CAS-ANLA workshop

April 23rd and 24th in N204

Scan me for more information

1 Introduction

Linear algebra kernels and algorithms are prevalent in scientific computing, machine learning, statistics, image processing, and much more. With the ever-increasing scale and complexity of some of those computational problems, the need for scalable and cost-effective linear algebra algorithms is critical. For this reason, approximate computing techniques have been increasingly used in modern linear algebra algorithms. These techniques can substantially reduce the algorithms' complexities by introducing (controlled) errors into the calculations. Those techniques include, but are not limited to, low-rank approximations, mixed precision, randomization, or tensor compression.

The Chinese Academy of Sciences Workshop on Approximate Computing in Numerical Linear Algebra (CAS-ANLA) aims to provide greater exposure to the topic of approximate computing techniques to the local Chinese and Asian scientific community. It will gather outstanding researchers from China working in approximate computing together with some of the best international experts in the field. The speakers will illustrate theoretically and experimentally the benefit of cutting-edge approximate computing techniques and present recent advances in the field. We also expect other topics to be represented, such as optimization, machine learning, or functions of matrices to name a few.

The CAS-ANLA workshop will be spread over two days and will feature 14 plenaries. In total, 13 different institutions are represented, 6 of them are from abroad, 3 of them are Chinese institutes outside Beijing.

2 Program 23rd of April

Time	Speakers	Title
8:30 - 9:00	Weiying Zheng	Opening Remarks
9:00 - 9:50	Weifeng Liu (chair: Xin Liu)	Block-Wise Mixed Precision Sparse Solvers
9:50 - 10:40	Tomohiro Sogabe (chair: Xin Liu)	Recent Developments in Krylov Subspace Methods for Nonsymmetric Shifted Linear Systems
10:40 - 11:00		Coffee break
11:00 - 11:50	Junfeng Yin (chair: Chensong Zhang)	Surrogate hyperplane Bregman – Kaczmarz methods for solving linear inverse problems
12:00 - 14:00		Lunch break
14:00 - 14:50	Luc Giraud (chair: Bastien Vieublé)	A journey through some numerical linear algebra algorithms with variable accuracy
14:50 - 15:40	Meiyue Shao (chair: Bastien Vieublé)	Exploiting low-precision arithmetic in eigensolvers
15:40 - 16:00		Coffee break
16:00 - 16:50	Chensong Zhang (chair: Junfeng Yin)	Learning-based Multilevel Solvers for Large-Scale Linear Systems
16:50 - 17:40	Xinye Chen (chair: Junfeng Yin)	Mixed-precision HODLR matrices

3 Program 24th of April

Time	Speakers	Title
8:30 - 9:00		Welcoming
9:00 - 9:50	Françoise Tisseur (chair: Qingna Li)	Computing accurate eigenvalues using a mixed-precision Jacobi algorithm
9:50 - 10:40	Qinmeng Zou (chair: Qingna Li)	Stability of modified Gram-Schmidt and its low-synchronization variants
10:40 - 11:00		Coffee break
11:00 - 11:50	Alena Kopanicakova (chair: Bastien Vieublé)	Towards trustworthy use of scientific machine-learning in large scale numerical simulations
12:00 - 14:00		Lunch break
14:00 - 14:50	Zaikun Zhang (chair: Tao Cui)	Solving 10,000-Dimensional Optimization Problems Using Inaccurate Function Values: An Old Algorithm
14:50 - 15:40	Qingna Li (chair: Tao Cui)	A Fast BB Reduced Minimization Algorithm for Nonnegative Viscosity Optimization in Optimal Damping
15:40 - 16:00		Coffee break
16:00 - 16:50	Mantas Mikaitis (chair: Bastien Vieublé)	Error Analysis of Matrix Multiplication with Narrow Range Floating-Point Arithmetic
16:50 - 17:40	Massimiliano Fasi (chair: Bastien Vieublé)	Analysis of Floating-Point Matrix Products Computed via Low-Precision Integer Arithmetic
17:40 - 17:50	Organizers	Closing Remarks

4   How to come?

地址: 北京市海淀区中关村东路 55 号 邮政编码：100190.
Address: China Academy of Sciences Fundamental Science Park, No.55, Zhong-guancun East Road, Haidian Beijing China, 100190.
Closest subway station: Zhichunli (line 10).

5   Our sponsors

- Academy of Mathematics and Systems Science, Chinese Academy of Sciences
- CAS AMSS-PolyU Joint Laboratory of Applied Mathematics
- National Key R&D Program of China
- National Natural Science Foundation of China
- State Key Laboratory of Mathematical Sciences, AMSS, CAS

6   Book of abstract

Block-Wise Mixed Precision Sparse Solvers

Speaker: Weifeng Liu
Abstract: Sparse direct and iterative solvers play a crucial role in scientific computing and engineering applications. Mixed precision computing, which integrates double, single, and half precision arithmetic, provides an effective approach to reducing computational costs while hopefully preserving solution accuracy. This talk will present our two recent advancements in mixed precision-enhanced open-source solvers: the sparse direct solver PanguLU (Fu et al., SC '23)[1] and the sparse iterative solver Mille-feuille (Yang et al., SC '24)[2]. Both solvers employ block-wise mixed precision strategies on GPUs, and exhibit obvious performance improvements over their original FP64-precision counterparts.

Recent Developments in Krylov Subspace Methods for Nonsymmetric Shifted Linear Systems

Speaker: Tomohiro Sogabe
Abstract: Shifted linear systems of the form $(A+s_kI)x_k = b$ for $k = 1, 2, \ldots, n$ frequently arise in quantum chromodynamics and large-scale electronic structure calculations. These linear systems also appear in subproblems of eigenvalue problems and matrix function computations. In this talk, we briefly review Krylov subspace methods for solving nonsymmetric linear systems, along with our recent solver that is joint work with R. Zhao.

Surrogate hyperplane Bregman–Kaczmarz methods for solving linear inverse problems

Speaker: Junfeng Yin

---

[1]https://dl.acm.org/doi/10.1145/3581784.3607050
[2]https://ieeexplore.ieee.org/abstract/document/10793204

Abstract: Linear inverse problems arise in many practical applications, for instance, compressed sensing, phase retrieval and regularized basis pursuit problem. We propose a residual-based surrogate hyperplane Bregman-Kaczmarz method (RSHBK) for solving this kind of problems. The convergence theory of the proposed method is investigated in details. When the data is contaminated by the independent noise, an adaptive version of our RSHBK method is developed. An adaptive relaxation parameter is derived for optimizing the bound on the expectation error. We demonstrate the efficiency of our proposed methods for both noise-free and independent noise problems by comparing with other state-of-the-art Kaczmarz methods in terms of computation time and convergence rate through synthetic experiments and real-world applications.

A journey through some numerical linear algebra algorithms with variable accuracy

Speaker: Luc Giraud
Abstract: In this talk, we will explore the numerical behavior of widely used numerical linear algebra methods when the errors introduced by the underlying hardware and working arithmetic are decoupled from those associated with the data representation of mathematical objects—primarily matrices and vectors—computed by the algorithms, a concept we refer to as variable accuracy.

We will present experimental results in fundamental contexts, including basis orthogonalization using Modified Gram-Schmidt variants, the solution of linear systems with GMRES, and eigenproblem solutions via the FEAST method.

Additionally, we will discuss the numerical quality of the computed results under variable accuracy and draw connections with well-established results in finite precision arithmetic.

Finally, we will showcase applications in low-rank tensor computations using the Tensor Train format, along with ongoing perturbation analysis to justify the observed experimental behavior.

Exploiting low-precision arithmetic in eigensolvers

Speaker: Meiyue Shao
Abstract: In recent years mixed-precision algorithms have received increasing attention in numerical linear algebra and high performance computing. Modern mixed-precision algorithms perform a significant amount of low-precision arithmetic in order to speed up the computation, while still providing the desired solution in working precision. A number of existing works in the literature focus on mixed-precision linear solvers—much less is known about how to improve eigensolvers. In this talk we discuss several techniques that can accelerate eigenvalue computations by exploiting low-precision arithmetic. These techniques lead to several mixed-precision symmetric eigensolvers, for both dense and sparse eigenvalue problems. Our mixed-precision algorithms outperform existing fixed-precision algorithms without compromising the accuracy of the final solution.

Learning-based Multilevel Solvers for Large-Scale Linear Systems

Speaker: Chensong Zhang

Abstract: This talk presents our efforts on developing learning-based solvers for large-scale linear systems arising from discretized PDEs. Our approach bridges traditional multilevel solvers with machine learning, automating solver design to enhance efficiency and scalability. The method generalizes across grid sizes, coefficients, and right-hand-side terms, enabling offline training and efficient generalization, with convolutional neural networks (CNNs) serving as the basic computational kernels. It utilizes multilevel hierarchy for rapid convergence and cross-level weight sharing to adapt flexibly to varying grid sizes. The proposed solver achieves speedup over classical geometric multigrid methods for convection-diffusion PDEs in preliminary numerical experiments. We further extend this framework by introducing a Fourier neural network (FNN) to accelerate source influence propagation in Helmholtz equations within heterogeneous media. Supervised experiments demonstrate superior accuracy and efficiency compared to other neural operators, while unsupervised scalability tests reveal significant speedups over other AI solvers, achieving near-optimal convergence for wave numbers up to $k \approx 2000$. Ablation studies validate the effectiveness of the multigrid hierarchy and the novel FNN architecture.

Mixed-precision HODLR matrices

Speaker: Xinye Chen

Abstract: Hierarchical matrix computations have attracted significant attention in the science and engineering community as exploiting data-sparse structures can significantly reduce the computational complexity of many important kernels. One particularly popular option within this class is the Hierarchical Off-Diagonal Low-Rank (HODLR) format. This talk demonstrates that off-diagonal low-rank blocks in HODLR matrices can be stored in reduced precision without degrading overall accuracy. An adaptive-precision scheme ensures numerical stability in key computations. We provide theoretical insights to guide precision selection, with experiments confirming effectiveness. An emulation software that facilitates the research on HODLR matrix computations is also presented.

Computing accurate eigenvalues using a mixed-precision Jacobi algorithm

Speaker: Françoise Tisseur

Abstract: Efforts on developing mixed precision algorithms in the numerical linear algebra and high performance computing communities have mainly focussed on linear systems and least squares problems. Eigenvalue problems are considerably more challenging to solve and have a larger solution space that cannot be computed in a finite number of steps

In this talk we present a mixed-precision preconditioned Jacobi algorithm, which uses low precision to compute the preconditioner, applies it in high precision (amounting to two matrix-matrix multiplications) and solves the eigenproblem using the Jacobi algorithm in working precision. Our analysis yields meaningfully smaller relative forward error bounds for the computed eigenvalues

compared with those of the standard Jacobi algorithm. We further prove that, after preconditioning, if the off-diagonal entries of the preconditioned matrix are sufficiently small relative to its smallest diagonal entry, the relative forward error bound is independent of the condition number of the original matrix. We present two constructions for the preconditioner that exploit low precision, along with their error analyses. Our numerical experiments confirm our theoretical results.

Stability of modified Gram-Schmidt and its low synchronization variants

Speaker: Qinmeng Zou

Abstract: Modified Gram-Schmidt (MGS) is a traditional QR factorization process that are widely used in solving linear systems and least squares problems. For example, the generalized minimal residual method (GMRES) usually employs QR factorization to orthogonalize the basis of Krylov subspace. This talk discusses a class of low-synchronization MGS algorithms, denoted as MGS-LTS, which can date back to Bjorck's work in 1967. We give a stability analysis of MGS-LTS, proving that the loss of orthogonality of its basic form as well as the block and normalization lagging variants is proportional to the condition number. We also discuss the probabilistic tools and investigate the causes of instability of the Cholesky-based block variant. Finally, we show numerical experiments for the mixed-precision low-synchronization GMRES algorithm with sparse approximate inverse preconditioning.

Towards trustworthy use of scientific machine-learning in large scale numerical simulations

Speaker: Alena Kopanicakova

Abstract: Recently, scientific machine learning (SciML) has expanded the capabilities of traditional numerical approaches by simplifying computational modeling and providing cost-effective surrogates. However, SciML models suffer from the absence of explicit error control, a computationally intensive training phase, and a lack of reliability in practice. In this talk, we will take the first steps toward addressing these challenges by exploring two different research directions. Firstly, we will demonstrate how the use of advanced numerical methods, such as multilevel and domain-decomposition solution strategies, can contribute to efficient training and produce more accurate SciML models. Secondly, we propose to hybridize SciML models with state-of-the-art numerical solution strategies. This approach will allow us to take advantage of the accuracy and reliability of standard numerical methods while harnessing the efficiency of SciML. The effectiveness of the proposed training and hybridization strategies will be demonstrated by means of several numerical experiments, encompassing the training of DeepONets and the solving of linear systems of equations that arise from the high-fidelity discretization of linear parametric PDEs using DeepONet-enhanced multilevel and domain-decomposition methods.

Solving 10,000-Dimensional Optimization Problems Using Inaccurate Function Values: An Old Algorithm

Speaker: Zaikun Zhang

Abstract: We reintroduce a derivative-free subspace optimization framework originating from Chapter 5 of [Z. Zhang, On Derivative-Free Optimization Methods, PhD thesis, Chinese Academy of Sciences, Beijing, 2012 (supervisor Yaxiang Yuan)]. At each iteration, the framework defines a low-dimensional subspace based on an approximate gradient, and then solves a subproblem in this subspace to generate a new iterate. We sketch the global convergence and worst-case complexity analysis of the framework, elaborate on its implementation, and present some numerical results on problems with dimensions as high as 10,000.

The same framework was presented by Zhang during ICCOPT 2013 in Lisbon under the title "A Derivative-Free Optimization Algorithm with Low-Dimensional Subspace Techniques for Large-Scale Problems", although it remained nearly unknown to the community until very recently. An algorithm following this framework named NEWUOAs was implemented by Zhang in MATLAB in 2011[3], ported to Modula-3 in 2016 by M. Nystroem, a Principal Engineer at the Intel Corporation, and made available in the open-source package CM3[4]. NEWUOAs has been used by Intel in the design of chips, including its flagship product Atom P5900.

More information about the framework can be found in the paper: https://arxiv.org/abs/2501.04536.

A Fast BB Reduced Minimization Algorithm for Nonnegative Viscosity Optimization in Optimal Damping

Speaker: Qingna Li

Abstract: We consider the fast optimization algorithm for optimal viscosities in damping system. Different from standard models that minimize the trace of the solution of parameterized Lyapunov equation, the nonnegative constraints for viscosities are added in the optimization model, which hasn't been considered before. To solve the new model, an efficient algorithm is then proposed, aiming at reducing the residuals of the corresponding KKT conditions. By combining with the Barzilai-Borwein stepsize, the proposed BB residual minimization algorithm (short for BBRMA) can further speed up to deal with large scale linear vibration systems. Extensive numerical results verify the high efficiency of the proposed algorithm. This is the joint work with Françoise Tisseur.

Error Analysis of Matrix Multiplication with Narrow Range Floating-Point Arithmetic

Speaker: Mantas Mikaitis

Abstract: High-performance computing hardware now supports many different floating-point formats, from 64 bits to only 4 bits. While the effects of reducing precision in numerical linear algebra computations have been extensively studied, some of these low precision formats also possess a very narrow

---

[3]https://github.com/newuoas/newuoas
[4]https://github.com/modula3/cm3/blob/master/caltech-other/newuoa/src/NewUOAs.m3

range of representable values, meaning underflow and overflow are very likely. The goal of this article is to analyze the consequences of this narrow range on the accuracy of matrix multiplication. We describe a simple scaling that can prevent overflow while minimizing underflow. We carry out an error analysis to bound the underflow errors and show that they should remain dominated by the rounding errors in most practical scenarios. We also show that this conclusion remains true when multiword arithmetic is used. We perform extensive numerical experiments that confirm that the narrow range of low precision arithmetics should not significantly affect the accuracy of matrix multiplication—provided a suitable scaling is used.

Analysis of Floating-Point Matrix Products Computed via Low-Precision Integer Arithmetic

Speaker: Massimiliano Fasi
Abstract: Matrix multiplication is arguably a fundamental kernel in scientific computing, and efficient implementations underpin the performance of many numerical algorithms for linear algebra. The Ozaki scheme is a method that computes matrix – matrix products by recasting them as sequences of error-free transformations. First developed in 2008 in the context of summation, this technique has recently seen a resurgence of interest, because it is particularly well suited to the mixed-precision matrix-multiplication units available on modern hardware accelerators (such as GPUs, TPUs, or NPUs). The latest generations of these accelerators are particularly efficient at computing products between matrices of low-precision integers. In scientific application, integers are typically not sufficient, but, in the last couple of years, variants of the Ozaki scheme that rewrite *floating-point* matrix multiplications in terms of *integer* matrix products have been proposed. We analyse the error incurred by these integer variants of the Ozaki scheme, and we characterise cases in which these methods can fail.