2016 Workshop on Optimization and Eigenvalue Computation

June 24-26, 2016

Beijing, China

http://bicmr.pku.edu.cn/EigOpt-2016

2016 Workshop on Optimization and Eigenvalue Computation

June 24-26, 2016

Beijing, China

http://bicmr.pku.edu.cn/EigOpt-2016

Information for Participants

Conference Hotel For Invited Speakers

Hotel:	"Zhong Guan Xin Yuan" Global Village, Building 1 中关新园1号楼
Address:	No. 216 Zhongguancun North Road, Haidian District 北京市海淀区中关村北大街126号
Dates:	By default, the hotel room is reserved from June 23th (check in) to June 27th (check-out). Please let us know if you have a different arrival-departure schedule.
Arrival:	By air, please see this link By subway: line 4 to "east gate of Peking University"
Website:	www.pkugv.com
Tel:	+86-10-62752288

Conference Venue

Venue:	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
Map:	PKU campus map

Meals

- Breakfasts will be complementary at the hotel.
- Lunches and dinners are provided by the workshop. Please let us know if you have any dietary restrictions or preferences.

Currency

Chinese currency is RMB. The current rate is about 6.56 RMB for 1 US dollar. The exchange of foreign currency can be done at the airport or the conference hotel. Please keep the receipt of the exchange so that you can change back to your own currency if you have RMB left before you leave China. Please notice that some additional processing fee will be charged if you exchange currency in China.

Parking at PKU Campus

If you plan to drive to PKU, please send us your license plate number; otherwise, your car cannot enter the PKU campus.

Contact Information

If you need any help, please feel free to contact

- Prof. Zaiwen Wen: wenzw@pku.edu.cn
- Jiang Hu: jianghu@pku.edu.cn

Sponsors

National Natural Science Foundation of China

Beijing International Center For Mathematical Research, Peking University

The Mathematical Programming Branch of OR Society of China

Committees

Organizing Committee

Zhaojun Bai, University of California, Davis

Yuhong Dai, Chinese Academy of Sciences

Xin Liu, Chinese Academy of Sciences

Zaiwen Wen, Peking University

Wotao Yin, University of California, Los Angeles

Yaxiang Yuan, Chinese Academy of Sciences

Scientific Committee

Yaxiang Yuan, Chinese Academy of Sciences

Weinan E, Peking University and Princeton University

Pingwen Zhang, Peking University

Conference Schedule

Each talk is 35 minutes + 5 minutes for questions.

June 24, Friday

08:30-09:00 Opening Ceremony

08:30-08:40 Welcome Address

08:40-09:00 Group Photo

09:00-10:20 Session F1
Chair: Zhaojun Bai

09:00-09:40 Wotao Yin, Coordinate update algorithms

09:40-10:20 Mengdi Wang, Stochastic composition optimization: algorithms and sample complexities

10:20-10:40 Coffee Break

10:40-12:00 Session F2
Chair: Wotao Yin

10:40-11:20 Haizhao Yang, Preconditioning orbital minimization method for planewave discretization

11:20-12:00 Lin Lin, Adaptively compressed exchange operator

12:00-13:30 Lunch

13:30-15:30 Session F3
Chair: Lin Lin

13:30-14:10 Weizhu Bao, Fundamental gaps and energy asymptotics of the Gross-Pitaevskii/nonlinear Schroedinger equation with repulsive interaction

14:10-14:50 Aihui Zhou, A conjugate gradient optimization method for electronic structure calculations

14:50-15:30 Chao Yang, Solving large-scale configuration interaction eigenvalue problem

15:30-15:50 Coffee Break

15:50-17:50 Session F4
Chair: Kim-Chuan Toh

15:50-16:30 Naihua Xiu, Solution analysis of low-rank matrix optimization

16:30-17:10 Xiaojun Chen, Sparse solutions of the LCP and low rank solutions of the semidefinite LCP

17:10-17:50 Chao Ding, Convex optimization learning of faithful Euclidean distance representations in nonlinear dimensionality reduction

18:10 Dinner

June 25, Saturday

09:00-10:20 Session S1
Chair: Qiang Ye

09:00-09:40 **Bart Vandereycken**, Robust methods for computing extremal points of real pseudospectra

09:40-10:20 **Zhaojun Bai**, Solving eigenvalue problems arising in Rayleigh quotient optimizations

10:20-10:40 Coffee Break

10:40-12:00 Session S2
Chair: Chao Yang

10:40-11:20 **Xin Liu**, An efficient Gauss-Newton algorithm for symmetric low-rank product matrix approximation

11:20-12:00 **Yin Zhang**, Accelerating convergence by augmented Rayleigh-Ritz projections For large-scale eigenpair computation

12:00-13:30 Lunch

13:30-15:30 Session S3
Chair: Defeng Sun

13:30-14:10 **Bo Jiang**, Tensor and its Tucker core: the invariance relationships

14:10-14:50 **Yimin Wei**, Generalized tensor eigenvalue problems

14:50-15:30 **Daniel Kressner**, Low-rank tensor completion by Riemannian optimization

15:30-15:50 Coffee Break

15:50-17:50 Session S4
Chair: Daniel Kressner

15:50-16:30 **Qiang Ye**, Computing singular values of large matrices with an inverse free preconditioned Krylov subspace method

16:30-17:10 **Ren-Cang Li**, Nonlinear eigenvalue problems from maximizing sum of trace ratios

17:10-17:50 **Zaiwen Wen**, Adaptive regularized method for optimization on Riemannian manifold

18:10 Dinner

June 26, Sunday

09:00-10:20 Session V1
Chair: Xiaojun Chen

09:00-09:40 **Kim Chuan Toh**, Fast algorithms for large scale generalized distance weighted discrimination

09:40-10:20 **Defeng Sun**, A two-phase proximal augmented Lagrangian method for convex quadratic semidefinite programming

10:20-10:40 Coffee Break

10:20-12:00 Session V2
Chair: Yin Zhang

10:40-11:20 **Yuhong Dai**, The steepest descent and conjugate gradient methods revisited

11:20-12:00 **Weiguo Gao**, Algorithms for group sparsity with overlap and beyond

12:00-13:30 Lunch

13:30-15:30 Session V3
Chair: Aihui Zhou

13:30-14:10 **Zhongxiao Jia**, The regularization theory of the Krylov iterative solvers LSQR, CGLS, LSMR and CGME for linear discrete ill-posed problems

14:10-14:50 **Yunkai Zhou**, Accelerating large eigenvalue calculations by spectrum-partition

14:50-15:30 **Yangfeng Su**, Two-level orthogonal Arnoldi procedure

17:30 Dinner

Abstracts

Solving eigenvalue problems arising in Rayleigh quotient optimizations

Zhaojun Bai
University of California, Davis
zbai@ucdavis.edu

Many computational science and data analysis techniques lead to optimizing Rayleigh-Quotient (RQ) and RQ-type objective functions, such as robust classification to handle uncertainty and constrained data clustering to incorporate a prior information. In this talk, we will discuss the origins of RQ and RQ-type optimizations and their reformulations to various algebraic eigenvalue problems. We will discuss how to exploit underlying properties of eigenvalue problems for design reliable and fast eigensolvers, and illustrate the efficacy of these eigensolvers in large-scale data analysis.

Fundamental gaps and energy asymptotics of the Gross-Pitaevskii/nonlinear Schroedinger equation with repulsive interaction

Weizhu Bao
National University of Singapore
matbaowz@nus.edu.sg

We study asymptotically and numerically fundamental gaps (i.e. difference between the first excited state and the ground state) in energy and chemical potential of the Gross-Pitaevskii equation (GPE) – nonlinear Schroedinger equation (NLSE) with cubic nonlinearity – with repulsive interaction under different trapping potentials including box potential and harmonic potential. Based on our asymptotic and numerical results, we formulate a gap conjecture on the fundamental gaps in energy and chemical potential of the GPE on bounded domains with the homogeneous Dirichlet boundary condition and in the whole space with a convex trapping potential growing at least quadratic in the far field. We extend these results to the GPE on bounded domains with either the homogeneous Neumann boundary conditionor periodic boundary condition. Finally, we obtain ground state approximations and their corresponding energy and chemical potential asymptotics of the NLSE with general nonlinearity and different trapping potentials for different parameter regimes and observe bifurcations in the ground states when the power of the nonlinearity becomes very large. Numerical results are reportedto support the asymptotic results. This is a joint work with Mr Xinran Ruan.

Sparse solutions of the LCP and low rank solutions of the semidefinite LCP

Xiaojun Chen
The Hong Kong Polytechnic University
xiaojun.chen@polyu.edu.hk

This talk considers the characterization and computation of sparse solutions and least-p-norm ($0 < p < 1$) solutions of the linear complementarity problem LCP(q, M). We show that the number of non-zero entries of any least-p-norm solution of the LCP(q, M) is less than or equal to the rank of M for any arbitrary matrix M and any number $p \in (0, 1)$, and there is $\bar{p} \in (0, 1)$ such that all least-p-norm solutions for $p \in (0, \bar{p})$ are sparse solutions. Moreover, we provide conditions on M such that a sparse solution can be found by solving convex minimization. Extension of sparse solutions of the LCP to low rank solutions of the semidefinite LCP is discussed.

The steepest descent and conjugate gradient methods revisited

Yuhong Dai
Chinese Academy of Sciences
dyh@lsec.cc.ac.cn

The steepest descent and conjugate gradient methods are basic first order methods for unconstrained optimization. More efficient variants have been proposed in recent decades by forcing them to approximate Newton's method (or quasi-Newton method). In this talk, I shall review some recent advances on steepest descent method and conjugate gradient method. While significant numerical improvements have been made, the behavior of these more efficient variants are still to be understood and more analysis are obviously required.

Convex optimization learning of faithful Euclidean distance representations in nonlinear dimensionality reduction

Chao Ding
Chinese Academy of Sciences
dingchao@amss.ac.cn

Classical multidimensional scaling only works well when the noisy distances observed in a high dimensional space can be faithfully represented by Euclidean distances in a low dimensional space. Advanced models such as Maximum Variance Unfolding (MVU) and Minimum Volume Embedding (MVE) use Semi-Definite Programming (SDP) to reconstruct such faithful representations. While those SDP models are capable of producing high quality configuration numerically, they suffer two major drawbacks. One is that there exist no theoretically guaranteed bounds on the quality of the configuration. The other is that they are slow in computation when the data points are beyond moderate size. In this talk, we propose a convex optimization model of Euclidean distance matrices. We establish a non-asymptotic error bound for the random graph model with sub-Gaussian noise, and prove that our model produces a matrix estimator of high accuracy when the order of the uniform sample size is roughly the degree of freedom of a low-rank matrix up to a logarithmic factor. Our results partially explain why MVU and MVE often work well. Moreover, we develop a fast inexact accelerated proximal gradient method. Numerical experiments show that the model can produce configurations of high quality on large data points that the SDP approach would struggle to cope with.

The regularization theory of the Krylov iterative solvers LSQR, CGLS, LSMR and CGME for linear discrete ill-posed problems

Zhongxiao Jia
Tsinghua University
jiazx@tsinghua.edu.cn

For the large-scale linear discrete ill-posed problem $\min \|Ax - b\|$ or $Ax = b$ with a noisy b, LSQR, which is a Krylov iterative solver based on Lanczos bidiagonalization, and its mathematically equivalent CGLS are most commonly used. They have intrinsic regularizing effects, where the number of iterations plays the role of regularization parameter. However, there has been no answer to the long-standing fundamental concern: *for which kinds of problems LSQR and CGLS can find best possible regularized solutions*. The concern was actually raised perspectively by Björck and Eldén in 1979. Here a best possible regularized solution means that it is at least as accurate as the best regularized solution obtained by the truncated singular value decomposition (TSVD) method, which and the best possible solution of standard-form Tikhonov regularization are both of the same order of the worst-case error and cannot be improved under the assumption that the solution to an underlying linear compact operator equation is continuous or its derivative squares integrable. In this paper we make a complete analysis on the regularization of LSQR for severely, moderately and mildly ill-posed problems. We first consider the case that the singular values of A are simple. We establish accurate $\sin\Theta$ theorems for the 2-norm distance between the underlying k-dimensional Krylov subspace and the k-dimensional dominant right singular subspace of A. Based on them and some follow-up results, for the first two kinds of problems, we prove the following results: (i) the k-step Lanczos bidiagonalization always generates a near best rank k approximation to A; (ii) the k Ritz values always approximate the first k large singular values in natural order; (iii) the k-step LSQR always captures the k dominant SVD components of A, so that LSQR can find a best possible regularized solution; (iv) the diagonals and subdi-

agonals of the bidiagonal matrices generated by Lanczos bidiagonalization decay as fast as the singular values of A. However, for the third kind of problem, the above results do not hold generally. The decay rates of diagonals and subdiagonals of the bidiagonal matrices can be used to decide if LSQR can find a best possible regularization solution. We also analyze the regularization of other two Krylov solvers LSMR and CGME that amount to MINRES and the CG method and MINRES applied to $A^TAx = A^Tb$ and $\min \|AA^Ty - b\|$ with $x = A^Ty$, respectively, proving that LSMR has similar regularizing effects to LSQR for each kind of problem and both are superior to CGME. We extend all the results to the case that A has multiple singular values. Numerical experiments confirm our theory on LSQR.

Tensor and its Tucker core: the invariance relationships

Bo Jiang
The Shanghai University of Finance and Economics
jiang.bo@mail.shufe.edu.cn

Hillar and Lim famously demonstrated that multilinear (tensor) analogues of many efficiently computable problems in numerical linear algebra are NP-hard". Despite many recent advancements, the state-of-the-art methods for computing such 'tensor analogues' still suffer severely from the curse of dimensionality. In this paper we show that the Tucker core of a tensor however, retains many properties of the original tensor, including the CP rank, the border rank, the tensor Schatten quasi norms, and the Z-eigenvalues. Since the core is typically smaller than the original tensor, this property leads to considerable computational advantages, as confirmed by our numerical experiments. In our analysis, we in fact work with a generalized Tucker-like decomposition that can accommodate any full column-rank factorization matrices.

Low-rank tensor completion by Riemannian optimization

Daniel Kressner

EPFL

daniel.kressner@epfl.ch

In tensor completion, the goal is to fill in missing entries of a partially known tensor under a low-rank constraint. We survey existing work in this area and discuss algorithms that perform Riemannian optimization techniques on the manifold of tensors of fixed multilinear rank or fixed tensor train rank. Paying particular attention to the efficient implementation, our algorithm scales linearly in the size of the tensor. Examples with synthetic data demonstrate good recovery even if the vast majority of the entries are unknown. We illustrate the use of the developed algorithm for a range of applications, including the recovery of multidimensional images and for the approximation of multivariate functions.

Nonlinear eigenvalue problems from maximizing sum of trace ratios

Ren-Cang Li
University of Texas at Arlington
rcli@uta.edu

We are concerned with the maximization of

$$\frac{trace(V^TAV)}{trace(V^TBV)} + trace(V^TCV)$$

over the Stiefel manifold $\{V \in \mathbb{R}^{m\times\ell} | V^TV = I_\ell\}$ ($\ell < m$), where B is a given symmetric and positive definite matrix, A and C are symmetric matrices, and $trace(\,\cdot\,)$ is the trace of a square matrix. Applications that give rise to such a problem include the the problem of balancing individual capacities in a multi-user MIMO downlink channel ($\ell = 1$) and the sparse Fisher discriminant analysis in pattern recognition ($C = 0$). We establish necessary conditions for both the local and global maximizers and connect the problem with a nonlinear extreme eigenvalue problem which can be solved by a self-consistent-field (SCF) iteration. We analyze the global and local convergence of the SCF iteration, and show that the necessary condition for the global maximizers is fulfilled at any convergent point of the sequences of approximations generated by the SCF iteration. This is one of the advantages of the SCF iteration over optimization-based methods. Numerical tests show that the SCF iteration is much more efficient than manifold-based optimization methods.

This is a joint work with Lei-hong Zhang (Shanghai University of Finance and Economics).

Adaptively compressed exchange operator

Lin Lin
University of California, Berkeley
linlin@math.berkeley.edu

The Fock exchange operator plays a central role in modern quantum chemistry, such as in Hartree-Fock calculations and Kohn-Sham density functional theory calculations with hybrid exchange-correlation functionals. The Fock exchange operator significantly increases the computational cost for solving the associated Kohn-Sham eigenvalue problem. We develop the adaptively compressed exchange ACE operator formulation, which greatly reduces the computational cost associated with the Fock exchange operator without loss of accuracy. The ACE formulation does not depend on the size of the band gap, and thus can be applied to insulating, semiconducting as well as metallic systems. Numerical results indicate that the ACE formulation can become advantageous even for small systems with tens of atoms.

An efficient Gauss-Newton algorithm for symmetric low-rank product matrix approximation

Xin Liu

Chinese Academy of Sciences

liuxin@lsec.cc.ac.cn

We derive and study a GaussNewton method for computing a symmetric low-rank product XX', where X is an n by k matrix for k < n, that is the closest to a given n by n symmetric matrix A in Frobenius norm. When A = B'B (or BB'), this problem essentially reduces to finding a truncated singular value decomposition of B. Our GaussNewton method, which has a particularly simple form, shares the same order of iteration-complexity as a gradient method when k¡¡n, but can be significantly faster on a wide range of problems. In this paper, we prove global convergence and a Q-linear convergence rate for this algorithm and perform numerical experiments on various test problems, including those from recently active areas of matrix completion and robust principal component analysis. Numerical results show that the proposed algorithm is capable of providing considerable speed advantages over Krylov subspace methods on suitable application problems where high-accuracy solutions are not required. Moreover, the algorithm possesses a higher degree of concurrency than Krylov subspace methods, thus offering better scalability on modern multi-core computers.

Two-level orthogonal Arnoldi procedure

Yangfeng Su
Fudan University
yfsu@fudan.edu.cn

The second-order Arnoldi (SOAR) procedure is an algorithm for computing an orthonormal basis of the second-order Krylov subspace. It has found applications in solving quadratic eigenvalue problems and model order reduction of second-order dynamical systems among others. Unfortunately, the SOAR procedure can be numerically unstable. The two-level orthogonal Arnoldi (TOAR) procedure has been proposed as an alternative of the SOAR to cure the numerical instability. In this paper, we provide a rigorous stability analysis of the TOAR procedure. We prove that under mild assumptions, the TOAR procedure is backward stable in computing an orthonormal basis of the associated linear Krylov subspace. The benefit of the backward stability of the TOAR is demonstrated by its high accuracy in structure-preserving model order reduction of second-order dynamical systems.

(Co-authored with Ding Lu and Zhaojun Bai)

A two-phase proximal augmented Lagrangian method for convex quadratic semidefinite programming

Defeng Sun
National University of Singapore
matsundf@nus.edu.sg

In machine learning, finance, statistics, and other areas, numerous interesting problems can be modelled into the form of convex composite quadratic conic programming. In this talk, we use the convex quadratic semidefinite programming (QSDP) problem as an example to introduce a two-phase proximal augmented Lagrangian method, called QSDPNAL, for solving convex composite quadratic conic programming problems with constraints consisting of a large number of linear equality, inequality constraints, a simple convex polyhedral set constraint, and a closed convex cone constraint. As the cornerstone of QSDPNAL, we first introduce the powerful and elegant inexact symmetric Gauss-Seidel (sGS) decomposition technique for solving a convex minimization problem whose objective is the sum of a multi-block (non-separable) quadratic function and a non-smooth function involving only the first block. A first order algorithm which takes advantage of our inexact sGS decomposition technique is adopted in our QSDPNAL-Phase I to generate a reasonably good initial point. Then, in QSDPNAL-Phase II, we design a proximal augmented Lagrangian method (ALM) where the inner sub-problem in each iteration is solved via the inexact accelerated block coordinate descent (ABCD) method, which again relies on our inexact sGS decomposition technique, together with the intelligent incorporation of the semi-smooth Newton-CG algorithm. Moreover, under certain suitable conditions, we are able to analyze the rate of convergence of the proposed algorithm. We further discuss the important numerical issues in the implementation of QSDPNAL. Extensive numerical results for various large scale QSDPs show that our two-phase framework is not only fast but also robust in obtaining accurate solutions. [This talk is based on a joint work with Xudong Li and Kim-Chuan Toh].

Fast algorithms for large scale generalized distance weighted discrimination

Kim-Chuan Toh
National University of Singapore
mattohkc@nus.edu.sg

High dimension low sample size statistical analysis is important in a wide range of applications. In such situations, the highly appealing discrimination method, support vector machine, can be improved to alleviate data piling at the margin. This leads naturally to the development of distance weighted discrimination (DWD), which can be modeled as a second-order cone programming problem and solved by interior-point methods when the scale (in sample size and feature dimension) of the data is moderate. Here, we design a scalable and robust algorithm for solving large scale generalized DWD problems. Numerical experiments on real data sets from the UCI repository demonstrate that our algorithm is highly efficient in solving large scale problems, and sometimes even more efficient than the highly optimized LIBSVM for solving the corresponding SVM problems.

(Co-authored with Xin-Yee Lam, J.S. Marron, Defeng Sun)

Robust methods for computing extremal points of real pseudospectra

Bart Vandereycken
University of Geneva
Bart.Vandereycken@unige.ch

We introduce a criss-cross type algorithm to compute the rightmost point in the real pseudospectrum of a given matrix, that is, the set of eigenvalues of all real, norm-bounded, additive perturbations of this matrix. Existing methods, which are based on eigenvalue optimization over the rank-2 matrix manifold using steepest descending schemes, may converge to only locally rightmost solutions and their convergence can be slow. To avoid these difficulties, we formally extend the criss-cross algorithm originally developed for unstructured pseudospectra but to reduce its computational cost, we exploit a supersets characterisation of the real pseudospectrum. Each criss and cross search is performed on carefully selected supersets, and involves solving only linear eigenvalue problems and singular value optimization problems. Furthermore, we also propose a subspace projection framework, which combines the criss-cross algorithm with subspace projection techniques to make the algorithm applicable to large-scale matrices. The developed algorithms are proven to be globally convergent and their robustness and efficiency are demonstrated by a number of numerical examples.

(Co-authored with Ding Lu)

Stochastic composition optimization: algorithms and sample complexities

Mengdi Wang
Princeton University
mengdiw@princeton.edu

Classical stochastic gradient methods are well suited for minimizing expected-value objective functions. However, they do not apply to the minimization of a composition of two expected-value functions, i.e., the stochastic composition optimization problem $\min_x \mathbf{E}_v \left[f_v \big(\mathbf{E}_w [g_w(x)] \big) \right]$ .Stochastic composition optimization finds wide application in learning, estimation, risk-averse optimization, dynamic programming, etc. In order to solve this problem, we propose a class of stochastic compositional first-order methods that can be viewed as stochastic versions of quasi-gradient method. The algorithms update the solutions based on noisy sample gradients of f_v, g_w and use auxiliary variables to track the unknown quantity $\mathbf{E}_w [g_w(x)]$. We prove that the algorithms converge almost surely to an optimal solution for convex optimization problems (or a stationary point for nonconvex problems), as long as such a solution exists. The convergence involves the interplay of two martingales with different timescales. We obtain rate of convergence results under various assumptions, and show that the algorithms achieve the optimal sample-error complexity in several important special cases. These results provide the best-known rate benchmarks for stochastic composition optimization. Indeed, stochastic composition optimization is very common in practice. We demonstrate its application to statistical estimation and reinforcement learning.

Generalized Tensor eigenvalue problems

Yimin Wei
Fudan University
ymwei@fudan.edu.cn

This talk is devoted to generalized tensor eigenvalue problems. We focus on the properties and perturbations of the spectra of regular tensor pairs. Employing different techniques, we extend several classical results from matrices or matrix pairs to tensor pairs, such as the Gershgorin circle theorem, the CollatzWielandt formula, the BauerFike theorem, the RayleighRitz theorem, backward error analysis, the componentwise distance of a nonsingular tensor to singularity, etc. Some of these results preserve their original forms, while others change when being extended.

Adaptive regularized method for optimization on Riemannian manifold

Zaiwen Wen

Peking University

wenzw@pku.edu.cn

Optimization on Riemannian Manifold widely arises in eigenvalue computation, density functional theory, Bose-Einstein condensates, image and signal processing. We propose a second-order type approximation to the original problem and apply a first-order type method solve it. Global convergence to the first-order optimality conditions is established. Preliminary numerical experiments show that our method is promising.

Algorithms for group sparsity with overlap and beyond

Weiguo Gao
Fudan University
wggao@fudan.edu.cn

We discuss the algorithms for group sparsity with overlap in which matrix-vector multiplications can be carried out efficiently. An accelerating technique to improve the performance of existing algorithms is proposed by solving a similar sub-problem. We further show that the new scheme can be formulated in the inner-outer regime by exploring the sparse structure. This can also be viewed from the dimensional reduction viewpoint. Convergence for ADMM (as an example) is guaranteed by a rigorous proof. Numerical experiments demonstrate the efficiency of the technique.

Solution analysis of low-rank matrix optimization

Naihua Xiu
Beijing Jiaotong University
nhxiu@bjtu.edu.cn

The Low-rank matrix optimization (LRMO) is to minimize a general continuously differentiable function subject to Low-rank and equality/inequality sets, and it arises in many applications, including matrix regression, matrix completion and matrix decomposition problems, etc. In this talk, we mainly discuss tangent cones and normal cones of low-rank set, and establish and analyze the optimality conditions for the LRMO.

(Co-authored with Xinrong Li)

Preconditioning orbital minimization method for planewave discretization

Haizhao Yang
Duke University
haizhao@math.duke.edu

We present an efficient preconditioner for the orbital minimization method when the Hamiltonian is discretized using planewaves (i.e., pseudospectral method). This novel preconditioner is based on an approximate Fermi operator projection by pole expansion, combined with the sparsifying preconditioner to efficiently evaluate the pole expansion for a wide range of Hamiltonian operators. Numerical results validate the performance of the new preconditioner for the orbital minimization method, in particular, the iteration number is reduced toO(1)and often only a few iterations are enough for convergence.

(Co-authored with Jianfeng Lu)

Solving large-scale configuration interaction eigenvalue problem

Chao Yang
Lawrence Berkeley Lab
cyang@lbl.gov

Configuration interaction (CI) is a technique often used in quantum chemistry and nuclear physics to obtain approximate solutions to a many-body Schrodinger eigenvalue problem. A typical CI calculation involves computing the smallest eigenvalues and the corresponding eigenvectors of a large sparse symmetric matrix. The dimension of the matrix, which increases rapidly with respect to the number of particles in the quantum system and the size of the model configuration space, can reach billions. To solve this type of problem efficiently, we must exploit the structure of problem. In this talk, I will describe properties of nuclear CI Hamiltonian matrices and discuss how some of these properties can be used to develop efficient iterative methods for solving this type of problem.

Computing singular values of large matrices with an inverse free preconditioned Krylov subspace method

Qiang Ye
University of Kentucky
qiang.ye@uky.edu

We present an efficient algorithm for computing a few extreme (largest and smallest) singular values and corresponding singular vectors of a large sparse m × n matrix C. Our algorithm is based on reformulation of the singular value problem as an eigenvalue problem for the normal equation matrix and, to address the clustering of singular values, we use an inverse-free preconditioned Krylov subspace method to accelerate convergence. We propose a new deflation scheme that is more efficient for the singular value problem and we present a convergence analysis. We consider preconditioning that is based on robust incomplete factorizations and we discuss various implementation issues. Extensive numerical tests are presented to demonstrate efficiency and robustness of the new algorithm.

Coordinate update algorithms

Wotao Yin
University of CaliforniaLos, Angeles
wotao.yin@gmail.com

This talk focuses on a class of algorithms, called coordinate update algorithms, which are useful at solving large-sized problems involving linear and nonlinear mappings, and smooth and nonsmooth functions. They decompose a problem to simple subproblems, where each subproblem updates one, or a small block of, variables each time. They have found applications throughout signal/imaging processing, differential equations, and machine learning. We abstract many problems to the fixed-point problem $x^{k+1} = Tx^k$. This talk discusses the favorable structures of the operator T that enable highly efficient coordinate update iterations. It can be carried out in sequential, parallel, or async-parallel fashions. We introduce new scalable coordinate-update algorithms to many problems involving coupling constraints $Ax = b$, composite nonsmooth functions $f(Ax)$, and large-scale data. We will present numerical examples. This is joint work with Zhimin Peng and Tianyu Wu (UCLA), Yangyang Xu (IMA), and Ming Yan (MSU).

Accelerating convergence by augmented Rayleigh-Ritz projections for large-scale eigenpair computation

Yin Zhang
Rice University
yzhang@rice.edu

Iterative algorithms for large-scale eigenpair computation are mostly based subspace projections consisting of two main steps: a subspace update (SU) step that generates bases for approximate eigenspaces, followed by a Rayleigh-Ritz (RR) projection step that extracts approximate eigenpairs. A predominant methodology for the SU step makes use of Krylov subspaces that builds orthonormal bases piece by piece in a sequential manner. On the other hand, block methods such as the classic (simultaneous) subspace iteration, allow higher levels of concurrency than what is reachable by Krylov subspace methods, but may suffer from slow convergence. In this work, we analyze the rate of convergence for a simple block algorithmic framework that combines an augmented Rayleigh-Ritz (ARR) procedure with the subspace iteration. Our main results are Theorem **??** and its corollaries which show that the ARR procedure can provide significant accelerations to convergence speed. Our analysis will offer useful guidelines for designing and implementing practical algorithms from this framework.

Subspace iteration of nonlinear eigenvalue problems for multi-view learning

Zhenyue Zhang
Zhejiang University
zyzhang@zju.edu.cn

This talk focuses on an nonlinear and nonconvex optimization problem for multi-view learning. Its first-order necessary condition of the optimal solutions result in *nonlinear* eigenvalue problem whose coefficient matrix depends on the solution. We will discuss an iterative algorithm to solve the nonlinear eigenvalue problem by solving a *linear* eigenvalue problem in the scale of the number of objects, using a previously estimated solution in the coefficient matrix. Its convergence is discussed, too. We will also give a fast implementation of the algorithm. Three strategies are used to reduce the computation cost and speed up the convergence: running the eigen-space iteration algorithm in a low-dimensional feasible space, extending the feasible space, and restarting the procedure after extending the space several times to control the ever-increasing scale. Numerical experiments show that the proposed methods have competitive performances on the real-world data sets reported.

Accelerating large eigenvalue calculations by spectrum-partition

Yunkai Zhou
Southern Methodist University
yzhou@smu.edu

We present two recently developed spectrum-partition methods for solving very large eigenvalue problems. Many eigen-algorithms have been proposed for large eigenvalue problems, the more efficient ones of them have the O(nk^2) complexity, where n is the matrix size and k is the number of eigenvalues to be computed. Therefore these algorithms inevitably become inefficient when k is very large. Spectrum-partition can be utilized to overcome the difficulty associated with large k. However, spectrum-partition methods have their own difficulties when not designed properly. We address several of the intrinsic difficulties associated with spectrum-partition methods, and present theoretical as well as numerical results to show that our methods are practical.

A conjugate gradient optimization method for electronic structure calculations

Aihui Zhou
Chinese Academy of Sciences
azhou@lsec.cc.ac.cn

In this presentation, we talk a conjugate gradient method for electronic structure calculations. We propose a Hessian based step size strategy, which together with three orthogonal approaches yields three algorithms for computing the ground state energy of atomic and molecular systems. Under some mild assumptions, we prove that our algorithms converge locally. It is shown by our numerical experiments that the conjugate gradient method performs quite well.

The organizing committee wishes you a pleasant stay in BICMR!