Optimization Day

Date: Thursday 5 September 2019

Time & Venue:

1.45pm to 6.00pm

Lecture Theatre 31, S16 Level 3

Organisers:

Toh Kim Chuan (NUS)

SPEAKERS

Jean B. Lasserre (LAAS-CNRS)

Li Xudong (Fudan)

Soh Yong Sheng (IHPC)

Lee Ching-pei (NUS)

Pang Chin-How Jeffrey (NUS)

Sandra Tan (NUS)

Antonios Varvitsiotis (NUS)

NUS National University of Singapore | Department of Mathematics Faculty of Science

Programme

1.45pm to 2.30pm	**Linking optimization with spectral analysis of tri-diagonal Hankel matrices** *Jean B. Lasserre, LAAS-CNRS*
2.30pm to 3.00pm	**Graph isomorphism: conic formulations and their physical interpretation** *Antonios Varvitsiotis, NUS*
3.00pm to 3.30pm	**Deterministic distributed asynchronous optimization with Dykstra's splitting** *Pang Chin How, Jeffrey, NUS*
3.30pm to 4.00pm	Break
4.00pm to 4.30pm	**Analysis of Optimization Algorithms via Sums-of-Squares** *Sandra Tan, NUS*
4.30pm to 5.00pm	**Common-directions Method and its Limited-memory version for Smooth Optimization** *Lee Ching-pei, NUS*
5.00pm to 5.30pm	**Fitting Convex Sets to Data** *Soh Yong Sheng, IHPC*
5.30pm to 6.00pm	**Exploiting Second Order Sparsity in Big Data Optimization** *Li Xudong, Fudan*
6.30pm	Dinner (for workshop speakers and invited guests only)

Abstract

Linking optimization with spectral analysis of tri-diagonal Hankel matrices

Jean B. Lasserre, LAAS-CNRS

(1.45pm to 2.30pm)

We consider the problem of minimizing a continuous function "f" on a compact set K. By considering the pushforward measure (on the real line) of the Lebesgue measure on K by the mapping "f", one may approximate the global minimum of "f" by computing a converging sequence of upper bounds (f_d), d in N. Each upper bound f_d is the smallest eigenvalue of a certain tri-diagonal Hankel matrix of size "d". When "f" is & polynomial and "K" is a simple set (e.g., box, ellipsoid, simplex, and their affine transformation) it yields a practical numerical scheme. In doing so, one essentially reduces a multivariate optimization problem to a univariate one. Such an approach can also be used to compute the Lebesgue volume of the (compact) sublevel set of a polynomial.

Graph isomorphism: conic formulations and their physical interpretation

Antonios Varvitsiotis, NUS
(Co-Authors: A. Atserias, L. Maninska, D. Roberson, Ramal, S. Severini, A. Varvitsiotis)

(2.30pm to 3.00pm)

We introduce the (G,H)-isomorphism game where classical players win with certainty if and only if G and H are isomorphic. We then define the notions of quantum and non-signalling isomorphism, by considering perfect quantum and non-signalling strategies for the (G,H)-isomorphism game, respectively. First, we prove that non-signalling isomorphism coincides with fractional graph isomorphism. Second, we show there exist graphs that are quantum isomorphic but not isomorphic. Lastly, we show that both classical and quantum isomorphism can be reformulated as feasibility programs over the completely positive and completely positive semidefinite cones respectively, and give a combinatorial interpretation of the relaxation obtained using the doubly nonnegative cone.

Deterministic distributed asynchronous optimization with Dykstra's splitting

Pang Chin How, Jeffrey , NUS

(3.00pm to 3.30pm)

Consider the setting where each vertex of a graph has a function, and communications can only occur between vertices connected by an edge. We wish to minimize the sum of these functions in a distributed manner. For the case when each function is the sum of a strongly convex quadratic and a convex function, we propose a distributed version of Dykstra's algorithm. This algorithm is distributed, decentralized, asynchronous, has deterministic convergence. We show how to extend to directed graphs with unreliable communications, and present convergence rates results and linear convergence conditions.

Analysis of Optimization Algorithms via Sums-of-Squares

Sandra Tan, NUS

(4.00pm to 4.30pm)

In this work, we introduce a new framework for unifying and systematizing the performance analysis of first-order black-box optimization algorithms for unconstrained convex minimization over finite-dimensional Euclidean spaces. The low-cost iteration complexity enjoyed by this class of algorithms renders them particularly relevant for applications in machine learning and large-scale data analysis. However, existing proofs of convergence of such optimization algorithms consist mostly of case-by-case analyses. On the other hand, our approach is based on sum-of-squares optimization and puts forward a promising framework for unifying the convergence analyses of optimization algorithms. Illustrating the usefulness of our approach, we recover several known convergence bounds for four widely-used first-order algorithms in a unified manner, and also derive one new convergence result for gradient descent with Armijo-terminated line search.

Common-directions Method and its Limited-memory version for Smooth Optimization

Lee Ching-pei, NUS

(4.30pm to 5.00pm)

In this talk, I will present a new algorithm for smooth optimization, called common-directions method, and its limited-memory version. This algorithm interpolates between first-order and second-order methods in the sense that at each iteration, a subspace is constructed using first-order information from the current and previous iterations, and an efficient second-order method is deployed to find an approximate minimizer within this subspace.

We show that the proposed algorithm achieves the optimal convergence rates for first-order methods while remaining a descent method when the subproblems are solved with sufficient accuracy, even if the subspace dimension is as small as two. This is the first "first-order" algorithm that achieves strict objective descent and optimal complexity simultaneously. When the subspace is the span of all previous gradients, local superlinear convergence of the original problem is inherited from the second-order subproblem solver.

Furthermore, even if the subproblem solves are obtained by running only one iteration of the subproblem solver, the algorithm still possesses good convergence speed on both convex and nonconvex problems and the empirical performance is superior.
Since most of the operations are dense matrix-matrix operations, the proposed method is suitable for both multicore parallelization within one machine and distributed optimization using multiple machines, as evidenced by our numerical study.

This is a joint work with Po-Wei Wang, Weizhu Chen, and Chih-Jen Lin.

Fitting Convex Sets to Data

Soh Yong Sheng, IHPC

(5.00pm to 5.30pm)

A number of problems in data analysis and signal processing, when viewed from the appropriate perspective, can be phrased in the language of fitting a convex set to data. Examples of such problems include the task of computing a common sparse representation for data often referred to as "dictionary learning" or "sparse coding" as well as the task of reconstructing a polytope from half-space information arising in tomographic and robotics applications.

In this talk, we describe an approach for fitting convex sets that are parameterized as linear projections of more structured convex sets residing in higher dimensional space. The lift-and-project perspective allows us to naturally generalize classical approaches to fitting problems to convex sets that are expressible via semidefinite programming. We show examples where fitting with semialgebraic descriptions lead to improved reconstructions, and we highlight conceptual connections between our geometric fitting task with some structured matrix factorization instances.

Exploiting Second Order Sparsity in Big Data Optimization

Li Xudong, Fudan

(5.30pm to 6.00pm)

In this talk, we shall demonstrate how second order sparsity (SOS) in important optimization problems can be exploited to design highly efficient algorithms. The SOS property appears naturally when one applies a semismooth Newton (SSN) method to solve the subproblems in a proximal point algorithm (PPA) designed for certain classes of structured convex optimization problems. With in-depth analysis of the underlying generalized Jacobians and sophisticated numerical implementation, one can solve the subproblems at surprisingly low costs. For lasso problems with sparse solutions, the cost of solving a single PPA subproblem by our second order method is comparable or even lower than that in a single iteration of many first order methods. Consequently, with the fast convergence of the SSN based PPA, we are able to solve many challenging large scale convex optimization problems in big data applications efficiently and robustly. For the purpose of illustration, we present a highly efficient software called SuiteLasso for solving various well-known Lasso-type problems.