frank-Wolfe method (also known as the conditional gradient method) implements efficient linear optimization oracles for minimizing smooth functions over compact convex sets. This method and its variants form a prominent class of projection-free first-order methods for a large variety of application problems such as matrix completion. In this talk, we extend this method to minimize smooth functions over a possibly nonconvex compact set, which is defined as the level set of a difference-of-convex function that satisfies mild regularity conditions. The key to our extension is the introduction of a new linear optimization oracle for the nonconvex compact constraint set. We also develop a variant of the classical away-step oracle in this nonconvex setting. We discuss convergence and present numerical experiments to illustrate the empirical performance of the proposed algorithm.