The Greater Bay Area Optimization Day

February 24, 2018

M1603, Core M, Li Ka Shing Tower

The Hong Kong Polytechnic University

Program

09:00–10:30 Morning Session I (Chair: Xiaojun Chen)

09:00–09:45 **Ya-xiang Yuan**, Theory and application of p-regularized subproblem with $p > 2$

09:45–10:30 **Zhiquan Luo**, A linearly convergent double stochastic Gauss-Seidel algorithm for linear systems

10:30–11:00 Tea Break & Group Photo

11:00–12:30 Morning Session II (Chair: Xiaoqi Yang)

11:00–11:45 **Duan Li**, Reformulations, relaxations and algorithms for QCQP

11:45–12:30 **Tong Zhang**, Composite functional gradient learning of generative adversarial models

12:45–13:45 Lunch

14:00–16:15 Afternoon Session I (Chair: Ting Kei Pong)

14:00–14:45 **Xiao Qing Jin**, A Riemannian inexact Newton-CG method for nonnegative inverse eigenvalue problems

14:45–15:30 **Anthony Man-Cho So**, Non-convex phase synchronization via projected gradient ascent with provable estimation and convergence guarantees

15:30–16:15 **Xiaojun Chen**, Two-stage stochastic variational inequalities

16:15–16:30 Tea Break

16:30–18:00 Afternoon Session II (Chair: Defeng Sun)

16:30–18:00 **Panel discussion** (Facilitators: Ya-xiang Yuan, Zhiquan Luo, Duan Li, Tong Zhang)

Coordinator: Zaikun Zhang, `zaikun.zhang@polyu.edu.hk`, (00852) 2766 4592

Theory and application of p-regularized subproblem with $p > 2$

Ya-xiang Yuan

yyx@lsec.cc.ac.cn

Academy of Mathematics and Systems Science, Chinese Academy of Sciences

The p-regularized subproblem (p-RS) is the key content of a regularization technique in computing a Newton-like step for unconstrained optimization. The idea is to incorporate a local quadratic approximation of the objective function with a weighted regularization term $(\sigma/p)||x||^p$ and then globally minimize it at each iteration. In this paper, we establish a complete theory of the p-RSs for general $p > 2$ that covers previous known results on $p = 3$ or $p = 4$. The theory features necessary and sufficient optimality conditions for the global and also for the local non-global minimizers of (p-RS). It gives a closed-form expression for the global minimum set of (p-RS) and shows that (p-RS), $p > 2$ can have at most one local non-global minimizer. Our theory indicates that (p-RS) have all properties that the trust region subproblems do. In application, (p-RS) can appear in natural formulation for optimization problems. We found two examples. One is to utilize the Tikhonov regularization to stabilize the least square solution for an over-determined linear system; and the other comes from numerical approximations to the generalized Ginzburg-Landau functionals. Moreover, when (p-RS) is appended with m additional linear inequality constraints, denoted by (p-RS$_m$), the problem becomes NP-hard. We show that the partition problem, the k-dispersion-sum problem and the quadratic assignment problem in combinatorial optimization can be equivalently formulated as special types of (p-RS$_m$) with $p = 4$. In the end, we develop an algorithm for solving (p-RS$_m$).

A linearly convergent double stochastic Gauss-Seidel algorithm for linear systems

Zhiquan Luo

luozq@cuhk.edu.cn

The Chinese University of Hong Kong (Shenzhen)

In this talk, we propose a double stochastic Gauss-Seidel (G-S) type algorithm for solving a linear system of equations $Ax = b$. Although it is generally known that the G-S scheme and many of its variants diverge when A is neither diagonal dominant nor symmetric positive definite, we show a positive result: with a random choice of equations and variables, the (over relaxed) G-S algorithm converges globally linearly (in expectation) for any feasible linear systems of equations. The key in the algorithm design is to introduce a nonuniform and double randomization rule for picking the equations and the variables in each update step. Our analysis also generalizes to certain iterative alternating projection algorithms for solving the linear inequality system $Ax \leq b$ with an arbitrary A. Our result demonstrates that properly introducing randomization can ensure global linear convergence of an (otherwise divergent) G-S type algorithm for arbitrary A.

This is a joint work with Mingyi Hong, Meisam Razaviyayn, Navid Reyhanian.

Reformulations, relaxations and algorithms for QCQP

Duan Li

dli226@cityu.edu.hk

City University of Hong Kong

We report in this talk our research progress on nonconvex QCQP. First, we derive a new convex relaxation for general nonconvex QCQP with nonconvex quadratic constraints and linear constraints, which yields tighter bounds than the existing ones. We consider next the singly-constrained QPQP, which is also known as the generalized trust region subproblem (GTRS). We finally propose two novel reformulations and algorithms to solve the GTRS more efficiently.

Composite functional gradient learning of generative adversarial models

Tong Zhang

tongzhang0@gmail.com

Tencent AI Lab

Generative adversarial networks (GAN) have become popular for generating data that mimic observations by learning a suitable variable transformation from a random variable. However, empirically, GAN is known to suffer from instability. Also, the theory provided based on the minimax optimization formulation of GAN cannot explain the widely-used practical procedure that uses the so-called logd trick. This paper provides a different theoretical foundation for generative adversarial methods which does not rely on the minimax formulation. We show that with a strong discriminator, it is possible to learn a good variable transformation via functional gradient learning, which updates the functional definition of a generator model, instead of updating only the model parameters as in GAN. The theory guarantees that the learned generator improves the KL-divergence between the probability distributions of real data and generated data after each functional gradient step, until the KL-divergence converges to zero. This new point of view leads to enhanced stable procedures for training generative models that can utilize arbitrary learning algorithms. It also gives a new theoretical insight into the original GAN procedure both with and without the logd trick. Empirical results are shown on image generation to illustrate the effectiveness of our new method.

Joint work with Rie Johnson.

A Riemannian inexact Newton-CG method for nonnegative inverse eigenvalue problems

Xiao Qing Jin

xqjin@umac.mo

University of Macau

This talk is concerned with the nonnegative inverse eigenvalue problem of finding a nonnegative matrix such that its spectrum is the prescribed self-conjugate set of complex numbers. We first reformulate the nonnegative inverse eigenvalue problem as an underdetermined constrained nonlinear matrix equation over several matrix manifolds. Then we propose a Riemannian inexact Newton-CG method for solving the nonlinear matrix equation. The global and quadratic convergence of the proposed method is established under some mild conditions. Finally, numerical experiments are reported to illustrate the efficiency of the proposed method.

Non-convex phase synchronization via projected gradient ascent with provable estimation and convergence guarantees

Anthony Man-Cho So

manchoso@se.cuhk.edu.hk

The Chinese University of Hong Kong

The problem of phase synchronization with its numerous applications has attracted much research in recent years. Currently, there are two main approaches for computing the maximum likelihood estimator of the unknown phase vector. One is semidefinite relaxation and the other is to directly solve the non-convex maximum likelihood formulation by the generalized power method (GPM), which is a projected gradient method. In this talk, we focus on the latter approach and study it from both statistical and optimization-theoretic viewpoints. Our contribution is twofold. First, we show that the ℓ_2- and ℓ_∞-bounds on the estimation error of the GPM decrease geometrically to the best provable bound under the least restrictive noise level requirement. Second, we prove that the objective value and the sequence of iterates of GPM converge linearly to the optimal value and the unique global optimizer, respectively, under a less restrictive requirement compared with that in the literature. This answers an open question raised in a recent work of Boumal.

Two-stage stochastic variational inequalities

Xiaojun Chen

xiaojun.chen@polyu.edu.hk

The Hong Kong Polytechnic University

The two-stage stochastic variational inequality (SVI) provides a powerful modeling paradigm for many important applications in which uncertainties and equilibrium are present. The two-stage SVI is to find a pair: here-and-now solution and wait-and-see solution. The here-and-now solution represents now-decisions, while the wait-and-see solution depends on future events described by random variables. This talk reviews new developments in theory and algorithms for two-stage SVI, including joint work with Ting-kei Pong, Alexander Shapiro, Defeng Sun, Hailin Sun, Roger Wets, Huifu Xu and Junfeng Yang.