2013 PKU Workshop On Optimization and Data Sciences

December 22, 2013

Lecture Hall, Jia Yi Bing Building, 82 Jing Chun Yuan, BICMR

北京大学镜春园 82 号甲乙丙楼二层报告厅

08:40am-08:50am: Opening remarks

08:50am-11:30am: Session 1 (Chair: Naihua Xiu)

- 08:50am-09:25am: Wotao Yin
  Sparse regularization by evolving the l1 subgradient
- 09:25am-10:00am: Lingchen Kong
  Optimization in censored quantile regression
- 10:00am-10:35am: Li Gao
  An Uzawa-type algorithm for non-symmetric saddle point problems
- 10:35am-11:10am: Yafeng Liu
  Joint Power and Admission Control
- 11:10am-11:30am: Panel discussions

01:00pm-03:40pm: Session 2 (Chair: Wotao Yin)

- 01:00pm-01:35pm: Yuhong Dai
  Unconstrained Optimization Models for Computing Several Extreme Eigenpairs of Real Symmetric Matrices
- 01:35pm-02:10pm: Zhouchen Lin
  Linearized alternating direction method: two blocks and multiple blocks
- 02:10pm-02:45pm: Ruibin Xi
  Computational analyses of high-throughput sequencing data
- 02:45pm-03:05pm: Panel discussions

03:25pm-05:30pm: Session 3 (Chair: Yuhong Dai)

- 03:25pm-04:00pm: Naihua Xiu
  Exact recovery for sparse signal via weighted l1 minimization
- 04:00pm-04:35pm: Hongchao Zhang
  Stochastic approximation methods for nonconvex stochastic composite optimization
- 04:35pm-05:10pm: Qingna Li
  Lower bound theory for Schatten-p quasi-norm least squares problem
- 05:10pm-05:30pm: Panel discussions

Sparse regularization by evolving the l1 subgradient

Wotao Yin
University of California, Los Angeles

This talk is about evolving the l1 subgradient by following a PDE in order to recover a sparse solution u to Au = b (when b has noise, Au ~= b). This approach is motivated by l1 minimization but does not minimize an l1-norm energy. Its solution is better than l1 when b has noise. The solution path u(t) is obtained by evolving the l1 subgradient p(t) as follows: partial_+ p(t) = - A*(A u(t) - b), subject to p(t) being an l1-norm subgradient evaluated at u(t). Under some right continuity conditions, we show that p(t) is unique and gives sparse u(t), which has important properties such as no bias and consistency while the LASSO model does not in general. Computational results will demonstrate that u(t) is either sparser or fits data better, or both, than LASSO. The PDE is known as inverse-scale space in image processing and related to Bregman regularization and linearized Bregman. This is joint work with people at UCLA and PKU.

Optimization in Censored Quantile Regression

Lingchen Kong
Beijing Jiaotong University

In this talk, we will review the basic concepts and results on censored quantile regression, which include the background, history development and recent works in statistics. Then we focus on the mathematical models appeared in statistic approaches to censored quantile regression. In order to understand the optimization models and establish the efficient algorithms, we finally study the properties of Lȍwner operator generated by the quantile function. More specifically, we will consider its continuity, (locally) Lipschitz continuity, directional differentiability, F-differentiability, continuous differentiability, and strong semismoothness.

An Uzawa-type algorithm for non-symmetric saddle point problems

Li Gao
Peking University

Non-symmetric saddle point systems arise in scientific and engineering problems, such as in certain discretization of Navier-Stokes equations, and mixed discretization of second order elliptic problems with convective terms. In this talk, we will review Uzawa-type algorithms for symmetric and non-symmetric saddle point problems, and then propose an Uzawa-type algorithm for non-symmetric saddle point problems.

Joint Power and Admission Control

Yafeng Liu
Chinese Academy of Sciences

In the first part of this talk, we shall consider the joint power and admission control problem for a wireless network consisting of multiple interfering links where the channel state information is assumed to be perfectly known. The goal is to support a maximum number of links at their specified signal to interference plus noise ratio (SINR) targets while using a minimum total transmission power. We first reformulate this NP-hard problem as a sparse $¥ell_0$-minimization problem and then approximate it by a linear program. Furthermore, we derive two easy-to-check necessary conditions for all links in the network to be simultaneously supported at their target SINR levels, and use them to iteratively remove strong interfering links (deflation). Numerical simulations show that the proposed approach compares favorably with the existing approaches in terms of the number of supported links, the total transmission power, and the execution time. In the second part of this talk, we shall discuss some interesting extensions of the first part. For instance, instead of doing linear programming approximation, we can do non-convex approximations to improve the approximation performance. Also, we shall discuss how to extend the algorithm developed in the first part to the imperfect channel state information scenario.

Unconstrained Optimization Models for Computing Several Extreme Eigenpairs of Real Symmetric Matrices

Yuhong Dai
Chinese Academy of Sciences

This paper considers the problem of computing several extreme eigenpairs of real symmetric matrices. Based on the variational principles, we put forward some new unconstrained models for this classical problem and further analyze some significant properties. It is shown that the extreme eigenpairs of any real symmetric matrix can be recovered from the global minimizers of our unconstrained quartic and beta-order models. The alternate Barzilai-Borwein method with the adaptive nonmonotone line search is then utilized for solving the unconstrained models. The preliminary numerical results indicate that our approach is promising. This is a joint work with Bo Jiang and Chunfeng Cui.

Linearized Alternating Direction Method: Two Blocks and Multiple Blocks

Zhouchen Lin
Peking University

Alternating direction method (ADM) is an intuitive yet powerful method for various convex optimization problems. However, the traditional ADM assumes that each subproblem is easy to solve and its convergence is proven only in the case of two blocks. Such limitations greatly prevent ADM from wider applications to more complex problems. So I generalize ADM in two aspects. First, I linearize the quadratic penalty term and update the penalty parameter adaptively, introducing linearized ADM (LADM) with adaptive penalty. Second, I modify LADM slightly to account for the multiple block case, introducing linearized ADM with parallel splitting and adaptive penalty. Although neither aspect is brand new, deeper results are achieved in the scenario of machine learning and signal processing and the proposed algorithms fit for engineering use much better.

Computational analyses of high-throughput sequencing data

Ruibin Xi
Peking University

Abstract: The breakthrough of the high-throughput sequencing technologies has revolutionized many fields in biological and biomedical research. These technologies can cost-effectively generate a huge amount of data in a short period of time and allows researcher to investigate various biological processes with unprecedented resolution. However, computational analysis is becoming the main bottleneck of biological assay with the HTS data. In this talk, I will be discussing several statistical models and computational algorithms for analyzing HTS data and discuss challenges we encountered for HTS data analysis.

Exact Recovery for Sparse Signal via Weighted L_1 Minimization

Naihua Xiu
Beijing Jiaotong University

Numerical experiments have indicated that the reweighted L1 minimization perform exceptionally well in recovering sparse signal. In this paper, we study exact recovery conditions for sparse signal via weighted L1 minimization from the insight of the classical NSP (null space property) and RIC (restricted isometry constant) bound. We first introduce the concept of WNSP (weighted null space property) and reveal that it is a necessary and sufficient condition for exact recovery. We then prove that the RIC bound by weighted L1 minimization is larger than the sharp one by L1 minimization, and also greater than 0.4343 via weighted L1 minimization under some mild cases.

Stochastic Approximation Methods for Nonconvex Stochastic Composite Optimization

Hongchao Zhang
Louisiana State University

We consider a class of constrained stochastic composite optimization problems whose objective function is given by the summation of a differentiable (possibly nonconvex) component, together with a certain nondifferentiable (but convex) component. A randomized stochastic projected gradient (or gradient-free) algorithm will be introduced. In these algorithms, proper mini-batch of samples are taken at each iteration depending on the total budget of stochastic samples allowed. Complexity and some preliminary numerical results of these algorithms will be also discussed.

Lower Bound Theory for Schatten-p quasi-norm Least Squares Problem

Qingna Li
Beijing Institute of Technology

Consider the Schatten-p regularized matrix least squares problem, where the matrix least square problem is regularized by the Schatten-p quasi-norm with $0 < p < 1$. Such problem has been receiving more and more attentions due to its wide applications in affine matrix rank minimization problem. For the diagonal matrix case, it reduces to the l2-lp vector minimization problem, which is getting more and more popular due to the numerical observations that it can lead to sparser solution than l1 minimization. Interesting theoretical results are established and numerical algorithms are discussed for l2-lp vector minimization problem. In particular, Chen, Xu and Ye established the lower bound theory for the nonzero entries in a sparse solution and a smoothing gradient algorithm is proposed to solve the l2-lp vector minimization problem. Therefore, a natural question is that whether such wonderful results for vector case can be extended to matrix case, i.e., the Schatten-p regularized matrix least squares problem. In this paper, we take a small step to answer part of the question: there is a similar lower bound theory for the Schatten-p regularized matrix least squares problem. Moreover, we characterize the first and second order necessary condition for the problem as well as its smoothing problem. A smoothing gradient method is proposed to solve the problem, where the developed lower bound is employed to the final solution to get a lower rank output. Numerical results are reported to confirm the efficiency of the proposed method.