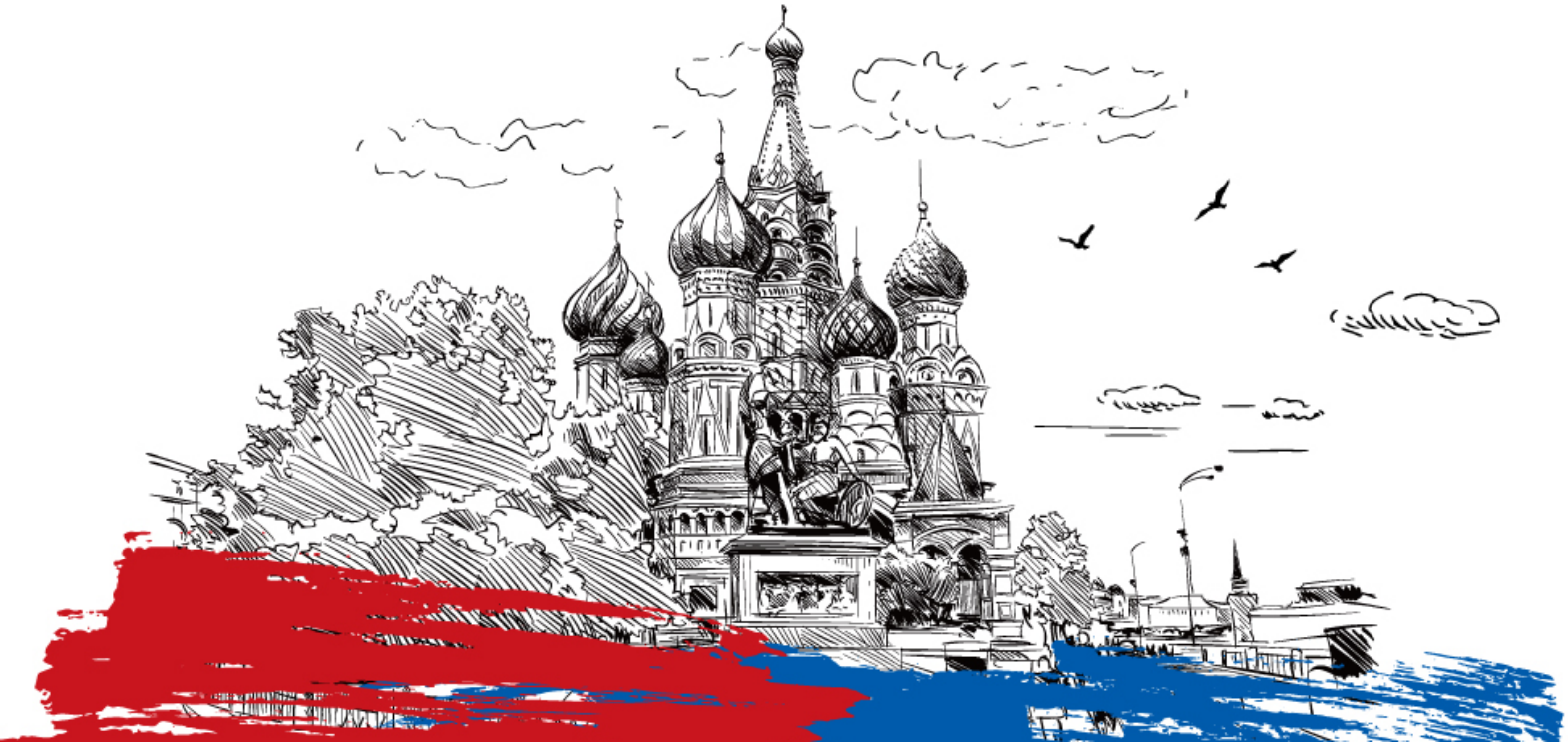
Skoltech
Skolkovo Institute of Science and Technology

HUAWEI

The 5th International Conference on Matrix Methods in Mathematics and applications (MMMA-2019)

19th - 23th Aug 2019 | Moscow, Russia



Conference organizers:

This is a book of abstracts of the 5-th international conference on Matrix Methods in Mathematics and Applications. Conference website is `www.deeptensor.ml/mmma2019`. Conference is hosted by Skolkovo Institute of Science and Technology and by Marchuk Institute of Numerical Mathematics, RAS. We are grateful to Huawei Technologies, our industrial co-organizer in 2019.

Program committee: Eugene Tyrtyshnikov (INM RAS), Ivan Oseledets (Skoltech), Vladimir Lyashev (Huawei).
Organizing committee: Sergey Matveev, Ekaterina Loginova, Alina Sitnikova, Maria Lazenina, Sergey Monin, Maxim Kuznetsov.

Technical editor of this book is Sergey Matveev.

Contents

Rational approximations to fractional powers of operators

Lidia Aceto

Department of Mathematics - University of Pisa (Italy)
Co-author: Paolo Novati - University of Trieste , Italy

In this talk we present rational approximations of fractional powers of self-adjoint positive operators attainable with a specific integral representation of the operator function. We provide accurate estimates of the truncation errors by exploiting classical results in Approximation Theory involving the Pade approximants. Finally, we report some numerical results to illustrate the effectiveness of the presented analysis.

Direct numerical simulation of multiphase flows at pore scale

Vladislav Balashov

Keldysh Institute of Applied Mathematics RAS
Co-author: Evgeny Savenkov, KIAM

Analysis of the multiphase flows in different natural and artificial porous materials is relevant and important problem. With the development of micro-CT technologies which allows to construct geometrical models of material microstructure, it becomes possible to analyze such flows numericaly.

We describe the computational framework for an analyses of multiphase fluid flows in the pore space. The approach is based on quasi-hydrodynamic regularization of Navier-Stokes-Cahn-Hilliard model which uses diffuse interface description of the inter-phase boundaries. The developed algorithms allows for highly parallel implementation. Description of the simulation domain is based on voxel representation wich allows to use micro-CT data with minimal preprocessing.

A number of simulations demonstrating consistency of the model and algorithms as well as “realistic“ flow simulations within realistic micro-CT models are presented. Additional attention is paid to the analysis of numerical algorithms with improved properties which minimize influence of the so-called “parasitic currents“.

ART: Adaptive residual-time restarting for Krylov subspace matrix exponential evaluations

Mike Botchev

Keldysh Institute of Applied Mathematics, RAS; Marchuk Institute of Numerical Mathematics, RAS
Co-author: Leonid Knizhnerman, Mathematical Modelling Department, Central Geophysical Expedition, Moscow, Russia

In this talk a new restarting method for Krylov subspace matrix exponential evaluations is discussed. Since our restarting technique essentially employs the residual, some convergence results for the residual are presented. We also show how the restart length can be adjusted after each restart cycle, which leads to an adaptive restarting procedure. Numerical tests are presented to compare our restarting with three other restarting methods. Some of the algorithms described in

this talk are a part of the Octave/Matlab package expmARPACK available at http://team.kiam.ru/botchev/expm/.

Globally Optimal Continuous Control with a Sparse Reward using Tensor Train

Alexey Boyko

Skoltech
Co-author: Ivan Oseledets, Skoltech

One of the important problems of Artificial Intelligence is finding best policy of taking actions for general complex systems. The general globally optimal solution for this problem is given by solving Bellman Equation. Due to its exponential curse of dimensionality, it is typically solved by either manual simplification and reward shaping (by control engineers) or by machine learning techniques (by reinforcement learning community). Ideally, an algorithm should be able to figure out an optimal policy for achieving a goal without manual tuning by human. Such a problem of sequential decision making (or control) is called Sparse Reward problem. For this kind of reward, classical control methods as well as reinforcement learning methods often fail to reliably provide an optimal policy. We study an alternative approach originally proposed by Gorodetsky that solves the unmodified Bellman Equation for a continuous system from the first principles, by applying Tensor Train low-rank decomposition. Our experiments demonstrate full capability to handle sparse rewards

GP for tensor completion

Evgeny Burnaev

Skoltech
Co-author: Ermek Kapushev

Methods based on GP for tensor completion.

Impurity in a quantum gas: exact diagonalization in the Bethe Ansatz basis.

Evgeni Burovski

Higher School of Economics, Moscow, Russia

We examine stationary state properties of an impurity particle injected into a one-dimensional quantum gas. The terminal velocity of an impurity turns out to be non-zero. For equal masses of the impurity and host particles, the

problem reduces to a variant of the Gaudin-Young model, which is integrable and admits a formal solution based on the Bethe Ansatz. Obtaining physical observables from this formal solution is still non-trivial, as it involves summations of infinite series of form-factors, represented by rank-N determinants for N particles. We develop a controllable numerical method for performing summations of form-factor expansions, which employs stochastic enumeration based on the Metropolis algorithm.

For unequal masses, the problem is no longer integrable, and Bethe Ansatz approach breaks down. To study effects due to integrability-breaking perturbations, we develop an exact diagonalization procedure in the basis of the Bethe Ansatz states for an integrable model.

Time integration of (anti-)symmetric Tucker tensors.

Gianluca Ceruti

University of Tuebingen

Co-author: Christian Lubich, University of Tuebingen.

A numerical integrator is presented that computes a symmetric or anti-symmetric low-rank approximation to large symmetric or anti-symmetric time-dependent Tucker tensors that are either given explicitly or are the unknown solution to a matrix/tensor differential equation. We show that low-rank time-dependent matrices and tensors are reproduced exactly, and the error behaviour is robust to the presence of small singular values of the solution or the approximate solution.

A Nuclear-Norm Model for Multi-Frame Super-Resolution Reconstruction from Video Clips

Raymond Chan

City University of Hong Kong

Co-author: Raymond Chan

Co-author: Hongfei Yang

We present a variational approach to obtain super-resolution images from multiple low-resolution frames extracted from video clips. First the displacement between the low-resolution frames and the reference frame is computed by an optical flow algorithm. Then a low-rank model is used to construct the reference frame in high resolution by incorporating the information of the low-resolution frames. The model has two terms: a 2-norm data fidelity term and a nuclear-norm regularization term. Alternating direction method of multipliers is used to solve the model. Comparison of our methods with other models on synthetic and real video clips shows that our resulting images are more accurate with less artifacts. It also provides much finer and discernable details.

Predicting dynamical system evolution with residual neural networks

Artem Chashchin

Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences, Skolkovo Institute of Science and Technology

Co-author: Mikhail Botchev, Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences

Co-author: Ivan Oseledets, Skolkovo Institute of Science and Technology, Institute of Numerical Mathematics of the Russian Academy of Sciences

Co-author: George Ovchinnikov, Skolkovo Institute of Science and Technology

Forecasting time series and time-dependent data is a common problem in many applications. One typical example is solving ordinary differential equation (ODE) systems x'=F(x). Oftentimes the right hand side function F(x) is not known explicitly and the ODE system is described by solution samples taken at some time points. Hence, ODE solvers cannot be used. In this paper, a data-driven approach to learning the evolution of dynamical systems is considered. We show how by training neural networks with ResNet-like architecture on the solution samples, models can be developed to predict the ODE system solution further in time. By evaluating the proposed approaches on three test ODE systems, we demonstrate that the neural network models are able to reproduce the main dynamics of the systems qualitatively well. Moreover, the predicted solution remains stable for much longer times than for other currently known models.

Matrix/tensor completetions and randomization algorithms algorithms for tensor train and tensor rings

Andrzej Cichocki

SKOLTECH

Co-author: Cichocki Andrzej

In this talk we discuss recent trends and progress in matrix/tensor completions and related topics like recommender systems, using tensor networks, especially tensor train and tensor rings architectures. The tensor train and tensor rings frameworks can relatively easily accommodate Total Variation or Tikhonov regularization and also other more sophisticated regularization due to their low-rank representations. Image and video inpainting experiments indicate the superiority of the recent completion schemes in terms of performance and scalability, Moreover, we demonstrate that the developed schemes are especially advantageous when only tiny portions (say, 1-2images/videos are known. Moreover, a new approach is presented for very large-scale low-rank tensor decomposition using randomized techniques.

On maximum volume submatrices and cross approximation

Alice Cortinovis

EPF Lausanne

Co-author: Daniel Kressner, EPF Lausanne

Co-author: Stefano Massei, EPF Lausanne

The problem of finding the k × k submatrix of maximum volume is connected to rank-k approximations of A constructed from k rows and k columns of A. We analyze the low-rank approximation error returned by a greedy method for volume maximization, cross approximation with complete pivoting. We obtain an error bound for general matrices which extends an existing result for symmetric positive definite matrices and yields new bounds for diagonally dominant matrices. We also get an improved error bound for cross approximation of functions.

Reduced-Order Modeling of Deep Neural Networks

Talgat Daulbaev

Skoltech

Co-author: Julia Gusak, Skoltech
Co-author: Evgeny Ponomarev, Skoltech
Co-author: Andrzej Cichocki, Skoltech
Co-author: Ivan Oseledets, Skoltech

We introduce a new method for speeding up the inference of deep neural networks. It is somewhat inspired by the reduced-order modeling techniques for dynamical systems. The cornerstone of the proposed method is the maximum volume algorithm. We demonstrate efficiency on VGG and ResNet architectures pre-trained on different datasets. We show that in many practical cases it is possible to replace convolutional layers with much smaller fully-connected layers with a relatively small drop in accuracy.

Householder-type matrix algebras in displacement decompositions

Carmine Di Fiore

Department of Mathematics, University of Rome "Tor Vergata"

Co-author: Piero Deidda, Department of Mathematics, University of Rome "Tor Vergata"

Projections on SDU algebras of matrices simultaneously diagonalized by an Householder matrix
$U = I - 2uu^h$ ($u \in \mathbb{C}^n$, $||u|| = 1$) or by a product U of a small number of Householder matrices, have been recently shown to be competitive both as preconditioners of structured linear systems [2],[3], and as low complexity Hessian approximations in quasi-Newton minimization iterative schemes [1],[3],[4].
In principle, matrices in such Householder SDU algebras could be also used as bricks in displacement-rank matrix decompositions [5], exactly as it happens for the better known SDU algebras diagonalized by matrices U of Fourier, Jacobi and Hartley type [6],[7],[8],[9],[10].
The investigation of this idea, commenced in [11], has yielded some new displacement formulas defined in terms of pairs of SDU algebras where U is a generic unitary matrix, one including a rank-1 perturbation of a matrix in the other one. Moreover, it has leaded us to face significant problems of numerical linear algebra whose study involves in a natural way unitary matrices of the form $I - \alpha uu^h$, $|\alpha|^2 = 2Re(\alpha)$, which are a simple extension of Householder matrices that make them more suitable when working in the complex field.

References

[1] S. Cipolla, C. Di Fiore, F. Tudisco, P. Zellini: *Adaptive Matrix Algebras in Unconstrained Minimization*, Linear Algebra and its Applications, 471, pp.544-568 (2015).

[2] S. Cipolla, C. Di Fiore, F. Tudisco: *Euler-Richardson method preconditioned by weakly stochastic matrix algebras: a potential contribution to Pagerank computation*, Electronic Journal of Linear Algebra, 32, pp.254-272 (2017).

[3] S. Cipolla, C. Di Fiore, P. Zellini: *Low complexity matrix projections preserving actions on vectors*, Calcolo, June 2019, 56:8 (2019).

[4] S. Cipolla, C. Di Fiore, P. Zellini: *A variation of Broyden Class methods using Householder adaptive transforms*, submitted, August 2017 (https://arxiv.org/abs/1906.00034).

[5] T.Kaylath, H.Sayed: *Displacement structure: Theory and applications*, SIAM Review, 37, pp. 297-386 (1995).

[6] I. Gohberg, V. Olshevsky: *Circulants, displacements and decompositions of matrices*, Integral Equations and Operator Theory, 15, pp 730-743 (1992).

[7] D. Bini, V. Pan: *Polynomial and Matrix Computations*, Vol.1 Fundamental Algorithms, Birkhäuser, Boston (1994).

[8] C. Di Fiore, P.Zellini: *Matrix Decompositions Using Displacement Rank and Classes of Commutative Matrix Algebras*, Linear Algebra and its Applications, 229, pp 49-99 (1995).

[9] E. Bozzo, C.Di Fiore: *On the use of certain matrix algebras associated with discrete trigonometric transforms in matrix displacement decomposition*, SIAM J. Matrix Anal. Appl. , 16, pp 312-326 (1995).

[10] C. Di Fiore: *Matrix Algebras and Displacement Decompositions*, SIAM J. Matrix Anal. Appl., 21, pp 646-667 (2000).

[11] P. Deidda: *Householder-type matrix algebras in displacement decompositions*, Master Thesis in Mathematics (supervisor C. Di Fiore), University of Rome "Tor Vergata", May 22, 2019
(https://www.mat.uniroma2.it/∼difiore/TesiPiero.pdf).

Low-rank tensor decompositions for sampling of high-dimensional probability distributions

Sergey Dolgov

University of Bath

Co-author: Karim Anaya-Izquierdo (U. Bath)
Co-author: Colin Fox (U. Otago)
Co-author: Robert Scheichl (U. Heidelberg)

Uncertainty quantification and inverse problems in many variables are pressingly needed tasks, yet high-dimensional functions are notoriously difficult to integrate in order to compute desired quantities of interest. Functional approximations, in particular the low-rank separation of variables into tensor product decompositions, have become popular for reducing the computational cost of high-dimensional integration down to linear scaling in the number of variables. However, tensor approximations may be inefficient for non-smooth functions. Sampling based Monte Carlo methods are more general, but they may exhibit a very slow convergence, overlooking a hidden structure of the function.

In this talk we review tensor product approximations for the problem of uncertainty quantification and Bayesian inference. This allows efficient integration of smooth PDE solutions, posterior density functions and quantities of interest. Moreover, we can use the low-rank approximation of the density function to construct efficient proposals in the MCMC algorithm, as well as in the importance weighting.

This combined tensor approximation - MCMC method is more accurate also if the quantity of interest is not smooth, such as the indicator function of an event.

Embedding properties of network realizations of reduced order models with applications to inverse scattering and data science

Vladimir Druskin

Worcester Poitechnic Institute
Co-author: Lilliana Borcea, University of Michigan, AnnArbor
Co-author: Alexander Mamonov, University of Houston
Co-author: Mikhail Zaslavsky, Schlumberger-Doll Research

Continued fractions are known since antiquity as the most compact representations of numbers. At the end of the 19th century Stieltjes connected them with physics. This connection gave rise to network syntheses in the first half of the 20th century that was at the base of modern electronics design and consecutively to model order reduction (MOR) that tremendously impacted many areas of engineering by enabling efficient compression of the underlining dynamical systems. In his seminal 1950s works Krein realized that in addition to their compressing properties, Stieltjes continuous fractions can be used to embed the data back into the state space of the underlying dynamical system via special mechanical networks known as Stieltjes strings. Such networks can learn the underlying PDE system from the data (transfer function) via rigorously chosen hyper-parameters. Among many application of this powerful approach we discuss the following two. 1. Imaging in strongly scattering media with waves (e.g.,seismic exploration) via data-driven ROMs. 2. Reduced order graph-Laplacians and efficient cluster analysis of big data sets.

Perturbation theory for the canonical polyadic decomposition

Eric Evert

KU Leuven
Co-author: Lieven De Lathauwer, KU Leuven

This talk gives perturbation theoretic bounds for the canonical polyadic decom- position (CPD) of a low rank tensor. In practice, one works with a measurement of a low rank tensor which is corrupted by additive noise and computes a low rank approximation of the measured tensor. Our results give an upper bound for the error between the computed CPD and the CPD for the original tensor. Additionally, we use these perturbation theoretic results to provide guarantees for the existence of a best low rank approximation of a measured tensor.

Ergodicity coefficients for stochastic higher-order tensors

Dario Fasino

University of Udine, Italy
Co-author: Francesco Tudisco (University of Edinburgh, UK)

Coefficients of ergodicity for stochastic matrices provide easily computable conditions to ensure the uniqueness of the Perron eigenvector and the ergodicity of the associated Markov chain.

Application of Style-GAN for microstructure synthesis

Daria Fokina

Skoltech CDISE
Co-author: Ivan Oseledets, Skoltech

The work considers the problem of microstructure synthesis. The task is the following: given several crops of the structure, a larger sample is produced. It is proposed to use recently developed style-based GAN architecture for this task. This architecture should be quite efficient for microstructures, as it uses images at different resolutions. However, it has a drawback: it generates only fixed-size samples. It is suggested to synthesize several examples and connect them via image quilting procedure to increase the size of the resulting structure. The effectiveness of the proposed method is shown for several microstructures.

Why to find out best low-rank TT approximation with L1 norm as well

Alain Franc

INRA BioGeCo & INRIA Pleiade, Bordeaux, France
Co-author: Coulaud, Olivier, INRIA HiePacs, Bordeaux, France
Co-author: Peyrard, Nathalie, INRA MIAT, Toulouse, France

Let Z be the partition function of distribution function of a system of many interacting particles in statistical physics. Its exact calculation is untractable most of the time. A standard approximate solution is to build the best approximation of a given tensor by a tractable TT, compute Z of the TT, and hope it is a good approximation of the partition function of the exact tensor as well. As the tensor under study is a joint law, all its coefficients are non-negative. Then, the partition function simply is the L1 norm of the tensor. It is known that, in Banach spaces, L1 and L2 norms are not equivalent, i.e. there is no constant bounding from above the L1 norm for a vector with unit L2 norm. We consider the sphere in of tensors of d modes with dimension n on each mode. We show that the expectation of the L1 norm of such tensors is in . This shows that even in the case where the low rank TT-approximation of a tensor A is excellent, it may happen that the approximation of Z(A) by considering the partition function of the approximation is poor by considering the L1 and L2 norms of the difference between the tensor A and its TT-approximation. This leads to the necessity of considering the best low rank TT approximation of a given tensor A with L1 norm, which is known to be a difficult, and to our knowledge still open, problem.

A straightforward generalization of low rank approximation approach for hybrid recommender systems

Evgeny Frolov

Skoltech
Co-author: Ivan Oseledets, Skoltech

We propose a new hybrid approach for matrix- and tensor-based recommender systems. Unlike the majority of hybrid recommenders, it directly ties collaborative user behavior with additional side information in an intuitive and straightforward way. It not only helps to address the problem of extreme data sparsity, but also allows to naturally exploit patterns in the observed interactions for constructing a compact and meaningful representation of user intents. We demonstrate the effectiveness of the proposed model on several standard benchmark datasets. The general formulation of the approach imposes no restrictions on the type of observed interactions and makes it potentially applicable for joint modelling of any type of contextual information along with side data.

New Algorithms for Solving Nonlinear Eigenvalue Problems

Walter Gander

ETH and HKBU

To solve a nonlinear eigenvalue problem we develop algorithms which compute zeros of $\det A(\lambda) = 0$. We show how to apply third order iteration methods for that purpose. The necessary derivatives of the determinant are computed by algorithmic differentiation. Since many nonlinear eigenvalue problems have banded matrices we also present an algorithm which makes use of their structure.

Different approaches for optimal multiband filter synthesis

Sergei Goreinov

Marchuk Institute of Numerical Math. RAS

A number of approaches to the synthesis of electrical filters is considered, including numerical one which uses Kolmogorov's (linear) optimality criterion for C-norm approximation, and several analytic approaches, working only for "simplified" formulations but in more complex cases e.g. many pass- and stopbands, narrow transition bands, high attenuation at stop bands etc. Comparison of these approaches is given.

Iterative tensor decomposition for neural network compression

Julia Gusak

Skoltech
Co-author: Maxim Kholyavchenko (Innopolis)
Co-author: Evgeny Ponomarev (Skoltech)
Co-author: Andrzej Cichocki (Skoltech)
Co-author: Ivan Oseledets (Skoltech)

The low-rank tensor approximation is very promising for the compression of deep neural networks. We propose a new simple and efficient iterative approach, which alternates low-rank factorization with a smart rank selection and fine-tuning. We demonstrate the efficiency of our method comparing to non-iterative ones. Our approach improves the compression rate while maintaining the accuracy for a variety of tasks.

On the parallel two-level multi-preconditioned iterative methods in Krylov subspace

Valery P. Ilin

Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk State University

We consider different parallel versions of the two level least squares methods in the multi-preconditioned Krylov subspaces based on computing various basis vectors. These algorithms are used for solving very large real, non-symmetric, in general, sparse systems of linear algebraic equations (SLAEs) which arise I grid approximations of multi-dimensional boundary value problems. In particular, the Chebyshev acceleration approaches, steepest descent and minimal residual, conjugate residual, as well as multi-preconditioned semi-conjugate direction methods applied as preliminary iterative processes. The resulting minimization of residuals is provided by the block, or implicit procedures. The properties of the Krylov approaches proposed are analyzed with various types of preconditioning, incomplete factorization, domain decomposition, algebraic multi-grid and coarse grid correction including. The main criteria of scalable paralleiezation based on the hybrid programming are estimated. The convergence rate and stability of the algorithms are demonstrated on the results of numerical experiments for the model SLAEs which present the exponential fitting approximation of diffusion-convection equations with various mesh steps and with different coefficients.

Gaussian Processes for Tensor Completion

Yermek Kapushev

Skoltech
Co-author: Evgeny Burnaev, Skoltech
Co-author: Ivan Oseledets, Skoltech

In this paper we consider the tensor completion problem in the tensor train (TT) format. It is assumed that tensor is high-dimensional and tensor values are generated by an unknown smooth function of its indices. The assumption allows us to develop an efficient initialization scheme based on Gaussian Process Regression and TT-cross approximation technique. The proposed approach can be used in conjunction with any optimization algorithm that is usually utilized in tensor completion problems. We empirically justify that in this case the reconstruction error improves compared to the tensor completion with random initialization. As an additional benefit, our technique automatically selects rank during TT-cross approximation.

Finding of maximum volume square and rectangular submatrices in block matrix case.

Gleb Karpov

Skoltech

Co-author: Gleb Ryzhakov, Skoltech

Co-author: Ivan Oseledets, Skoltech; Marchuk Institute of Numerical Mathematics, RAS.

This work is devoted to finding submatrices with certain extreme property, so-called D-optimality criterion, which is extremely useful in a variety of tasks, such as recommender systems [5] or wireless communication [4]. We expand the criterion to the block matrix case (when one physical entity corresponds to several rows, which are connected together, and this entity cannot be chosen partially: it is either fully selected, or not selected at all). The proposed approach allows finding square and rectangular block D-optimal submatrices of a matrix by greedy updates and extensions of the initial square submatrix. Algorithm for selecting the last one is also proposed and based on PLUQ decomposition of a matrix. This work originates from the results of [1] and [3], which are concerned with finding D-optimal submatrices in a standard case, and appears to be their natural expansion to the block case.

As a practical application, we consider selection nodes for multivariate function approximation with use of its derivatives. One way to approximate a function is to construct a polynomial expansion. Coefficients of the former can be found with the Least Squares Method (LSM). Values of derivatives give additional equations by extending the matrix and right-hand side of LSM, which is the case of utilizing block matrices. Thus, we reduce the number of points where to evaluate the function. Based on the notion that using D-optimal submatrices in the LSM has a positive effect on the approximation quality (see [2] in a square case, minimizing variance in overdetermined case) we use proposed algorithm to select the proper nodes. Accuracy of approximation is compared with well-known samplings such as Latin Hypercubes Sampling (LHS), Sobol' and random.

References: [1] Goreinov, S. A., Savostyanov, D., Tyrtyshnikov, E. E., Zamarashkin, N.L.:How to Find a Good Submatrix*, Matrix Methods: Theory, Algorithmsand Applications, 247-256., 2010. [2] Goreinov, S.A., Tyrtyshnikov, E.E. :The maximal-volume concept in ap-proximation by low-rank matrices., Contemporary Mathematics, 280, 47-51, 2001. [3] Mikhalev, A., Oseledets, I. V. :Rectangular maximum-volume submatricesand their applications., Linear Algebra and Its Applications, 538, 187-211,2018. [4] Hong Wang, Bu & Hui, Hon & Seng Leong, Mook. :Global and FastReceiver Antenna Selection for MIMO Systems., IEEE Transactions on Communications. 58. 2505-2510, 2010. [5] Nathan N. Liu, Xiangrui Meng, Chao Liu, and Qiang Yang. :Wisdom ofthe better few: cold start recommendation via representative based ratingelicitation., In Proceedings of the fifth ACM conference on Recommendersystems (RecSys 11). ACM, New York, NY, USA, 37-44, 2011.

Practical shift choice in the shift-and-invert Krylov subspace evaluations of the matrix exponential

Alexandr Katrutsa

Skolkovo Institute of Science and Technology

Co-author: Mike Botchev, Keldysh Institute of Applied Mathematics, Russian Academy of Sciences

Co-author: Ivan Oseledets, Skolkovo Institute of Science and Technology

We propose two methods to find a proper shift parameter in the shift-and-invert method for computing matrix exponential matrix-vector products. These approaches are useful in the case matrix exponential action has to be computed for a number of vectors. The first approach is based on the zero-order optimization of the mean residual norm for a given number of initial states. The second approach processes the vectors one-by-one and estimates, for each vector,

the derivative of the residual norm as a function of the shift parameter. The estimated derivative value is then used to update the shift parameter for the next vector. To demonstrate the performance of the proposed approaches we perform numerical experiments for two-dimensional non-stationary convection-diffusion problems with discontinuous coefficients. The experiments show that extra costs for optimization or derivative estimation pay off already for a small number of the initial vectors.

Resolving the structure of multiscale diffusion using low-rank tensors: approximation and preconditioning

Vladimir Kazeev

Stanford University

As a model problem, we consider a boundary-value problem for a linear second-order diffusion equation with high-frequency oscillations in the diffusion coefficient. At small scales, the multiscale structure of the problem renders it intractable for classical finite-element methods. Specifically, they require mesh refinement beyond the finest scale of the problem to achieve convergence, which means that the dimension of the discrete problem has to grow algebraically with respect to the scale parameter. One way to mitigate this is to retain a problem-nonspecific discretization but recast it in a high-dimensional space and apply a low-rank tensor representation to achieve the adaptive, data-driven compression of the solution in the course of its computation. This approach has been successfully applied to various types of problems but is contrary to special finite-element methods for multiscale problems, such as gFEM, MsFEM and HMM, which rely on constructing adaptive basis functions or quadrature rules by solving auxiliary problems. As in the previous works by Khoromskij and Repin (RJNAMM 2015) and Kazeev, Oseledets, Rakhuba and Schwab (ACOM 2017), we consider the low-rank tensor decomposition known as matrix product states (MPS) or tensor train (TT). For the aforementioned multiscale problem, with several scales and posed in two or three dimensions, we use the TT-MPS decomposition as a low-parametric representation for the discretizations recast in high-dimensional spaces. Building on the recent result of Bachmayr and Kazeev (arXiv 1802.09062) on preconditioning in the TT-MPS format, we develop a numerical multiscale solver that computes approximate solutions in low-rank decompositions whose number of parameters is polylogarithmic with respect to the scale parameter. We investigate, as an alternative to the direct tensor-structured approach, the approximation of the multiscale solution using a high-dimensional single-scale limit problem, developed by Hoang and Schwab (MMS 2005). Under additional assumptions, we theoretically analyze the approximability of the solution by low-rank approximations in the TT-MPS format and prove exponential convergence of such approximations. The talk is based on joint works with Markus Bachmayr, Ivan Oseledets, Maxim Rakhuba and Christoph Schwab.

Large Matrices in Superconductor Circuits Inductances Extraction

Mikhail Khapaev

Lomonosov Moscow State University, VMK, dep. of Numerical Methods

Co-author: Mikhail Kupriyanov, Lomonosov Moscow State University, Skobeltsyn Institute of Nuclear Physics

Future of high performance computing is most likely associated with one of alternative so-called Post-Moore

Technologies where energy dissipation is drastically lower. One of the most promising candidates is the superconductor digital technology. In contrast with semiconductor circuits where in simulation RC model circuits dominate,

superconductor circuits are based on inductances. In the report we consider the problem of extraction of inductances for superconductor circuits. Mathematical models different from those for normal conductors and numerical algorithms for this problem are discussed. Typically the problem needs solution of large dense matrix equations. We discuss applicability of direct approach and implementation of fast matrix methods.

Variational block term-decomposition as a structured mixing model

Pavel Kharyuk

Skoltech / INM RAS

Co-author: Ivan Oseledets, Skoltech / INM RAS

Block-term decomposition (BTD) was proposed by Lieven de Lathauwer in 2007, and later applied by the research community to different practical problems, to name a few: modelling electroencephalographic measurements of epileptic seizures, blind deconvolution of telecommunication signals, extracting atrial activity from electrocardiograms. In this decomposition input tensor is represented as a sum of structured terms (other tensors). Original BTD model assumes that these terms are consistent - represented in either (constrained) canonical or Tucker format, though examples of using non-consistent formats are also known. In our previous work simultaneous use of terms in different formats was explored in standard machine learning tasks, clustering and classification. This work contributes to extension of block-term decomposition to probabilistic mixing model, namely mixture of probabilistic PCAs with structured mixing matrices.

Tensor numerical methods in computational quantum chemistry

Venera Khoromskaya

Max-Planck Institute for Mathematics in the Sciences, Leipzig, Germany

Novel grid-based tensor numerical methods providing accuracy at the level of analytical computations are now used in grid-based solution of many problems in natural sciences. They are based on the accurate low-rank tensor-structured representation of d-variate functions and operators on large $n^{\otimes d}$ grids (in view of results from nonlinear approximation theory) providing $O(dn)$ complexity of numerical calculus [6].

A starting point was the tensor-based solution of the nonlinear three-dimensional integro-differential Hartree-Fock equation, which is a classical model for ab-initio calculation of the ground state energy of molecules [3,6]. In calculation of all involved operators, including 3D convolution integral operators, the 3D analytical integration is completely avoided, and substituted by the grid-based tensor algorithms in 1D complexity [1,3]. Moreover, the grid-based low-rank two-electron integrals tensor [2] enables economical computational scheme for the excitation energies and density of states for optical spectra of molecules in the framework of the Bethe-Salpeter equation [5,6].

Numerical modeling of the collective electrostatics for large many particle systems is a challenging problem in scientific computing. Our tensor method for summation of the long-range potentials on $L \times L \times L$ 3D lattices [6] provides computational cost of the order of $O(L)$ which outperforms the traditional $O(L^3)$-algorithms like Ewald-type summation. Recent range-separated (RS) tensor format [4] is an efficient tool for computation of the collective electrostatics for multiparticle systems of general type.

[1] B. N. Khoromskij and V. Khoromskaia. *Multigrid Tensor Approximation of Function Related Arrays.* SIAM J. Sci. Comput., **31**(4), 3002-3026 (2009).

[2] V. Khoromskaia, B.N. Khoromskij and R. Schneider. *Tensor-structured Calculation of the Two-electron Integrals in a General Basis.* SIAM J. Sci. Comp., **35** (2), A987-A1010, 2013.

[3] V. Khoromskaia. *Black-box Hartree-Fock solver by tensor numerical methods.* Comp. Methods Appl. Math., v.14 (1) , 2014, pp. 89-111.

[4] P. Benner, V. Khoromskaia and B. N. Khoromskij. *Range-Separated Tensor Format for Many-particle Modeling.* SIAM J. Sci. Comput., 40 (2), A1034-A1062, 2018.

[5] P. Benner, S. Dolgov, V. Khoromskaia and B. N. Khoromskij. *Fast Iterative Solution of the Bethe-Salpeter Eigenvalue Problem Using Low-rank and QTT Tensor Approximation.* J. Comput. Phys., 334, pp. 221-239, 2017.

[6] Venera Khoromskaia and Boris N. Khoromskij. *Tensor Numerical Methods in Quantum Chemistry.* Research monograph, De Gruyter, Berlin, 2018.

Tensor numerical methods in scientific computing and data science

Boris N. Khoromskij

Max-Planck Institute for Mathematics in the Sciences, Leipzig, Germany

Rank-structured tensor approximation of functions and operators by using the traditional canonical (CP), Tucker and tensor train (TT) formats allows the almost linear complexity scaling in dimension. Further data-compression to the logarithmic scale can be achieved by using the method of quantized-TT (QTT) tensor approximation. The novel range-separated (RS) tensor format provides the efficient low-rank representation of highly non-regular function related data in $\mathbb{R}^d$.

We discuss how the tensor numerical methods based on modern rank structured approximation techniques apply to the solution of complicated multi-dimensional problems in the PDE driven modeling, and for the efficient representation of unstructured physical data. We consider the application of RS tensor format for calculation of many-particle electrostatic potentials of bio-molecules (the Poisson-Boltzmann equation). We also show how the RS tensor format can be gainfully applied in machine learning for modeling and analysis of scattered data in $\mathbb{R}^d$.

An overview of recent results on tensor numerical methods in the optimal control problems constrained by multi-dimensional elliptic PDEs and for stochastic homogenization of the elliptic PDEs in random media, will be discussed.

Talk is based on the results presented in the research monographs [1,2], as well as in [3]-[5].

[1] Boris N. Khoromskij. *Tensor Numerical Methods in Scientific Computing.* De Gruyter Verlag, Berlin, 2018.

[2] Venera Khoromskaia, Boris N. Khoromskij. *Tensor Numerical Methods in Quantum Chemistry.* De Gruyter Verlag, Berlin, 2018.

[3] P. Benner, V. Khoromskaia and B. N. Khoromskij. *Range-separated tensor formats for many-particle modeling.* SIAM J. Sci. Comp., (2): A1034-A1062, 2018.

[4] G. Heidel, V. Khoromskaia, B. Khoromskij, and V. Schulz. *Tensor approach to optimal control problems with fractional d-dimensional elliptic operator in constraints.* E-Preprint, arXiv:1809.01971v2, 2018.

[5] V. Khoromskaia, B. N. Khoromskij, and F. Otto. *Numerical study in stochastic homogenization for elliptic PDEs: convergence rate in the size of representative volume elements.* E-Preprint, arXiv:1903:12227v1, 2019.

On the Nonlinear Matrix Equation $X^p = A + M^T(X\#B)M$

Hyun-Min Kim

Pusan National University, South Korea

Co-author: Hosoo Lee, Jeju National University, South Korea
Co-author: Jie Meng, Pusan National University, South Korea

The nonlinear matrix equation $X^p = A + M^T(X\#B)M$, where $p \geq 1$ is a positive integer, M is an $n \times n$ nonsingular matrix, A is a positive semidefinite matrix and B is a positive definite matrix, is considered. We call C#D the geometric mean of positive definite matrices C and D. Based on the properties of the Thompson metric, we prove that the nonlinear matrix equation always has a unique positive definite solution and the fixed-point iteration method can be efficiently applied. In addition, estimates of the positive definite solution and perturbation analysis are investigated. Numerical experiments are given to confirm the theoretical analysis.

Application of the PML-type absorbing boundary condition for indefinite Helmholtz problems to numerical solution of the Helmholtz problem for a vertically inhomogeneous medium.

Leonid Knizhnerman

Central Geophysical Expedition, Moscow, Russia

It is shown how the PML-type absorbing boundary condition for indefinite Helmholtz problems, earlier proposed by Druskin, Guttel and Knizhnerman, can be applied to numerical solution of the Helmholtz problem for a vertically inhomogeneous medium. The implemented scheme is semi-discrete: cutting the computational domain or other discretization in the horizontal directions is not carried out. The computational algorithm and results of numerical experiments are presented.

Besides that, the case of a quasi-3D Helmholtz equation is theoretically considered.

Channel Knowledge Equilibrium

Vladimir Lyashev

Huawei Technologies
Co-author: Kirichenko Mikhail, Huawei Technologies
Co-author: Vasily Usatyuk, Huawei Technologies

The importance of channel estimation accuracy in 4G/5G and beyond wireless communication systems cannot be overemphasized - one of the basic kind of physical layer knowledge holding on its shoulders the quality of all the others signal processing algorithms taking places higher in a system hierarchy and protocol stack. This become particularly challenging when only partial and noisy observations of the signal are available, where current methods fail to handle uncertainty appropriately. The opportunities of classical and modern mathematical methods utilization for channel processing algorithms will be raised up during the speech together with their issues and drawbacks which lead to permanent problem of seeking for equilibrium between full knowledge and absolute cost.

The Dirichlet-Dirichlet iteration method with QTT

Larisa Markeeva

Skoltech
Co-author: Ivan Tsybulin
Co-author: Ivan Oseledets

The implementation of Dirichlet-Dirichlet iteration method using Quantized Tensor Train (QTT) and z-kron operation for high memory and time consumption.

TBA

Nicola Mastronardi

Istituto per le Applicazioni del Calcolo

TBA

Newton-Krylov methods for finding solutions of aggregation-fragmentation equations

Sergey Matveev

Skolkovo Institute of Science and Technology
Co-author: Ivan Timokhin, Lomonosov Moscow State University
Co-author: Eugene Tyrtyshnikov, Marchuk Institute of Numerical Mathematics, RAS
Co-author: Alexander Smirnov, Lomonosov Moscow State University
Co-author: Nikolai Brilliantov, Skolkovo Institute of Science and Technology

In this work we present efficient implementation of Newton-Krylov methods finding stationary and quasi-stationary solutions of Smoluchowski typed kinetic equations. The proposed approach allows to obtain numerical solutions of problem of out interest in modest times with high precision without exploitation of supercomputers. There are also preliminary observations about use of proposed iterative method for implicit time integration of dynamic Cauchy problem for the same class of mathematical models.

Eigenvalue computation in inexact or tensor arithmetic

Volker Mehrmann

TU Berlin

Eigenvalue problems for partial differential equations are at the heart of many problems in all areas of science and engineering. It is essential to balance the errors in discretization and eigenvalue approximation and to also guarantee that the approximated spectrum properly reflects the physics of the underlying problem, which should be encoded in the structure of the matrices or matrix pencils. We discuss two classes of (generalized) eigenvalue problems; problems associated with Hamiltonian dynamical systems arising in adiabatic quantum computing as well as finite element analysis of disk brakes. For very high dimensional problems it is necessary to encode the matrix operations in low precision arithmetic or approximate tensor formats. We will present special techniques and a backward error analysis for inexact Krylov methods for the solution of eigenvalue problems that shows that classical Krylov space techniques can be employed in such a setting.

From splitting scheme to stochastic optimization

Daniil Merkulov

Skoltech, MIPT

Co-author: Ivan Oseledets, Skoltech

Splitting methods are well-known methods for solution of ODEs. A good systematic review is presented in ([**?**]). However, it is well known, that simple splitting schemes do not preserve steady state. That's why the rebalancing is needed. We show, that vanilla SGD could be considered as a splitting scheme for a full gradient flow. We demonstrate the connection between rebalancing splitting and stochastic average gradient method.

Fast greedy algorithms for super-large least squares problem

Stanislav Morozov

Lomonosov Moscow State University

We discuss approaches to finding an approximate solution of the least squares problem in the case when the size of the matrix does not allow to store it in memory and to perform calculations by standard greedy OMP algorithms.

The accelerated augmented Lagrangian method for yield-stress fluid flows with threshold slip boundary conditions.

Larisa Muravleva

Lomonosov Moscow State University

Viscoplastic or yield stress fluids are materials which behave like a solid below critical yield stress and flow like a viscous fluid for stresses higher than this threshold. The numerical solution of yield stress fluid flows involves nonsmooth convex optimisation problems. Traditionally, augmented Lagrangian methods (ALM) developed in the 1980-s have been used for this purpose. The main drawback of these algorithms is their frustratingly slow convergence. Beck and Teboulle (2009) present fast iterative shrinkage-thresholding algorithms (FISTA) for solving linear inverse problems arising in signal/image processing. This method, which can be viewed as an extension of the classical gradient algorithm, is attractive due to its simplicity and thus is adequate for solving large-scale problems even with dense matrix data. The proposed acceleration is of the form first proposed by Nesterov, for gradient descent methods. Later was present accelerated variant of ALM that exhibit faster convergence than ALM. Many viscoplastic fluids slip at the wall with a yield slip. The fluid slip when the tangential stress exceeds a critical value called the yield slip, and otherwise, the fluid sticks at the wall. We exploit the analogy of structure between the slip law and the viscoplastic constitutive law and apply accelerated ALM to both the viscoplastic model and the yield slip equation. The traditional ALM converges with rate $O(1/\sqrt{k})$, an accelerated variant converges with the higher and provably optimal bound O(1/k) convergence, where k is the iteration counter. This accelerated version is obtained at a negligible extra computational cost. The proposed method is used to simulate the axisymmetric squeeze flow of Bingham, Casson, and Herschel-Bulkley fluids with the slip yield boundary condition at the wall.

Matrix methods and machine learning

Ivan Oseledets

Skolkovo Institute of Science and Technology

We review different connections between matrix methods and machine learning, including compression and optimization of deep neural networks.

Low-rank based ODE and Monte-Carlo methods for temperature dependent aggregation.

Alexander Osinsky

Skoltech

The solution of the temperature-dependent Smoluchowski equations is discussed, where the kernel matrix changes after each time step in the ODE model and after each collision in Monte-Carlo simulations. Fast low-rank methods are presented with O(Nr(r + log N)) complexity of each time step for the ODE model and O(r log N) complexity of each

collision for Monte-Carlo simulation, where N is the number of equations (particle sizes) and r is the approximation rank.

Comparative Study of AVX-based Implementations of Cholesky Factorization Algorithm

George Ovchinnikov

Skoltech

Co-author: Aleksandr Katrutsa
Co-author: Ivan Oseledets
Co-author: Pavel Uydanov
Co-author: Vladimir Lyashev

This work presents comprehensive investigation of the properties of Cholesky factorization implementations with respect to their vectorization potential. Performing Cholesky factorization of some positive definite matrix is a crucial step in many linear algebra algorithms and, therefore, arises in many applications. At the same time, modern CPUs propose a set of extensions that implement Single Instruction Multiple Data (SIMD) paradigm. This paradigm claims that CPU can operate with multiple data by single instruction, which means that timing of any elementary operation can be significantly decreased. The property of any algorithm to be reformulated in the way appropriate to use SIMD operations is called vectorization potential of the algorithm. We propose study, which compares different implementations of Cholesky factorization algorithm with respect to possibility to reformulate them in the SIMD-appropriate way. To illustrate the comparison we provide numerical experiments for a wide range matrix dimensions. We compare our own implementations with Eigen library, which uses MKL subroutines as backend.

Fundamental Matrix Computations at Sublinear Cost

Victor Pan

City University of New York (Lehman College and the Graduate Center)

We cover a number of old and new algorithms for some fundamental matrix computations that run at sublinear cost. We prove that they output accurate solutions with a high probability for random inputs. Our formal analysis is in good accordance with numerical tests.

Singular Value Projection for Matrix Completion with fast approximate projectors.

Sergey Petrov

INM RAS
Co-author: Alexander Osinsky, Skoltech
Co-author: Olga Lebedeva, LMSU

'Matrix completion' refers to the problem of finding a whole matrix knowing only a small fraction of it's elements, under the assumption that the considered matrix has low-rank structure. A 'Singular Value Projection' (SVP) algorithm for solving matrix completion problem is known in literature, and strong theoretical convergence bounds have been established for that algorithm under assumption of 'incoherence' of the considered matrix, which means that the singular vectors of the unknown matrix should not be sparse. The most computationally complex part of the algorithm is a projection of a matrix to the set of low-rank matrices, which is carried out using SVD. In this work, a modification of the SVP algorithm will be considered, which is based on using approximate low-complexity SVD projection algorithms. Theoretical results will be presented that show that similar convergence bounds can be established if fast approximate SVD procedure is used. Then, a comparison of some possible approximate projection algorithms, including cross-approximation and randomized SVD, with numerical experiments, will be presented.

Tensor Completion using Singular Value Projection

Sergey Petrov

INM RAS
Co-author: Nikolai Zamarashkin, INM RAS

"Matrix Completion" refers to the problem of finding a whole matrix knowing only a small fraction of its elements, under the assumption that the considered matrix has a low-rank structure. A "Singular Value Projection" (SVP) algorithm for solving matrix completion problem is known in literature, and strong theoretical convergence bounds have been established for that algorithm under certain assumptions. In this work, a generalization of the SVP algorithm will be proposed for the "Tensor Completion" problem, the problem of finding a whole tensor knowing only a small fraction of it's elements, under the assumption that the considered tensor has small Tucker ranks. Complexity reduction ideas will be discussed, and numerical experiments will be presented, that show that the high-dimensional tensor structure allows successful completion using smaller fractions of known elements compared to the matrix completion case.

Modified MaxVol algorithm for soil sampling design based on topographical features

Anna Petrovskaia

Skolkovo Institute of Science and Technology
Co-author: Gleb Ryzhakov
Co-author: Ivan Oseledets

A soil map is a geographical representation showing a diversity of soil types and/or soil properties in the area of interest. The necessity of large-scale soil maps is increasing rapidly as humankind is facing tremendous challenges, including food security, climate change, land degradation, biodiversity loss, water resource shortage, and ecosystem sustainability.

Spatial soil sampling is an integral part of a soil survey aimed at creating a soil map. This step dramatically affects the quality and accuracy of a map, as well as the cost of a survey. Thus, there is a strong motivation to develop an algorithm, which will produce an appropriate sampling scheme. An optimal layout of soil sampling points should contain as minimum points as possible, and at the same time, it should capture the variability of soil cover sufficient for creating precision soil map.

We would like to present a sampling design, the fundamental idea of which is to select sample locations by performing Rectangular MaxVol algorithm on terrain features of a landscape. MaxVol is a fast and efficient algorithm for obtaining submatrices of a maximum second volume, so-called D-optimality criterion in the case of square matrices. The second volume of a general tall rectangular matrix is defined as a square root of the determinant of the product of the transpose matrix on itself. Thus, Rectangular MaxVol algorithm selects locations on a study site with the dissimilarities in topographic features that are the most significant for soil mapping.

We adjust our algorithm with two heuristics to make it more useful in real soil mapping practice. The first heuristic lets a user restrict minimum distance between sampling points. The second heuristic represents a penalty for placing sampling points very near to site's edges.

Two numerical experiments were conducted. The first simple experiment was performed on the natural site, where sufficient soil data, that describes the whole distribution of soil types, is known. On this site quality of sampling scheme was measured by the accuracy classification score of prediction based on selected points. The data is presented by tall matrix, in which rows corresponded to the number pixels of terrain features (approximately 56 thousand), columns - to the number of terrain features (7 features including coordinates). Developed technique let us reach 84

The second experiment was conducted on the dataset from the big agricultural field. The general approach was the same as in the first experiment. The data matrix consists of about 52 million rows and 7 columns. There was no soil information to measure the quality of the obtained sampling scheme on this site, but soil mapping experts highly appreciated the result.

Computing Low Rank Eigenspaces

Max Pfeffer

Max Planck Institute for Mathematics in the Sciences, Leipzig

Co-author: Christian Krumnow

Co-author: Andre Uschmajew

We consider the task of approximating the eigenspace belonging to the lowest eigenvalues of a self-adjoint operator on a space of matrices, with the condition that it is spanned by low rank matrices that share a common row space of small dimension. Such a problem arises for example in the DMRG algorithm in quantum chemistry. We propose a Riemannian optimization method based on trace minimization that takes orthogonality and low rank constraints into account simultaneously, and shows better numerical results compared to other current methods.

Low-rank parametrization of turbulent flow

Vladislav Pimanov

Skoltech

Co-author: Ivan Oseledets, Skoltech
Co-author: Maxim Rakhuba, ETH Zurich
Co-author: Vladimir Pimanov
Co-author: Andrzej Cichocki, Skoltech

With the rapid increase in numerical capabilities, the direct numerical simulation (DNS) of the Navier-Stokes equations become an essential tool in the computational fluid dynamics (CFD) area. However, the substantial non-linearities presented in the model lead to various (numerical) instabilities and make the study of the phenomenon of turbulence a challenging task from both physical and computational points of view.

Long-term developments in the field of the CFD have led to the formation of an increasingly large amount of data, attracting more interest from the Data Science community. For example, recently, there have been increasing attempts to describe and explain the mechanisms behind the turbulence using Deep Neural Networks. Also, the first attempts have been made to parameterize turbulent flows using Tensor Decompositions and Networks - the advanced techniques from the numerical multilinear algebra.

In the course of our study, we investigate the applicability of the Tensor Train (TT) decomposition to reduce the order of the model of a turbulent flow. Using precomputed (DNS) database, we computed TT decompositions of the velocity field at different times and we observed that TT-ranks remained almost unchanged (over time). This is quite promising observation since it explicitly suggests that the simulation (iteration in time) should be performed staying within a certain low-parametric manifold. First of all, this would significantly reduce the computational complexity of the DNS. However, our main hope is to find a physically tractable low-rank model (e.g. a convenient architecture of tensor network) that would allow us to extract the essential statistical characteristics of the turbulent flow while removing its chaotic components.

Robust solver in a quantized tensor format for electronic structure calculations

Maxim Rakhuba

ETH Zurich

The idea of reshaping an array with 2^d elements into a multidimensional $2 \times \cdots \times 2$ array and then applying tensor-train (TT) decomposition is known under the name quantized TT decomposition (QTT).

It has been shown in a number of works that arrays arising in the discretization of certain PDEs allow for QTT representation with a small number of parameters. However, the quest for robust and at the same time efficient QTT algorithm to solve PDEs with three (and more) physical dimensions is not over yet. In this talk, we address this problem using the example of PDEs arising in electronic structure calculations with a new algorithm. The proposed algorithm is capable of solving PDEs discretized using 2^{100} grid points within minutes of computational time on a laptop.

Matrix Shanks Transformations

Michela Redivo-Zaglia

University of Padova, Italy, Department of Mathematics "Tullio Levi-Civita"
Co-author: Claude Brezinski, University of Lille, France

Shanks' transformation is a well know sequence transformation for accelerating the convergence of scalar sequences. It has been extended to the case of sequences of vectors and sequences of square matrices satisfying a linear difference equation with scalar coefficients. In this talk, a more general extension to the matrix case where the matrices can be rectangular and satisfy a difference equation with matrix coefficients is proposed and studied. In the particular case of square matrices, the new transformation can be recursively implemented by the matrix ε-algorithm of Wynn. Numerical experiments showing the interest of this transformation end the talk.

Function approximation using gradient information

Gleb Ryzhakov

Skoltech

We consider the problem of multivariate function approximation in basis of smooth functions. We assume that the function is sufficiently complex computationally, so our task is to reduce the number of points at which the function is evaluated. In order to solve this problem, we adjust the least squares method (LSM) by adding information about derivatives of the function to the system matrix. This modification allows reducing the number of evaluations of approximating function while keeping the accuracy at the appropriate level. The classical result by Baur and Strassen states that we need at most 4*N operations for calculating all derivatives of a rational function, where N is the number of operation required for calculation the value of the function itself. We propose several techniques for time-efficient calculation of derivatives, which do not rely on the theorem. As the main result, we present a quasi-optimal method for selecting a given number of points at which the function is calculated. Numerical examples are given for comparison of the approximation accuracy between the standard LSM and the proposed approach.

Parallel cross interpolation for high-precision calculation of high-dimensional integrals

Dmitry Savostyanov

University of Brighton
Co-author: Sergey Dolgov (University of Bath)

High–dimensional integrals appear in problems with uncertainty and noise and are common in stochastic calculus, mathematical finance, quantum physics, etc. Textbook quadratures applied to d-dimensional integrals require exponential in d number of function evaluations, which is unfeasible for dimensions exceeding ten (this is notoriously known as curse of dimensionality). The use of Smolyak's sparse grids relaxes, but does not completely remove the problem. Currently such integrals are predominantly treated with Monte Carlo algorithm or its variants, e.g. quasi-MC, MCMC,

but their relatively slow convergence leads to excessive numerical costs and sometimes limits the accuracy of results.

Can we do better by using more structure of the original problem? We propose a new algorithm which interpolates the given function with a low–rank tensor product format using separation of variables. Our method is based on adaptive cross interpolation and maximum–volume principle — well–established algorithms for matrices, which we generalised to high–dimensional tensors [1]. For functions that admit low-rank tensor–product representation, tensor interpolation converges faster than MC and qMC, and hence can be a promising new algorithm for accurate evaluation of high–dimensional integrals. To be competitive with Monte Carlo, a parallel version of tensor interpolation algorithm is necessary, and it is presented in this talk.

We demonstrate the efficiency of the proposed algorithm for a class of Ising integrals, which appear in Ising theory of mathematical physics in relation to magnetic susceptibility of two–dimensional spin lattices. This application encourages evaluation of integrals in dimensions up to 1000 with very high precision. Monte Carlo methods are not up to the challenge. Using tensor low–rank interpolation we compute integrals accurately to 100 decimal digits [2].

We hope that this example encourages further study and exploitation of low–rank tensor product structure for problems in other subject areas where the curse of dimensionality stands in the way of delivering highly accurate results at reasonable cost.

[1] Dmitry Savostyanov, Quasioptimality of maximum–volume cross interpolation of tensors, Linear Algebra Appl. 458:217-244, 2014. [2] Sergey Dolgov, Dmitry Savostyanov, Parallel cross interpolation for high-precision calculation of high-dimensional integrals, ArXiv:1903.11554.

TBA

Reinold Schneider

TU Berlin

TBA

Efficient tensor contraction for quantum computing simulations

Roman Schutski

Skoltech, CDISE

Quantum computers are the next step in the evolution of computational devices. As the capabilities of hardware quantum processors are still inferior compared to their classical counterparts, the simulation of their operation is crucial for developing quantum algorithms. In this talk we will show that the result of a quantum program can be efficiently simulated through a sequence of tensor contractions. We will show how to use graphical models to analyze the structure of tensor networks and to estimate the amount of resources needed to evaluate its contraction in sequential and parallel algorithms.

The method of boundary hypersingular integral equations in electromagnetic scattering problems for a complex shape and structure bodies

Alexey Setukha

Lomonosov Moscow State University

The problem of a monochromatic electromagnetic wave scattering by objects, which are a combination of homogeneous dielectric regions, ideally conducting bodies and screens, is considered. For the numerical solution of such a problem, a method has been developed based on reducing the problem to a system of boundary integral equations with hypersingular integrals solved on the surfaces of ideally conducting objects and on the boundaries between dielectric medias. The arising integral equations are solved by the methods of piecewise constant approximations and collocations. To increase computational capabilities in solving problems for objects of complex shape in a wide range of wavelengths, an integrated approach is used, including: improving the quality of quadrature formulas, parallel implementations of the numerical algorithm, using low-rank approximations of the arising matrix of a linear equations system.

Two examples of matrix methods solving machine-learning problems

Alexander Shapeev

Skoltech

Machine-learning interatomic potentials are models of interatomic interaction constructed on large amounts of quantum-mechanical data, accelerating quantum-mechanical calculations by several orders of magnitude. A number of successful applications of such interatomic potentials rely on two algorithms: one predicting the energy of interatomic interaction and another one detecting atomic configurations on which a prediction involves a large extrapolation with respect to the existing training set and which need to be added to the training set.

In my talk I will give two examples of algorithms solving these problem. The first algorithm describing the interaction energy is based on the tensor-train low-rank representation of the interaction tensor. The second algorithm for determining extrapolation is motivated by the maxvol algorithm.

This work was supported by the Russian Science Foundation (grant number 18-13-00479).

Tensor factorization with nonnegativity

Elena Shcherbakova

Lomonosov Moscow State University

Co-author: Eugene Tyrtyshnikov, Marchuk Institute of Numerical Mathematics of Russian Academy of Sciences

Tensor factorization is a powerful technique for speeding up computations and reducing memory usage. But when tensor consists of only nonnegative elements sometimes we want the approximant to keep this property. In this work we propose new methods for nonnegative tensor train decomposition, research their properties, compare their performance with other known algorithms and consider several applications.

Scalability of matrix multiplication with Intel Scalable processors

Lev Shchur

Dorodnicyn Computing Centre, FRC CSC RAS
Co-author: Alexander Russkov
Co-author: Sergey Matveev

We examine the scalability of the matrix multiplication as a function of the number of threads. We check the speed of matrix multiplication on the Intel Scalable CPU with 12, 22, and 26 cores and hyperthreading. We found linear grow of the speed up to the number of threads 24, 44, and 52, respectively. We will present other findings.

Preconditioner for some integral equation of a electromagnetic problem

Stanislav Stavtsev

INM RAS

Hypersingular integral equations are used in different areas of applied mathematics and engineering. We propose development of algorithms for solution of electromagnetic wave diffraction problem at perfectly conducting object of complex form. For solving of the diffraction problem at the object with the large wave number we have to deal with the huge dense matrix. We utilize low-rank approximation of the dense matrix for construction of the solution for appropriate integral equations. This method allows to construct special approximations and performs multiplication of matrix of size $N \times N$ by vector of size N within $\mathcal{O}(N \log(N))$ or $\mathcal{O}(N)$ operations instead of original $\mathcal{O}(N^2)$. In our work, we use the iterative method (GMRES) for solution of system with big dense matrix represented in low-rank format. The matrix is ill-conditioned in the case of the large wave numbers. Hence, we apply preconditioning techniques. We build the preconditioner on the base of uncompressed matrix blocks and decrease number of GMRES-iterations. It is worth to notice that the preconditioner is a sparse matrix. All in all, we exploit MUMPS package to solve the system with sparse matrix at high performance computers in parallel.

Implementation of Tensor-based Solvers for Smoluchowski-type Equations

Daniil Stefonishin

Skoltech CDISE
Co-author: Sergey Matveev (Skoltech)
Co-author: Dmitry Zheltkov (INM RAS)

In this talk we describe a C/C++ implementation for fast finite difference solvers for Smoluchowski-type population balance equations. The considered equations correspond to a model of aggregation (also many-particle) and fragmentation process with sources and sinks in a huge spatially-homogeneous system of chaotically moving particles. Computational algorithms and data representation are based on usage of tensor decompositions for approximate representation of kinetic

coefficients, such as polyadic (multilinear) and tensor train (TT) decompositions. The implementation allows user to easily modify the whole set of model parameters describing physical processes, utilized tensor representations and to choose different compilers.

Sensing Less While Computing Tensor Decompositions with Non-negativity

Eugene Tyrtyshnikov

Marchuk Institute of Numerical Mathematics, RAS and Lomonosov MSU

We follow the paradigm of using only small part of matrix and tensor entries that allows one to construct a sufficiently accurate and fast appoximation for matrices and tensors that are astronomicaly large, i.e. cannot be placed in any available computer memory and are accessed implicitly through calls to a procedure producing any individual entry on demand. In this talk we focus on how this approach can be used in the cases when we need to maintain nonnegativity of the elements.

Chebyshev polynomials and best rank-one approximation ratio

Andre Uschmajew

MPI MiS Leipzig

We establish a new extremal property of the classical Chebyshev polynomials in the context of the theory of rank-one approximations of tensors. We also give some necessary conditions for a tensor to be a minimizer of the ratio of spectral and Frobenius norms. This is joint work with Andrei Agrachev and Khazhgali Kozhasov.

Low-rank factorization of multivariate functions: tensor-based approaches

Konstantin Usevich

CNRS, Nancy (France)

We consider the problem of factorizing multivariate maps in sums of nonlinear transformations of linear forms. This factorization is motivated by nonlinear system identification and also can be interpreted as a shallow neural network with flexible nonlinearities. As we will show in the talk, the problem can be viewed as an X-rank decomposition, and can be reformulated as a structured CP decomposition of a certain tensor. We will report results on properties of such factorizations and on different tensorizations.

This talk is based on joint works with P. Comon, Y. Qi, P. Dreesen, and M. Ishteva.

Industry trends and optimization challenged from wireless communication

Andrey Vorobyev

Huawei Technologies

In this work:

1. We want to review new Wireless industry trends, like 5G, MIMO, high efficiency PA, mmWave, antenna array.
2. We want to show challenging optimization tasks from our research.

An Efficient Preconditioner for Low-frequency Electromagnetic Modeling on Structured Hexahedral grids

Nikolay Yavich

CDISE, Skoltech

Co-author: Mikhail Malovichko, Skoltech

Co-author: Michael Zhdanov, University of Utah

This talk addresses the problem of efficient solution of large sparse linear systems arising in 3D finite-element (FE) low-electromagnetic modeling on structured hexahedral grids. To design an efficient preconditioner, we exploit the relation between finite-element and finite-difference (FD) matrices. We first considered a preconditioner based on the FD matrix that corresponds to a layered conductivity distribution. This matrix has been used to precondition FD systems in the past. For FE systems, however, it did not provide convergence of an iterative solver whenever hexahedrons are deformed. To gain robustness, we combined the FD preconditioner with a smoothing procedure. The obtained preconditioner happens to be fast, robust, and applicable to complex electromagnetic modeling. We demonstrated effectiveness of this approach on a real geophysical modeling example.

Low-rank approximations and HPC for advection-coagulation problems

Rishat Zagidullin

Skolkovo Institute of Science and Technology

The work is devoted to analysis of acceleration opportunities for numerical algorithm solving the advection-coagulation equation. At the algorithmic level we decrease computational time using approximations of low rank coagulation kernels. The rest of the speedup is achieved via hybrid parallel calculations performed with CPU and GPU utilization. The results are illustrated for a problem statement on two-dimensional unstructured spatial grid and one-dimensional particle size grid. We achieve almost linear speedup using Zhores and Lomonosov supercomputers.
References:
[1] Zagidullin, R. R., Smirnov, A. P., Matveev, S. A., Tyrtyshnikov, E. E. (2017). An efficient numerical method for a mathematical model of a transport of coagulating particles. Moscow University Computational Mathematics and Cybernetics, 41(4), 179-186.
[2] Matveev, S. A., Zagidullin, R. R., Smirnov, A. P., Tyrtyshnikov, E. E. (2018). Parallel numerical algorithm for solving advection equation for coagulating particles. Supercomputing Frontiers and Innovations, 5(2), 43-54.
[3] Zacharov, Igor, et al. "'Zhores'–Petaflops supercomputer for data-driven modeling, machine learning and artificial intelligence installed in Skolkovo Institute of Science and Technology." arXiv preprint arXiv:1902.07490 (2019).

Probabilistic estimates for matrix cross approximations

Nikolai Zamarashkin

INM RAS
Co-author: A.I. Osinsky (Skoltech)

We will discuss new probabilistic estimates for matrix cross approximations based on the maximum volume principle.

D-optimal experimental design using a gradient-based sampling method with application to the least-squares polynomial approximation

Vitaly Zankin

Skoltech, MIPT
Co-author: Gleb Ryzhakov, Skoltech
Co-author: Ivan Osledets, Skoltech, INM RAS

In this work, we introduce a novel sampling method for finding a D-optimal experimental design. This method is developed in the context of finding the least-squares polynomial approximation of computationally costly multidimensional function, having a restricted number of evaluations. The resulting algorithm allows to sample such a set of points, called a design, from the function domain of interest that obtained least-squares polynomial approximation will have the least possible error. The efficiency of the proposed method is illustrated by conducting a comparative study with other sampling techniques (namely, LHS, Sobol' sequence sampling, and Maxvol sampling). Additionally, numerical experiments for the Lebesgue constant growth for the points sampled by the proposed method are carried out.

Efficient Nonnegative Tensor Ring Decomposition via A New Algorithm Framework

Ning Zheng

Tensor Learning Unit, RIKEN Center for Advanced Intelligence Project, Japan
Co-author: Yuyuan Yu, School of Automation, Guangdong University of Technology, China
Co-author: Chao Li, RIKEN Center for Advanced Intelligence Project, Japan
Co-author: Qibin Zhao, RIKEN Center for Advanced Intelligence Project, Japan

Tensor decomposition has been widely used for the dimensional reduction and extraction of the meaningful latent features of high dimensional tensor data. In many applications, the underlying data ensemble is nonnegative and consequently the nonnegative tensor decomposition is proposed to achieve additive parts-based representation and to learn more physically interpretable results. As the corresponding tensor optimization problem has computational difficulty due to nonconvex, together with sparse, smooth, graph based Tikhonov regularization, the construction and analysis of the reliable, efficient and robust algorithms are required. Under the framework of block coordinate descent method, we aim to present a new iterative algorithm which is based on the modulus type variable transformation. The theoretical analysis of the proposed method is discussed. Numerical experiments including the synthetic data and image data show the efficiency and superiority of the proposed method comparing with the state-of-the-art methods.

On the properties of parallel discrete event simulation algorithms

Liliia Ziganurova

National Research University Higher School of Ecomonics, Landau Institute for theoretical physics, Scientific center in Chernogolovka RAS
Co-author: Shchur Lev

Parallel discrete event simulation is an efficient method of large-scale simulations. It allows to run a single simulation program on thousands of computational nodes and/or cores. The synchronization between parallel processes is provided by special synchronization protocols. We study the models of the evolution of local virtual times of parallel processes in conservative and optimistic synchronization algorithms. We assume that parallel processes are arranged on regular and small-world communication topology. We measured the average speed of the local virtual times profile and the average width of the profile. These observables can be mapped on the utilization of processing time and the desynchronization degree of the algorithms, respectively. It is known that the conservative algorithm on regular topology belongs to the Kardar-Parisi-Zhang universality class, while the optimistic algorithm belongs to the universality class of Directed Percolation. We studied the influence of long-range communications between processes on the properties of the algorithms. We show that the presence of a small number of long-range communications enhances the synchronization between parallel processes, while slightly reducing the average utilization.