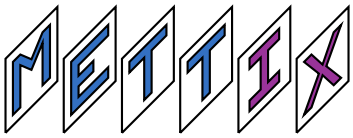
Book of Abstracts

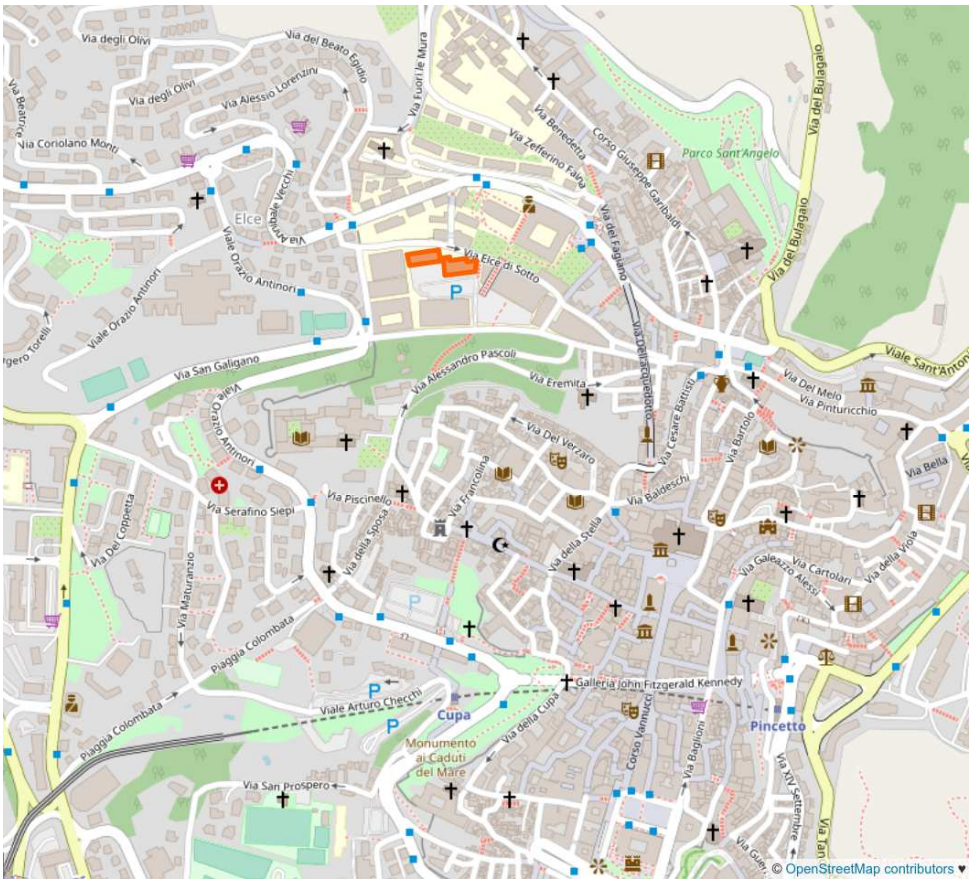
Matrix Equations and Tensor Techniques IX



Matrix Equations and Tensor Techniques IX

Perugia, September 9-10, 2021

Venue The conference is scheduled as a hybrid in-person/virtual conference. It will take place in the Dipartimento di Matematica e Informatica dell'Università di Perugia, Via Luigi Vanvitelli 1, 06123, Perugia (PG).



The METT Conferences

2005 Leipzig

2007 Chemnitz

2009 Braunschweig

2011 Aachen

2013 Lausanne

2015 Bologna

2017 Pisa

2019 Magdeburg

2021 Perugia

Steering Committee

Peter Benner (Max Planck Institute, Magdeburg, Germany)

Heike Faßbender (Technische Universität Braunschweig, Germany)

Lars Grasedyck (Aachen University, Germany)

Daniel Kressner (École Polytechnique Fédérale de Lausanne, Switzerland)

Beatrice Meini (Università di Pisa, Italy)

Valeria Simoncini (Università di Bologna, Italy)

Organizing Committee

Bruno Iannazzo (Università di Perugia, Italy)

Elena Addis (Università di Perugia, Italy)

Massimiliano Fasi (Örebro University, Sweden)

Beatrice Meini (Università di Pisa, Italy)

Federico Poloni (Università di Pisa, Italy)

Giulia Riganelli (Università di Perugia, Italy)

Leonardo Robol (Università di Pisa, Italy)

Programme

Thursday, 9 September

09:00 – 09:10	Opening Remarks
09:10 – 09:50	Keynote Session I
09:50 – 10:40	Blitz Session I
10:40 – 11:10	Coffee Break
11:10 – 12:50	Talk Session I
12:50 – 14:40	Lunch Break
14:40 – 16:20	Talk Session II
16:20 – 16:50	Coffee Break
16:50 – 17:50	Blitz Session II

Friday, 10 September

09:00 – 09:40	Keynote Session II
09:40 – 10:40	Talk Session III
10:40 – 11:10	Coffee Break
11:10 – 12:50	Talk Session IV
12:50 – 14:40	Lunch Break
14:40 – 16:20	Talk Session V
16:20 – 16:50	Coffee Break
16:50 – 18:10	Talk Session VI
18:10 – 18:20	Concluding Remarks

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Elementwise accurate doubling algorithm for shifted M-matrix algebraic Riccati equations

B. Iannazzo *E. Addis* F. Poloni

Università di Perugia, Italy

We consider the nonsymmetric algebraic Riccati equation (NARE)

$$XBX - XA - DX - C = 0, \tag{1}$$

where A, B, C, D are real matrices of sizes $n \times n$, $n \times m$, $m \times n$, $m \times m$, respectively. We focus on the case in which the matrix

$$M = \begin{bmatrix} A & -B \\ C & D \end{bmatrix},$$

is an irreducible singular M-matrix. The problem of finding the minimal non negative solution of such Riccati equations arises in applied probability, transportation theory, fluid queues.

A *structure-preserving doubling algorithm (SDA)* for computing the minimal solution of (1) has been proposed in [3]. This iterative method relies on the fact that the problem of solving (1) can be reduced to the computation of certain invariant subspaces of the matrix $H = \begin{bmatrix} I_n & 0 \\ 0 & -I_m \end{bmatrix} M$, and its convergence properties are connected with a quotient involving the eigenvalues of H.

In [2] Guo et al. studied the doubling algorithm in the case where M is an irreducible singular M-matrix, and in order to speed up the convergence proposed a shift technique to move one zero eigenvalue of H to a positive real number. This approach modifies the equation (1) introducing a shifted equation that shares with (1) the solution, and leads to a reduction of the quotient that controls the convergence so as to produce a dramatic decrease of the number of steps of the algorithm.

In the case where M is a non singular M-matrix or an irreducible singular M-matrix, algorithms computing the minimal non negative solution of (1) with high elementwise relative accuracy have been proposed in [6], [4], [5] . The general approach is based on the idea that a non singular M-matrix can be inverted by the GTH-like algorithm [1],

that consists in a modification of the standard Gaussian elimination in a cancellation-free fashion, when a *triplet representation* of the matrix is known. A triplet representation of A is a triple (P, u, v) such that $P \geq 0$, $u > 0$, $v \geq 0$ with P matrix with null diagonal entries and $A = D - P$, where D is diagonal, and $Au = v$.

Unfortunately, the shifted matrix $\hat{M}$ constructed in [2] in general is no longer a M-matrix, so the known elementwise accurate algorithms can not be applied directly together with the shift technique in order to improve the accuracy and also accelerate the convergence.

We present an elementwise accurate algorithm using the shift technique for the computation of the minimal non negative solution of (1), when M in an irreducible singular M-matrix.

We propose the idea of *delayed shift* and some results that guarantee the applicability and the convergence of structured doubling algorithm based only on the properties of the matrix of the initial setup of doubling algorithm instead of matrix M or $\hat{M}$. We provide a componentwise error analysis for the algorithm and we also show some numerical experiments that illustrate the advantage in terms of accuracy and convergence speed.

References

[1] Alfa, A.S., Xue, J., Ye, Q.: Accurate computation of the smallest eigenvalue of a diagonally dominant M-matrix, Math. Comput. **71**(237), 217–236 (2002)

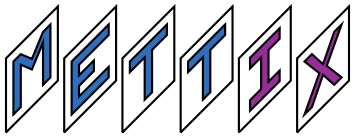
[2] Guo, C.H., Iannazzo, B., Meini, B.: On the doubling algorithm for a (shifted) nonsymmetric algebraic Riccati equation. SIAM J. Matrix Anal. Appl. **29**(4), 1083–1100 (2007)

[3] Guo, X.-X., Lin,W.-W., Xu, S.-F.: A structure-preserving doubling algorithm for nonsymmetric algebraic Riccati equation. Numer. Math. **103**, 393–412 (2006)

[4] Nguyen, G.T., Poloni, F.: Componentwise accurate fluid queue computations using doubling algorithms. Numer. Math. **130**(4), 763–792 (2015)

[5] Xue, J., Li, R.-C. Highly accurate doubling algorithms for M-matrix algebraic Riccati equations. Numer. Math., **135**(3):733–767 (2017)

[6] Xue, J., Xu, S., Li, R.-C.: Accurate solutions of M-matrix algebraic Riccati equations. Numer. Math. **120**(4), 671–700 (2012)



Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Randomized Algorithms for Rounding the Tensor Train Format

Grey Ballard Paul Cazeaux Eric Hallman Agnieszka Miedlar

Mirjeta Pasha Tim Reid Arvind Saibaba
Hussam Al Daas

Rutherford Appleton Laboratory, STFC, Harwell Campus, Oxfordshire

The Tensor Train (TT) format is a highly compact low-rank representation for high-dimensional tensors. TT is useful in particular in representing approximations to the solution of certain types of parametrized partial differential equations. The fundamental operation used to maintain feasible memory and computational time is called rounding, which truncates the internal ranks of a tensor already in TT format. We propose several randomized algorithms for this task that are generalizations of randomized low-rank matrix approximation algorithms and provide significant reduction in computation compared to deterministic TT rounding algorithms. Randomization is particularly effective in the case of rounding a sum of TT tensors, which is the bottleneck computation in the adaptation of GMRES to vectors in TT format. In this talk, we will present the randomized algorithms and compare their empirical accuracy and computational time with deterministic alternatives.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Solving matrix equations encountered in stochastic processes by means of fixed point iterations

Dario A. Bini Guy Latouche Beatrice Meini

Dipartimento di Matematica, Università di Pisa, Italy

We consider the problem of computing special solutions of certain nonlinear matrix equations encountered in stochastic processes. A first class of equations is of the kind

$$X = \sum_{i=-1}^{\infty} A_i X^{i+1}, \tag{1}$$

where A_i, for $i \geq -1$, are nonnegative square matrices such that $\sum_{i=-1}^{\infty} A_i$ is stochastic. This class of equations is fundamental in the analysis of M/G/1-type Markov chains which model a large variety of queuing problems in applied probability. The minimal nonnegative solution G, besides having an interesting probabilistic interpretation, provides an effective tool for representing the invariant probability measure of the chain.

Another class of equations, encountered in the analysis of Markov-modulated Lévy processes, has a matrix integral form of the kind

$$\Delta_a Y + \frac{1}{2}\Delta_{\sigma^2} Y^2 + \int_0^{\infty} \Delta_\nu(x)(e^{Yx} - I)dx + Q \circ U(0) + \int_0^{\infty} (Q \circ \mu(x))e^{Yx}dx = 0 \tag{2}$$

where the $n \times n$ matrix Y is the unknown, Q is an irreducible generator, Δ_v is the $n \times n$ diagonal matrix with diagonal entries v_i, $i = 1, \ldots, n$, the vectors $a = (a_i)_{i=1,\ldots,n}$ and $\sigma^2 = (\sigma_i^2)_{i=1,\ldots,n}$ are parameters defining a Brownian motion, $\nu_i(\cdot)$, $i = 1, \ldots, n$, are continuous densities, and $U_{ij}(\cdot)$, $i, j = 1, \ldots, n$, are distribution functions, such that $U_{ij}(x) = U_{ij}(0) + \int_0^x \mu_{ij}(u)du$. A matrix denoted as G, that solves (2), plays a fundamental role in the analysis of such stochastic processes. In the positive recurrent case, the matrix G is a generator and is the unique solution in the set of real matrices of order n having a simple eigenvalue equal to 0 and $n-1$ eigenvalues with strictly negative real parts [4, 5].

Classical fixed point iterations have been designed and analyzed in the literature for equation (1) and are customarily used for the computation of the matrix G [2], [6]. Concerning equation (2), few authors have considered the question of computing G. Breuer [4] and Simon [7] suggest some functional iterations but no convergence analysis nor experimental evidence is given.

In this talk we introduce a new family of fixed point iterations for the numerical computation of the minimal nonnegative solution G of equation (1) that includes the

classical iterations, and a new approach for computing the solution of interest of (2). Concerning (1), the idea relies on rewriting the original equation as a polynomial matrix equation of degree $q+1$ of the form

$$X = B_{-1}(X) + B_0(X)X + B_1(X)X^2 + \cdots + B_q(X)X^{q+1}, \tag{3}$$

where the coefficients $B_i(X)$ are suitable power series of X. The sequence $\{X_k\}_{k\geq 0}$ generated by the iteration, given X_0, is such that the value of X_{k+1} is the minimal nonnegative solution of the polynomial equation (3) with coefficients $B_i(X_k)$. The value of q and the choice of $B_i(X)$ characterize the specific fixed-point iteration in the class.

We show that the classical iterations are obtained for $q = 0$. Moreover, by means of a general convergence analysis, for any $q > 0$ we determine the coefficients $B_i(X)$ that maximize the convergence speed. As a result, we obtain new fixed-point iterations which are much faster than the classical ones. Numerical experiments confirm the effectiveness of our extension. More details can be found in [1].

Inspired by this approach, we design some fixed point iterations for solving (2). The idea consists in solving a sequence of quadratic matrix equations, where the matrix coefficients defining the matrix equation depend on the current approximation to the solution G. The numerical methods differ in the way the quadratic matrix equations are generated. In fact, relying on some changes of variable, we transform (2) into a sequence of either unilateral quadratic matrix equations or special nonsymmetric algebraic Riccati equation. A theoretical algorithmic analysis together with numerical tests are performed. Comparisons with the algorithms proposed by Breuer [4] and Simon [7] show the effectiveness of our approach. More details can be found in [3].

References

[1] D.A. Bini, G. Latouche, B. Meini, *A family of fast fixed point iterations for M/G/1 Markov chains* IMA J. of Numerical Analysis 2021.

[2] D.A. Bini, G. Latouche, B. Meini, *Numerical Methods for Structured Markov Chains*. Oxford University Press 2005.

[3] D. A. Bini, G. Latouche, and B. Meini, *Numerical solution of a matrix integral equation arising in Markov Modulated Lévy processes*, submitted for pubblication, 2021, http://arxiv.org/abs/2107.11611.

[4] L. Breuer, *First passage times for Markov additive processes with positive jumps of phase type*, J. Appl. Prob., 45 (2008), pp. 779–799.

[5] B. D'Auria, J. Ivanovs, O. Kella, and M. Mandjes, *First passage of a Markov additive process and generalized Jordan chains*, J. Appl. Probab., 47 (2010), pp. 1048–1057.

[6] G. Latouche, V. Ramaswami, *Introduction to Matrix Analytic Methods in Sochastic Modeling*. SIAM Philadelphia 1999.

[7] M. Simon, *Markov-Modulated Processes: Brownian Motions, Option Pricing and Epidemics*, PhD thesis, Université libre de Bruxelles, 2017.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

A Tensor-Train Dictionary Learning algorithm based on spectral proximal alternating linearized minimization.

Domitilla Brandoni Margherita Porcelli Valeria Simoncini

Department of Mathematics, Alma Mater Studiorum-Università di Bologna,
Piazza di Porta San Donato 5, Bologna

Dictionary Learning (DL) is one of the leading sparsity promoting techniques in the context of image classification, where the "dictionary" matrix D of images and the sparse matrix X are determined so as to represent a redundant image dataset Y. The resulting constrained optimization problem $\min_{D,X} \|Y - DX\|_F$ is nonconvex, non-smooth and NP-hard, providing several computational challenges for its solution (see e.g. [1]). To preserve multidimensional data features, various tensor DL formulations have been introduced, adding to the problem complexity (see e.g. [2]). Unfortunately all the tensor-based DL methods in the literature are not supported with theoretical convergence analysis. We propose a new tensor formulation of the DL problem using a Tensor-Train decomposition ([3]) of the multi-dimensional dictionary, together with a new alternating algorithm for its solution. The new method belongs to the Proximal Alternating Linearized Minimization (PALM) algorithmic family (see e.g. [4]), with the inclusion of second order information to enhance efficiency. We discuss a rigorous convergence analysis, and report on the new method performance on the image classification of several benchmark datasets.
This talk is based on [5].

References

[1] Mairal J., Bach F. and Ponce J. 2014 *Sparse Modeling for Image and Vision Processing*. Found. Trends. Comput. Graph. Vis. 8(2-3), 85–283.

[2] Dantas, C. F., Cohen, J. E. and Gribonval, R. 2019 *Learning tensor-structured dictionaries with application to hyperspectral image denoising*. In *27th European Signal Processing Conference (EUSIPCO)* pp. 1-5, IEEE.

[3] Oseledets, I. V. 2011 *Tensor-train decomposition*, pp. 2295–2317, SIAM J. Sci. Comput., 33.

[4] Bolte, J., Sabach, S. and Teboulle, M. 2014 *Proximal alternating linearized minimization for nonconvex and nonsmooth problems*. Math. Prog., 146(1), 459–494.

[5] Brandoni D. , Porcelli M. and Simoncini V. 2021 *A Tensor-Train Dictionary Learning algorithm based on Spectral Proximal Alternating Linearized Minimization*, pp.1-26, arXiv: 2107.11644.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

On the number of elements needed for low-rank tensor train completion

Stanislav Budzinskiy

Marchuk Institue of Numerical Mathematics RAS, 119333, Moscow

In this talk we explore the problem of recovering a tensor $\mathbf{A} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$ with low tensor train ranks $\mathbf{r} = (1, r_1, \ldots, r_{d-1}, 1)$ from a small portion of its entries indexed by $\Omega \subset [n_1] \times \ldots \times [n_d]$:

$$\|\mathcal{R}_\Omega \mathbf{X} - \mathcal{R}_\Omega \mathbf{A}\|_F^2 \to \min \quad \text{s.t.} \quad \mathbf{X} \in \mathbb{R}^{n_1 \times \ldots \times n_d}, \quad \mathrm{rk}_{TT}(\mathbf{X}) = \mathbf{r}.$$

By $\mathcal{R}_\Omega$ we denote the sampling operator that sets to zero all entries that are not in Ω.

In the matrix case one knows how many entries $\Omega \subset [n_1] \times [n_2]$ are needed to complete an incoherent low-rank matrix, albeit in the nuclear-norm-minimization formulation:

$$\|X\|_* \to \min \quad \text{s.t.} \quad X \in \mathbb{R}^{n_1 \times n_2}, \quad \mathcal{R}_\Omega X = \mathcal{R}_\Omega A.$$

Namely, if Ω is chosen uniformly at random with replacement then

$$O\left(r(n_1 + n_2)\log^2(n_1 + n_2)\right)$$

entries are sufficient to recover a matrix with high probability [1]. Fewer elements [2]

$$O\left(r(n_1 + n_2)\log(n_1 + n_2)\right)$$

guarantee local convergence of the Riemannian gradient descent applied to

$$\|\mathcal{R}_\Omega X - \mathcal{R}_\Omega A\|_F^2 \to \min \quad \text{s.t.} \quad X \in \mathcal{M}_r = \{X \in \mathbb{R}^{n_1 \times n_2} : \mathrm{rk}(X) = r\}.$$

The method exploits the geometric structure of the set $\mathcal{M}_r$, which is an embedded submanifold of $\mathbb{R}^{n_1 \times n_2}$.

For tensor trains the problem is less understood (there is progress in the Tucker case [3, 4]). We follow the geometric route and establish local convergence guarantees—in terms of the sample size $|\Omega|$—of the Riemannian gradient descent for tensor train completion:

$$\|\mathcal{R}_\Omega \mathbf{X} - \mathcal{R}_\Omega \mathbf{A}\|_F^2 \to \min \quad \text{s.t.} \quad \mathbf{X} \in \mathcal{M}_\mathbf{r} = \{\mathbf{X} \in \mathbb{R}^{n_1 \times \ldots \times n_d} : \mathrm{rk}_{TT}(\mathbf{X}) = \mathbf{r}\}.$$

To this we end we extend the notion of incoherence from matrices (their column and row subspaces) to tensor trains.

We then consider tensor completion with side information. In this problem, we are additionally given low-dimensional subspaces that contain the mode-k fiber spans of the tensor. The presence of side information makes it possible to significantly lower the number of entries sufficient for matrix completion in the nuclear norm formulation [5, 6]. We obtain similar reduction in the tensor train case.

This work has been supported by Russian Science Foundation Project (21-71-10072).

References

[1] B. Recht, *A simpler approach to matrix completion*, Journal of Machine Learning Research, (2011) pp. 3413-3430.

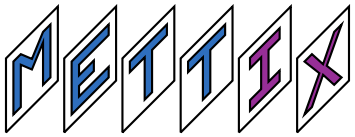
[2] K. Wei et al., *Guarantees of Riemannian optimization for low rank matrix completion*, arXiv:1603.06610 [math], 2016.

[3] C. Mu et al., *Square deal: Lower bounds and improved relaxations for tensor recovery*, International Conference on Machine Learning, (2014) pp. 73-81.

[4] M. Yuan, C.-H. Zhang, *On tensor completion via nuclear norm minimization*, Found Comput Math, (2016) pp. 1031-1068.

[5] P. Jain, I.S. Dhillon, *Provable inductive matrix completion*, arXiv:1306.0626 [cs, math, stat], 2013.

[6] M. Xu, R. Jin, Z. Zhou, *Speedup matrix completion with side information: Application to multi-label learning*, in Advances in Neural Information Processing Systems 26, Curran Associates Inc., 2013.



Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Solving the Discrete Euler-Arnold Equations for the Generalized Rigid Body Motion

João R. Cardoso Pedro Miraldo

Coimbra Polytechnic–ISEC,
and Center for Mathematics, University of Coimbra, Portugal

In [4], Moser and Veselov proposed the following equations to discretize the classical Euler-Arnold differential equations for the motion of a rigid body:

$$\begin{aligned} M_{k+1} &= \omega_k M_k \omega_k^T \\ M_k &= \omega_k^T J - J\omega_k, \end{aligned} \tag{1}$$

where M_k is the angular momentum with respect to the body (here represented by a skew-symmetric matrix), J is the inertia matrix (symmetric positive definite), and ω_k (orthogonal matrix) is the angular velocity. Rigid body equations arise in several applications, e.g., celestial mechanics, molecular dynamics, mechanical robotics and flight control, where they are used in particular to understand the body–body interactions of particles like planets, atoms and molecules.

The main challenge of solving (1) is to find an orthogonal matrix ω_k in the second equation, by assuming that J and M_k are given. Mathematically, the problem may be formulated as finding a special orthogonal matrix X ($X^TX = I$, $\det(X) = 1$) such that

$$XJ - JX^T = M, \tag{2}$$

where J is a given symmetric positive definite matrix, and M is a known skew-symmetric matrix. The matrix equation (2) was firstly investigated in [4], where the authors based their developments on factorizations of certain matrix polynomials. A different approach, but computationally more efficient, was provided later in [2], where the authors noted that (2) can be connected with a certain algebraic Riccati equation and, in turn, with the *Hamiltonian* matrix

$$\mathcal{H} = \begin{bmatrix} M/2 & I \\ M^2/4 + J^2 & M/2 \end{bmatrix}. \tag{3}$$

It is stated in [2] that (2) has a solution $X \in \mathcal{SO}(n)$ (the special orthogonal or rotation group of order n) if and only if the size of the Jordan blocks associated to the pure imaginary eigenvalues of $\mathcal{H}$ (if any) is even. The existing algorithms for solving (2) only work when $\mathcal{H}$ does not admit any pure imaginary eigenvalue. Moreover, the algorithms

based on solving the associated algebraic Riccati equation require the strong condition that the matrix $M^2/4 + J^2$ must be symmetric positive definite. These issues have motivated us to investigate methods whose applicability does not require those restrictive conditions.

The problem of finding a special orthogonal solution X in (2) can be formulated as an optimization problem in the following way:

$$\min_{X \in \mathcal{SO}(n)} \left\| XJ - JX^T - M \right\|_F^2, \tag{4}$$

where $\|.\|_F$ denotes the Frobenius norm. Techniques from Riemannian geometry to solve optimization problems with orthogonal constraints have attracted the interest of many researchers in the last decades; see [1, 3], and the references therein. An essential feature of those techniques is that they allow the transformation of a constrained optimization problem into an unconstrained one. Since the set of orthogonal matrices is a manifold and provided that the objective function satisfies some smoothness requirements, we can make available tools such as Euclidean gradients, Riemannian gradients, retractions, and geodesics.

In this talk, we propose two iterative methods for solving (4). They evolve on the orthogonal manifold and belong to the family of line search methods on matrix manifolds described in [1, Ch. 4]. They are constraint-preserving, in the sense that, starting with a matrix $X_0 \in \mathcal{SO}(n)$, all the iterates X_k also stay in $\mathcal{SO}(n)$. The first one splits the orthogonal constraints using the Bregman method, whereas the second method is of steepest-descent type, based on a Cayley-transformation to preserve the constraints and on a Barzilai-Borwein step size. A set of numerical experiments are carried out to compare the performance of the proposed algorithms, suggesting that the first algorithm has the best performance in terms of accuracy and number of iterations. An essential advantage of these two iterative methods is that they work even when the conditions for applicability of the direct methods available in the literature are not satisfied. That is, they allow the computation of special orthogonal solutions, even when $M^2/4 + J^2$ is not symmetric positive definite. Those iterative algorithms may also be used in problems where $\mathcal{H}$ has purely imaginary eigenvalues associated with Jordan blocks of even size, but, as will be illustrated with experiments, the convergence may slow down.

References

[1] P.-A. Absil, R. Mahony, R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, New Jersey, 2007.

[2] J. R. Cardoso, F. S. Leite, *The Moser-Veselov equation*, Linear Alg. Appl. 360, (2003), pp. 237-248.

[3] A. Edelman, T. Arias, S. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl. 20 (2), (1999) pp. 303-353.

[4] J. Moser, A. P. Veselov, *Discrete versions of some classical integrable systems and factorization of matrix polynomials*, Commun. Math. Phys. 139, (1991), pp. 217-243.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

A μ-mode integrator for solving evolution equations in Kronecker form

Marco Caliari *Fabio Cassini* Lukas Einkemmer

Alexander Ostermann Franco Zivcovich

Department of Mathematics, University of Trento,
Via Sommarive 14 – Povo (TN), Italy

Due to the importance of simulation in various fields of science and engineering, devising efficient numerical methods for solving high–dimensional evolutionary partial differential equations is of considerable interest. In this talk, we present a μ-mode integrator for computing the solution of stiff evolution equations. It is based on a d-dimensional splitting approach and it suitably combines, in a tensor framework, one-dimensional matrix exponentials (usually precomputed).

We show that our integrator solves exactly linear problems in Kronecker form with time-invariant coefficients, i.e. problems which can be written as

$$\mathbf{u}'(t) = M\mathbf{u}(t), \quad \mathbf{u}(0) = \mathbf{u}_0 \tag{1}$$

where

$$M = \sum_{\mu=1}^{d} A_{\otimes\mu}$$

and

$$A_{\otimes\mu} = I_d \otimes \cdots \otimes I_{\mu+1} \otimes A_\mu \otimes I_{\mu-1} \otimes \cdots \otimes I_1,$$

being A_μ an $n_\mu \times n_\mu$ matrix and I_μ the identity matrix of size n_μ. More in detail, it computes efficiently the exact solution of (1), i.e. $\mathbf{u}(t) = \exp(tM)\mathbf{u}_0$, by means of tensor techniques without explicitly forming the matrix M. This scheme can also be used as a building block for numerically solving more general classes of PDEs compared to (1), for example in the context of a splitting method or an exponential integrator.

We further explain how the needed actions of matrix exponentials can be implemented efficiently on modern computer systems, such as multithreaded CPUs and GPUs. In particular, the overall computational cost of our algorithm is $\mathcal{O}(N\max_\mu n_\mu)$, with $N = n_1 \cdots n_d$, while the storage requirement scales as $\mathcal{O}(N)$. Hence, the scheme is ideally suited to modern hardware, especially for GPUs.

We finally illustrate the features and the performances of the μ-mode integrator, both on CPUs and on GPUs, by numerically solving a range of problems from physics, such as three-dimensional heat equations (see Figure 1) and three-dimensional linear and nonlinear Schrödinger equations. In particular, we show that our integrator can significantly outperform numerical methods well established in the field and that we can obtain performance improvements between a factor of 10 and 20 by performing computations on GPUs rather than on CPUs.

If time allows, we also present how μ-mode products can be employed to compute spectral transforms efficiently even if no *fast* transform is available. This technique is useful, for example, in the context of a Hermite pseudospectral method.

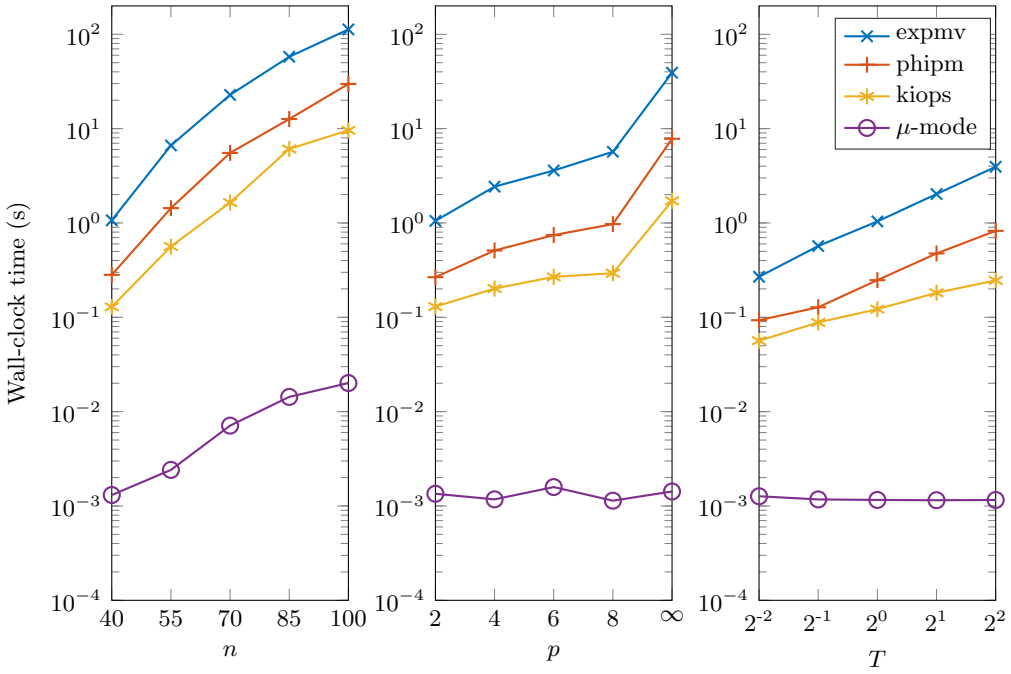


Figure 1: The wall-clock time for solving a three–dimensional heat equation is shown as a function of the size $n_\mu = n$ (left), of the order of the finite difference scheme p (middle), and of the final time T (right).

References

[1] M. Caliari, F. Cassini, L. Einkemmer, A. Ostermann, F. Zivcovich, *A μ-mode integrator for solving evolution equations in Kronecker form*, Submitted.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Decoupling Dominant Z-eigenvectors of Tensor Kronecker Products

Charles Colley Huda Nassar David Gleich

Computer Science at Purdue University, 305 N University St, West Lafayette, IN

We present a new extremal Z-eigenvector theorem for the Kronecker products of tensors which shows that the dominant eigenvector of $\underline{\boldsymbol{B}} \otimes \underline{\boldsymbol{A}}$ decouples to the dominant eigenvectors of $\underline{\boldsymbol{B}}$ and $\underline{\boldsymbol{A}}$, a surprising generalization of the matrix counterpart. Formally, we show [1]

> *Let $\underline{\boldsymbol{A}}$ be a symmetric, k-mode, m-dimensional tensor and $\underline{\boldsymbol{B}}$ be a symmetric, k-mode, n-dimensional tensor. Suppose that $(\lambda_A^*, \mathbf{u}^*)$ and $(\lambda_B^*, \mathbf{v}^*)$ are any dominant tensor Z-eigenvalues and vectors of $\underline{\boldsymbol{A}}$ and $\underline{\boldsymbol{B}}$, respectively. Then $(\lambda_A^* \lambda_B^*, \mathbf{v}^* \otimes \mathbf{u}^*)$ is a dominant eigenpair of $\underline{\boldsymbol{B}} \otimes \underline{\boldsymbol{A}}$.*

Decoupling the tensor Kronecker product gives us a new foothold to facilitate the use of *motifs* – small repeating subgraphs – in network alignment algorithms. We explore how this new theorem impacts TAME [2], an unsupervised method for network alignment – which is used to identify protein ortholog candidates from protein interaction data. The key to TAME was a tensor that encoded the presence of motifs among groups of nodes, and the algorithm sought to use the eigenvector of the Kronecker product of two such tensors to indicate how the original networks ought to align. Our new algorithms $\Lambda-$TAME & LowRankTAME outperform the original TAME when aligning the LVGNA [3] protein interaction networks, both running orders of magnitude faster and matching, at the minimum, as many triangles motifs as TAME.

Specifically, the decoupling allows our new $\Lambda-$TAME to use the power iterations of $\underline{\boldsymbol{T}}_A$ and $\underline{\boldsymbol{T}}_B$ independently to access useful information on a possible dominant eigenspace. This moves the computational bottleneck from computing iterations with the Kronecker product of tensors, as in TAME, to the step of producing an alignment from the possible eigeninformation. This means that $\Lambda-$TAME runs two orders of magnitude faster than TAME on the largest problems tested and aligns more triangles. When finding the maximum matching is infeasible, searching over the quadratic pairs of vectors computed in any step of the power iterations, is a simple algorithm that runs 3 orders of magnitude faster than TAME in exchange for a modest reduction in triangles matched.

The theory allows further improvements to the computations. Note that the above dominant eigenvector result strongly suggests that iterates of a power method applied to a

Kronecker product of tensors should have low-rank structure (when viewed as a matrix). We show how to use low-rank structure within the iterates of TAME itself, in concert with a tensor-generalized mixed product property [4, Thm 3.1], to produce a method that uses low-rank structure when appropriate. Thus, even without changing the underlying iteration, we show how to improve the computation TAME's iterates. In our experiments we see that we can run an order of magnitude faster than TAME, producing the same iterates with less susceptibility to numerical imprecision and opening the door to low-rank matching strategies.

We hope that our work will encourage a wider dissemination of these techniques, revealing the seemingly daunting runtime costs of tensor Kronecker products as nothing more than a paper tiger.

References

[1] C. Colley, H. Nassar, D. Gleich, *Addressing Computational Bottlenecks in Higher-Order Graph Matching with Tensor Kronecker Product Structure*, arXiv:2011.08837, 2020.

[2] M. Shahin, D. Gleich, T.Kolda, A. Grama, *Triangular Alignment TAME: A Tensor-Based Approach for Higher-Order Network Alignment*, IEEE/ACM Trans. Comput. Biol. Bioinform., (2017), pp.1446–1458.

[3] L. Meng, A. Striegel, T. Milenković, *Local versus global biological network alignment*, Bioinformatics, 2016, pp.3155–3164.

[4] J. Shao, *A general product of tensors with applications*, Linear Algebra and its applications, 2013, pp. 2350–2366,

[5] L. Lim, *Singular values and eigenvalues of tensors: a variational approach*, 1st Intl. Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2005, pp. 129–132.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Divide and conquer methods for functions of matrices with banded or hierarchical low-rank structure

Alice Cortinovis Daniel Kressner Stefano Massei

Institute of Mathematics, EPFL, 1015 Lausanne, Switzerland

This talk is concerned with approximating matrix functions for banded matrices, hierarchically semiseparable matrices, and related structures [1]. We propose new divide-and-conquer methods – in the spirit of the divide-and-conquer algorithms developed in [2] for linear matrix equations – which exploit the fact that these matrices can be (recursively) decomposed as a sum $A = D + R$ of a block diagonal matrix D and a low-rank correction R. While the update $f(A) - f(D)$ often has low numerical rank and can be approximated via (rational) Krylov subspace projections [3, 4], the block diagonal part $f(D)$ is computed recursively for each diagonal block.

We present a convergence analysis that relates the accuracy attained by the algorithm with best polynomial or rational approximations of the function. For the special case of a banded matrix, we show that the divide-and-conquer method reduces to a much simpler algorithm, which proceeds by computing matrix functions of small submatrices of A. When only the trace or the diagonal of the matrix function is of interest, we demonstrate – in practice and in theory – that convergence can be faster.

Finally, we test the algorithms on a variety of matrices and functions; the numerical results demonstrate that, most of the time, the proposed methods outperform state-of-art techniques with respect to time consumption and offer a comparable accuracy.

References

[1] A. Cortinovis, D. Kressner, S. Massei, *Divide and conquer methods for functions of matrices with banded or hierarchical low-rank structure*, https://arxiv.org/abs/2107.04337.

[2] D. Kressner, S. Massei, L. Robol, *Low-rank updates and a divide-and-conquer method for linear matrix equations*, SIAM Journal on Scientific Computing 41.2 (2019): A848-A876.

[3] B. Beckermann, D. Kressner, M. Schweitzer, *Low-rank updates of matrix functions*, SIAM Journal on Matrix Analysis and Applications (2018): 539-565.

[4] B. Beckermann, A. Cortinovis, D. Kressner, M. Schweitzer, *Low-rank updates of matrix functions II: Rational Krylov methods*, SIAM Journal on Numerical Analysis (2021): 1325-1347.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

The Extended Aluthge Transform

Chafiq Benhida *Raúl E. Curto*

Department of Mathematics, University of Iowa, Iowa City, Iowa, U.S.A.

Consider a bounded linear operator T acting on a complex separable Hilbert space $\mathcal{H}$ (finite or infinite dimensional), and let $T \equiv V|T|$ be the canonical polar decomposition of T (that is, V is a partial isometry and $\ker T = \ker V$). The Aluthge transform of T is the operator $\Delta(T) := |T|^{1/2} V |T|^{1/2}$. For P an arbitrary positive operator such that $VP = T$, we define the *extended Aluthge transform* of T associated with P, as follows: $\Delta_P(T) := P^{1/2} V P^{1/2}$.

First, we establish some basic properties of $\Delta_P(T)$. For instance, whenever $P \geq 0$ and $VP = T$, one automatically has

(i) $|T| \leq P$,

(ii) $\ker P \subseteq \ker |T|$,

(iii) P commutes with $|T|$, and

(iv) the restrictions of P and $|T|$ to the range of $|T|$ agree.

We also derive a 3×3 operator matrix representation for $\Delta(T)$ and $\Delta_P(T)$ relative to the orthogonal decomposition $\mathcal{H} = \overline{\text{Ran}}\ |T| \oplus \overline{\text{Ran}}\ (P|_{\ker T}) \oplus \ker P$. Along the way, we prove the following Intertwining Property: $|T|^{1/2} \Delta_P(T) P^{1/2} = P^{1/2} \Delta(T) |T|^{1/2}$.

Second, we study the fixed points of the extended Aluthge transform.

Third, we consider the case when T is an idempotent, and prove an optimal result for its associated Aluthge transform and extended Aluthge transform.

Next, we discuss whether the extended Aluthge transform leaves invariant the class of complex symmetric operators.

We also study how $\Delta_P(T)$ transforms the numerical radius and numerical range of T.

Finally, as a key application, we prove that the spherical Aluthge transform of a commuting pair of operators corresponds to the extended Aluthge transform of a 2×2 operator matrix built from the pair; thus, the theory of extended Aluthge transforms yields results for spherical Aluthge transforms.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

From analysis to learning: Tensor-based assessment of latent similarity

Lieven De Lathauwer

KU Leuven

In data analysis, tensor decompositions such as the Canonical Polyadic Decomposition (CPD) and Block Term Decomposition (BTD) may be used as basic tools. They allow one to break a single data set into interpretable components. In CPD, the terms are rank-1, while, in the more general BTD, they have low multilinear rank. Instrumental is the mildness of the conditions under which the tensor decompositions are unique (for instance, no orthogonality requirements, as in the QR-factorization or singular value decomposition of matrices). Countless applications have been reported in telecommunication, array processing, audio and image processing, chemometrics, psychometrics, astrophysics, biomedical signal processing, bio-informatics, ... The mildness of the conditions goes together with the possibility of computational issues: the lack of properties such as orthogonality makes that the terms in a decomposition can be arbitrarily close to each other, and hence numerically difficult to separate; in some cases the problem can even be ill-posed.

In this talk we take the step from data analysis to data comparison, and from the decomposition of a single tensor to the assessment of the similarity between components of different tensors. Assessing similarity is a key task in pattern recognition and machine learning. We will show that, also in the latter setting, tensors provide fundamentally new possibilities beyond matrix techniques. Moreover, under mild conditions, the assessment of similarity can be done by conventional linear algebra, i.e. the estimation of angles between subspaces, solving sets of linear equations in least-squares sense and matrix eigenvalue decomposition. The number of terms and their multilinear rank (in the case of BTD) can be found as well. The results will be illustrated with applications.

References

[1] I. Domanov and L. De Lathauwer, *From computation to comparison of tensor decompositions*, SIAM J. Matrix Anal. Appl., 42.2, (2021), pp. 449–474.

[2] F. Van Eeghem and L. De Lathauwer, *Tensor similarity in chemometrics*, in *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*, (S. Brown, R. Tauler, and B. Walczak, eds.), Elsevier, 2020.

[3] F. Van Eeghem, O. Debals, and L. De Lathauwer, *Tensor similarity in two modes*, IEEE Trans. Signal Process., 66.5, (2018), pp. 1273–1285.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

On the consistency of $X^\top AX = B$ when B is either symmetric or skew

Alberto Borobia Roberto Canogar *Fernando De Terán*

Departamento de Matemáticas, Universidad Carlos III de Madrid, Leganés, Spain

In this talk, we analyze the consistency of the matrix equation

$$X^\top AX = B, \tag{1}$$

where $A \in \mathbb{C}^{m\times m}$, $X \in \mathbb{C}^{m\times n}$ (unknown), and $B \in \mathbb{C}^{n\times n}$ is either symmetric or skew-symmetric. In particular, we will first provide a necessary condition for (1) to have a solution X. Then, we will prove that this condition is also sufficient for most matrices A and an arbitrary symmetric (or skew) matrix B. To be more precise, we will first show that, in order to analyze the consistency of (1), we can restrict ourselves to the case where A and B are in *Canonical form for congruence*. We use the canonical form for congruence introduced in [3], which is a direct sum of blocks of three types. Then, we will show that the condition mentioned above is sufficient when A does not contain any blocks of some of these types with certain size.

We want to emphasize that the question on the consistency of (1), when B is symmetric (respectively, skew), is equivalent to the following problem: given a bilinear form over $\mathbb{C}^n$ (represented by the matrix A), find the maximum dimension of a subspace such that the restriction of the bilinear form to this subspace is a symmetric (resp., skew) non-degenerate bilinear form.

The results presented in this talk for B being symmetric have been published in [1], whereas the ones for B being skew-symmetric are contained in the submitted manuscript [2].

References

[1] A. Borobia, R. Canogar, F. De Terán, *On the consistency of the matrix equation $X^\top AX = B$ when B is symmetric*, Mediterr. J. Math, (2021) pp. 18-40.

[2] A. Borobia, R. Canogar, F. De Terán, *The equation $X^\top AX = B$ with B skew-symmetric: How much of a bilinear form is skew-symmetric?*, submitted (2021).

[3] R. A. Horn, V. V. Sergeichuk. *Canonical forms for complex matrix congruence and $*$-congruence*, Linear Algebra Appl., 216 (2006) pp. 1010-1032.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Computing the condition number of tensor decompositions through Tucker compression

Paul Breiding *Nick Dewaele* Nick Vannieuwenhoven

KU Leuven

In this talk, we investigate the condition number of *structured block term decompositions* [1], which are a general class of tensor decompositions encompassing the tensor rank decomposition, the block term decomposition, and sums of tensor train products. These decompositions express a tensor $\mathcal{A}$ as a sum of R "simple" terms $\mathcal{A} = \sum_{r=1}^{R} \mathcal{A}_r$. Specifically, the summand $\mathcal{A}_r$ can be expressed as a multilinear product $(U_1, \ldots, U_D) \cdot \mathcal{C}$ where $U_1, \ldots, U_D$ are full-rank matrices and $\mathcal{C}$ lies on a manifold $\mathcal{M}_r \subseteq \mathbb{R}^{l_1 \times \cdots \times l_D}$ that satisfies two assumptions:

1. Every $\mathcal{X} \in \mathcal{M}_r$ has multilinear rank $(l_1, \ldots, l_D)$.
2. $\mathcal{M}_r$ is invariant under changes of basis.

In practice, the given tensor $\mathcal{A}$ is almost always corrupted by noise, so it is essential to quantify how sensitive the summands $\mathcal{A}_1, \ldots, \mathcal{A}_R$ are to perturbations of $\mathcal{A}$. We show how Rice's condition number [2] can be applied to this decomposition. If $\widetilde{\mathcal{A}}$ is sufficiently close to $\mathcal{A}$ and $\widetilde{\mathcal{A}}$ has a structured block term decomposition $\widetilde{\mathcal{A}} = \widetilde{\mathcal{A}}_1 + \cdots + \widetilde{\mathcal{A}}_R$, then the sensitivity is bounded by

$$\|(\widetilde{\mathcal{A}}_1 - \mathcal{A}_1, \ldots, \widetilde{\mathcal{A}}_R - \mathcal{A}_R)\| \lesssim \kappa \|\widetilde{\mathcal{A}} - \mathcal{A}\|$$

where κ is the condition number and $\|\cdot\|$ is the Frobenius norm.

The computation of structured block term decompositions is often sped up by applying a dimensionality technique known as Tucker compression. That is, one expresses $\mathcal{A} = (Q_1, \ldots, Q_D) \cdot \mathcal{G}$ where $Q_1, \ldots, Q_D$ are matrices with orthonormal columns and $\mathcal{G}$ is a tensor with much smaller dimensions than $\mathcal{A}$. Then, $\mathcal{G}$ is decomposed as $\mathcal{G} = \sum_{r=1}^{R} \mathcal{G}_r$, which corresponds to a structured block term decomposition $\mathcal{A} = \sum_{r=1}^{R} (Q_1, \ldots, Q_D) \cdot \mathcal{G}_r$.

Since $\mathcal{G}$ has fewer possible perturbations than $\mathcal{A}$, one would expect the condition number of its decomposition to be smaller than that of $\mathcal{A}$'s decomposition. However, our main result is that the two condition numbers are equal. This is in contrast to other problems, where the condition number of the structured problem is much lower than that of the unstructured one [3].

Our result implies an algorithm to compute the condition number of the decomposition of $\mathcal{A}$ based on Tucker compression. This algorithm can reach a speedup of over four orders of magnitude relative to the state of the art in practical cases, so that it is now possible to compute the condition numbers of decompositions of large tensors.

References

[1] N. Dewaele, P. Breiding, N. Vannieuwenhoven, *The condition number of many tensor decompositions is invariant under Tucker compression*, arXiv 2106.13034, (2021), pp. 1-18

[2] J. Rice, *A theory of condition*, SIAM J. Numer. Anal., (1966), pp.287-310

[3] B. Arslan, V. Noferini, F. Tisseur, *The structured condition number of a differentiable map between matrix manifolds, with applications*, SIAM J. Matrix Anal. Appl., (2019), pp. 774-799

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

A predictor-corrector method in quantized tensor train format for the equilibrium of the chemical master equation

Trang Dinh Trang Dinh Brandon M. Reid Roger B. Sidje

Department of Mathematics, University of Alabama, Tuscaloosa, USA

It is well known that networks with complex interacting components arise in a variety of disciplines [4], and that a master equation can help describe the joint probability function of the components over time. This work focuses on the chemical master equation (CME) that results from modeling the system using a continuous-time Markov chain. It is of our interest to study the system's long-term behavior, which can be explored by computing the stationary solution to the CME. However, the number of states involved grows exponentially with the number of chemical species tracked. To cope with an extremely large or even infinite state space when solving the CME in biological problems, a potent strategy is to restrict to a finite state projection (FSP) [10] and represent the transition matrix and probability vector in quantized tensor train (QTT) [11] format, leading to savings in storage while retaining accuracy. In summary, we propose a method that further advances the efficiency of using tensor-based approaches to estimate the probability mass function when the system is in statistical equilibrium. We do so by providing a computationally inexpensive initial approximation from which to efficiently obtain a more robust estimation of the stationary solution. Our methodology is to use the reaction rate equations of the system to roughly estimate the probability function of each chemical species. This cheap estimate is then fed to a predictor-corrector pair of solvers aimed at quickly reaching equilibrium by alternating between a linear system solver (predictor) and a transient solver (corrector). Numerous techniques have been developed that leverage tensors in recent years [11, 6, 8, 7, 2, 5, 9, 12] but here we use the AMEn solver the alternating minimal energy (AMEn) algorithm [3] for the predictor part and our adaptive FSP tensor-based solver with sliding windows [1] for the corrector part.

References

[1] T. Dinh and R.B. Sidje. An adaptive solution to the chemical master equation using quantized tensor trains with sliding windows. Phys. Biol., 17(6):065014, 2020.

[2] S. Dolgov and B. Khoromskij. Simultaneous state-time approximation of the chemical master equation using tensor product formats. Numer. Linear Algebra Appl., 22:197–219, 2015.

[3] S.V. Dolgov and D. Savostyanov. Alternating minimal energy methods for linear systems in higher dimensions. SIAM J. Sci. Comput., 36(5):A2248–A2271, 2014.

[4] J. Goutsias and G. Jenkinson. Markovian dynamics on complex reaction networks. Phys. Rep., 529(2):199 – 264, 2013.

[5] L. Grasedyck, D. Kressner, and C. Tobler. A literature survey of low-rank tensor approximation techniques. GAMM Mitteilungen, 36(1):53–78, 2013.

[6] M. Hegland and J. Garcke. On the numerical solution of the chemical master equation with sums of rank one tensors. ANZIAM J., 52, 2011.

[7] V. Kazeev, M. Khammash, M. Nip, and C. Schwab. Direct solution of the chemical master equation using quantized tensor trains. PLoS Comput. Biol., 10(3), 2014.

[8] B.N. Khoromskij. Tensors-structured numerical methods in scientific computing: Survey on recent advances. Chemometr. Intell. Lab., 110(1):1–19, 2012.

[9] T.G. Kolda and B.W. Bader. Tensor decompositions and applications. SIAM Rev., 51(3):455–500, 2008.

[10] B. Munsky and M. Khammash. The finite state projection algorithm for the solution of the chemical master equation. J. Chem. Phys., 124(4):044104, 2006.

[11] I.V. Oseledets. Tensor-train decomposition. SIAM J. Sci. Comput., 33(5):2295–2317, 2011.

[12] H. D. Vo and R. B. Sidje. An adaptive solution to the chemical master equation using tensors. J. Chem. Phys., 147:044102, 2017.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

A recursive eigenspace computation for the Canonical Polyadic decomposition

Eric Evert Michiel Vandecappelle Lieven De Lathauwer

ESAT, KU Leuven Kulak, Kortrijk

Tensors, or multiindexed arrays, play an important role in fields such as machine learning and signal processing. These higher-order generalizations of matrices allow for preservation of higher-order structure present in data, and low rank decompositions of tensors allow for compression of data and recovery of underlying information [12, 2, 3]. One of the most popular decompositions for tensors is the canonical polyadic decomposition (CPD) which expresses a tensor as a sum of rank one tensors.

An important feature of the CPD is that, with mild assumptions [6, 8, 11], the CPD of a low rank tensor is unique. It is this uniqueness that allows for extraction of component information from a signal tensor. Furthermore, for a tensor with a unique CPD, the CPD can often be found algebraically. Such an algebraic solution can typically be obtained with limited computation time, hence is often used as an initialization for optimization based methods when the tensor is noisy.

One of the most popular algorithms for algebraic computation of a CPD of a tensor is the generalized eigenvalue decomposition (GEVD) [4, 5, 10, 9]. The key idea behind the classical GEVD is that a factor matrix of a tensor may be obtained by computing the generalized eigenvectors of any subpencil of the tensor.

While in the noiseless setting GEVD can exactly recover a CPD, it has recently been shown that pencil based algorithms such as GEVD are unstable [1]. That is, the condition number for computing a generalized eigenvalue decomposition of a subpencil can be arbitrarily larger than the condition number [14] for computing the CPD of the tensor. In this talk we present an extension of the GEVD algorithm which significantly improves the stability of algebraic computation of the CPD.

The stability of computing generalized eigenvectors of a matrix pencil is heavily dependent on the separation between the generalized eigenvalues and the generalized eigenvectors of the pencil [13, 7]. In the case that the generalized eigenvalues and eigenvectors are well separated, a GEVD may be stably computed. However, when either a pair of generalized eigenvalues or generalized eigenvectors are near parallel, computation of the generalized eigenvectors becomes unstable. As such, the GEVD algorithm performs well when there is a subpencil for the tensor in which all generalized eigenvalues are well separated; however, GEVD runs into challenges if one is unable to find a subpencil in which all generalized eigenvalues are well separated.

It is not hard to show that the difficulty caused by poorly separated generalized eigenvalues necessarily occurs as tensor rank and dimensions increase. Indeed, taking a subpencil of a tensor is equivalent to projecting the columns of one of the tensor's factors to be vectors of length 2. The original columns lie in a vector space of dimension I where typically $I >> 2$. Roughly speaking, the stability of the original CPD computation is dependent on the separation of the columns of the original factor matrix, while the stability of the generalized eigenvalue decomposition used to compute the CPD is dependent on the separation between the projected columns. Of course, the separation between the columns can significantly decrease under a projection. This in turn causes instability for the GEVD algorithm.

We address this fundamental issue by using many different pencils to compute the CPD. Intuitively, this allows us to consider many projections of the original factor columns, and allows us to take advantage of the fact that given clusters of columns will be better separated under some projections than others. More precisely rather than using a single pencil and computing all of its generalized eigenvectors, we use many different pencils and in each pencil compute generalized eigenspaces corresponding to sufficiently well separated generalized eigenvalues. The generalized eigenspaces we compute are then used to decompose the tensor in question as a sum of tensors with reduced rank. This is done in a way so that the CPD of the original tensor can be recovered by computing the CPDs of the summand tensors. Though the resulting "generalized eigenspace decomposition" is still fundamentally pencil based, it is significantly more robust to noise than the classical GEVD.

We will present a detailed explanation of the generalized eigenspace decomposition algorithm, and we will compare the performance in terms of accuracy and computational time of the generalized eigenspace decomposition to GEVD. In addition, we will examine stability of the generalized eigenspace decomposition both empirically and theoretically.

References

[1] C. Beltrán, P. Breiding, N. Vannieuwenhoven, *Pencil-Based Algorithms for Tensor Rank Decomposition are not Stable*, SIAM J. Matrix Anal. Appl., (2019) pp. 739-773.

[2] A. Cichocki, D. P. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. F. Caiafa, A.-H. Phan, *Tensor decompositions for signal processing applications: From two-way to multiway component analysis* IEEE Signal Process. Mag., (2015) pp. 145-163

[3] P. Comon, C. Jutten, *Handbook of blind source separation: Independent component analysis and applications* Academic Press, 2009.

[4] I. Domanov, L. De Lathauwer, *Canonical Polyadic Decomposition of Third-Order Tensors: Reduction to Generalized Eigenvalue Decomposition*, SIAM J. Matrix Anal. Appl., (2014) pp. 636-660.

[5] I. Domanov, L. De Lathauwer, *Canonical polyadic decomposition of third-order tensors Relaxed uniqueness conditions and algebraic algorithm*, Linear Algebra Appl., (2017) pp. 342-375.

[6] I. Domanov, L. De Lathauwer, *On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part I: Basic results and uniqueness of one factor matrix*, SIAM J. Matrix Anal. Appl., (2013) pp. 855-875.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Matrix-oriented FEM formulation for stationary and time-dependent PDEs on x-normal domains

Massimo Frittelli Ivonne Sgura

Dept. of Mathematics and Physics "E. De Giorgi", University of Salento,
Via per Arnesano, 73100 Lecce (Italy)

The usual spatial finite element discretisation, of arbitrary order $k \in \mathbb{N}$, of elliptic and parabolic partial differential equations takes the form of a linear system or a system of ODEs in the parabolic case. On the so-called *x-normal domains*, we show that the method allows for a Matrix-Oriented formulation, called MO-FEM, [1]. In the elliptic case, the discrete problem takes the form of a *multiterm Sylvester equation*, in the parabolic case a sequence of multiterm Sylvester equations after time discretisation. The proposed framework encompasses the special case $k = 1$ on square and rectangular domains [2], where the discrete problem is a standard (two-term) Sylvester equation.

On square domains, each Sylvester equation can be solved very efficiently with the so-called *reduced method* in the spectral space. On general x-normal domains, when the reduced approach does not apply, we solve each multiterm Sylvester equation apply through the matrix-oriented form of the Preconditioned Conjugate Gradient (MO-PCG) with an ad-hoc preconditioner. The MO-PCG proves more efficient, in terms of computational time and memory occupation, than its standard counterpart in vector form and than MATLAB's built-in direct solver.

As an application, we consider reaction-diffusion PDE systems, where the coupling between diffusion and nonlinear kinetics can lead to the so-called Turing instability. To capture the morphological peculiarities of the Turing patterns, a very fine space discretisation is required, limiting the use of standard (vector-based) ODE solvers in time because of excessive computational costs. To show the advantages of the MO-FEM-PCG to approximate Turing patterns with high spatial resolution, we apply the MO-FEM to a two-species reaction-diffusion system for battery modeling [3].

References

[1] M. Frittelli, I. Sgura. 2021, *Matrix-oriented FEM discretization for stationary and time-dependent PDEs on x-normal domains*, submitted

[2] M.C. D'Autilia, I. Sgura, and V. Simoncini 2020, *Matrix-oriented discretization methods for reaction-diffusion PDEs: comparisons and applications*, Computers and Mathematics with Applications CAMWA 79, 2067–2085. DOI: 10.1016/j.camwa.2019.10.020

[3] B.Bozzini, D.Lacitignola, I.Sgura 2013, *Spatio-temporal organization in alloy electrodeposition: a morphochemical mathematical model and its experimental validation*, Journal of Solid State Electrochemistry 17(2), 467-479. DOI: 10.1007/s10008-012-1945-7

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Solving large-scale Riccati equations with indefinite quadratic terms

Peter Benner *Jan Heiland* Steffen W. R. Werner

Max Planck Institute for Dynamics of Complex Technical Systems,
Sandtorstraße 1, 39106 Magdeburg, Germany

Algebraic Riccati equations with indefinite quadratic terms of the form

$$A^\mathsf{T} XE + E^\mathsf{T} XA + E^\mathsf{T} X \left(B_1 B_1^\mathsf{T} - B_2 B_2^\mathsf{T}\right) XE + C^\mathsf{T} C = 0, \tag{1}$$

with $A, E \in \mathbb{R}^{n\times n}$, $B_1 \in \mathbb{R}^{n\times m_1}$, $B_2 \in \mathbb{R}^{n\times m_2}$, $C \in \mathbb{R}^{p\times n}$ and E invertible, play an important role in applications related to robust controller design and differential games; see, e.g., [5, 7].

A particular interest lies in the existence and computation of a symmetric positive semi-definite, stabilizing solution $X_\infty \in \mathbb{R}^{n\times n}$ of (1). In other words, we want to compute an X_∞ that solves (1), that is symmetric positive semi-definite, and that ensures that the eigenvalues of the matrix pencil $\lambda E - (A + B_1 B_1^\mathsf{T} - B_2 B_2^\mathsf{T}) X_\infty E$ all lie in the left open half-plane.

While there are some established approaches to that in the case of small-scale dense coefficient matrices [1, 6, 8], there is no approach available to compute solutions in the large-scale sparse setting. In our work, we propose an extension of the iterative procedure developed in [6] to efficiently compute the requested solution of (1) in the large-scale sparse case via low-rank approximations such that $Z_\infty Z_\infty^\mathsf{T} \approx X_\infty$, with $Z_\infty \in \mathbb{R}^{n\times r}$ and $r \ll n$. The approach is based on considering the Riccati operator

$$\mathcal{R}(X) := A^\mathsf{T} XE + E^\mathsf{T} XA + E^\mathsf{T} X(B_1 B_1^\mathsf{T} - B_2 B_2^\mathsf{T}) XE + C_1^\mathsf{T} C_1.$$

For two symmetric matrices $X_1 = X_1^\mathsf{T}$ and $X_2 = X_2^\mathsf{T}$, one can show that

$$\mathcal{R}(X_1 + X_2) = \mathcal{R}(X_1) + \widetilde{A}^\mathsf{T} X_2 E + E^\mathsf{T} X_2 \widetilde{A} + E^\mathsf{T} X_2 (B_1 B_1^\mathsf{T} - B_2 B_2^\mathsf{T}) X_2 E$$

holds, where $\widetilde{A} := A + (B_1 B_1^\mathsf{T} - B_2 B_2^\mathsf{T}) X_1 E$. In the case that X_2 is a solution to the algebraic Riccati equation with negative semi-definite quadratic term

$$0 = \mathcal{R}(X_1) + \widetilde{A}^\mathsf{T} X_2 E + E^\mathsf{T} X_2 \widetilde{A} - E^\mathsf{T} X_2 B_2 B_2^\mathsf{T} X_2 E, \tag{2}$$

the residual reads

$$\mathcal{R}(X_1 + X_2) = E^\mathsf{T} X_2 B_1 B_1^\mathsf{T} X_2 E.$$

This leads to an iterative procedure, where in each step a Riccati equation of the form (2) needs to be solved. Using low-rank approximations for the intermediate Riccati equations (2) and some clever formulations of the iteration matrices allows the use of classical large-scale sparse solvers for (2), like the ones described, e.g., in [2, 3, 10]. Together with a reformulation of the overall iteration approach in [6], this leads to our new *low-rank Riccati iteration (LR-RI)* method.

Implementations of this new approach are available in [4] for dense systems and in [9] for the large-scale sparse case.

References

[1] W. F. Arnold and A. J. Laub, *Generalized eigenproblem algorithms and software for algebraic Riccati equations*, Proc. IEEE, (1984) pp. 1746–1754.

[2] P. Benner, Z. Bujanović, P. Kürschner, and J. Saak, *RADI: A low-rank ADI-type algorithm for large scale algebraic Riccati equations*, Numer. Math., (2018) pp. 301–330.

[3] P. Benner, J.-R. Li, and T. Penzl, *Numerical solution of large-scale Lyapunov equations, Riccati equations, and linear-quadratic optimal control problems*, Numer. Lin. Alg. Appl., (2008) pp. 755–777.

[4] P. Benner and S. W. R. Werner, *Model reduction of descriptor systems with the MORLAB toolbox*, (2019) see http://www.mpi-magdeburg.mpg.de/projects/morlab.

[5] J. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, *State-space solutions to standard H_2 and H_∞ control problems*, IEEE Trans. Autom. Control, (1989) pp. 831–847.

[6] A. Lanzon, Y. Feng, B. D. .O. Anderson, and M. Rotkowitz, *Computing the positive stabilizing solution to algebraic riccati equations with an indefinite quadratic term via a recursive method*, IEEE Trans. Autom. Control, (2008) pp. 2280–2291.

[7] D. Mustafa and K. Glover, *Controller reduction by $\mathcal{H}_\infty$-balanced truncation*, IEEE Trans. Autom. Control, (1991) pp. 668–682.

[8] J. D. Roberts, *Linear model reduction and solution of the algebraic Riccati equation by use of the sign function*, Internat. J. Control, (1980) pp. 677–687.

[9] J. Saak, M. Köhler, and P. Benner, *M-M.E.S.S.-2.1 – The Matrix Equations Sparse Solvers library*, (2021) see https://www.mpi-magdeburg.mpg.de/projects/mess.

[10] V. Simoncini, *Analysis of the rational Krylov subspace projection method for large-scale algebraic Riccati equations*, SIAM J. Matrix Anal. Appl., (2016) pp. 1655–1674.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Palindromic linearization and numerical solution of nonsymmetric algebraic $\top$-Riccati equations

Peter Benner *Bruno Iannazzo* Beatrice Meini Davide Palitta

Università degli Studi di Perugia, Italy

We consider the Nonsymmetric algebraic $\top$-Riccati equation ($\top$-NARE)

$$DX + X^\top A - X^\top BX + C = 0, \tag{1}$$

where X is the unknown matrix and $A, B, C, D \in \mathbb{R}^{n\times n}$ are the coefficients, while the superscript $\top$ denotes transposition. Equation (1) has been considered in [1], with applications to solving large-scale Dynamic Stochastic General Equilibrium models.

The $\top$-NARE takes its name from the nonsymmetric algebraic Riccati equation (NARE)

$$DX + XA - XBX + C = 0, \tag{2}$$

whose analysis and numerical solution has been of great interest in the literature in the last decades (see the books [2], [3]), and from the transposition $\top$ of the unknown X in (1), when it premultiplies a matrix coefficient. Indeed, recently, the $\top$ counter-part of classical linear matrix equations has been widely studied [4], [5].

A standard procedure when dealing with an algebraic Riccati equation is the "linearization" that relates its solutions to the invariant (deflating) subspaces of a matrix (pencil). This greatly improves the analysis and the numerical solution of the equation. For instance, Equation (2) is associated with the matrix

$$H = \begin{bmatrix} A & -B \\ -C & -D \end{bmatrix},$$

and X is a solution to (2) if and only if there exists an n-dimensional invariant subspace of H spanned by the columns of $\begin{bmatrix} I \\ X \end{bmatrix}$.

Here, we introduce a linearization for the $\top$-Riccati equation (1). More specifically, using the coefficients of the matrix equation, we construct a $\top$-palindromic pencil $\varphi(z) = M + zM^\top$ of size $2n$, that linearizes the equation: if $\varphi(z)$ is regular and if X is a solution to (1), then the columns of $\begin{bmatrix} I \\ X \end{bmatrix}$ span a deflating subspace of $\varphi(z)$ and also a kind of converse result holds.

This linearization, besides being interesting per se, opens the way to find solutions of $\top$-Riccati equations by relying on algorithms that compute bases of deflating subspaces of a matrix pencil, such as the QZ algorithm and the Doubling algorithm.

In our tests we show that the two latter algorithms are more efficient, in terms of computational cost and CPU time, than Newton's method, the reference algorithm in [1], keeping the same accuracy.

Another interesting feature of our linearization of the $\top$-NARE is that it captures the peculiar structure of the problem, and this structure can be exploited by applying the palindromic QZ algorithm [6], [7], a structured variant of the QZ. We develop a structured ordering procedure for the palindromic QZ algorithm that allows us to find the required basis and gives computational advantages, being superior, in terms of forward error, in some difficult problems.

References

[1] P. Benner, D. Palitta, *On the solution of the nonsymmetric T-Riccati equation*, Electron. Trans. Numer. Anal., 54 (2021) pp. 68-88.

[2] D. A. Bini, B. Iannazzo, B. Meini, *Numerical Solution of Algebraic Riccati Equations*, SIAM, Philadelphia, 2012.

[3] T.-M. Huang, R.-C. Li, W.-W. Lin, *Structure-Preserving Doubling Algorithms for Nonlinear Matrix Equations*, SIAM, Philadelphia, 2018.

[4] F. De Terán, B. Iannazzo, F. Poloni, L. Robol, *Nonsyngular systems of generalized Sylvester equations: an algorithmic approach*, Numer. Linear Algebra Appl., 26-5 (2019) e2261, 29.

[5] F. M. Dopico, J. Gonzáles, D. Kressner, V. Simoncini, *Projection methods for large-scale T-Sylvester equations*, Math. Comput., 85 (2016) pp. 2427-2455.

[6] D. Kressner, C. Schröder, D. S. Watkins, *Implicit QR algorithm for palindromic and even eigenvalue problems*, Numer. Algorithms, 51-2 (2009) pp. 209-238.

[7] D. S. Mackey, N. Mackey, C. Mehl, V. Mehrmann, *Numerical methods for palindromic eigenvalue problems: computing the anti-triangular Schur form*, Numer. Linear Algebra Appl., 16-1 (2009) pp. 63-86.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Optimal Constant for Generalized Diagonal Update Method

Contributing Author *Jeong-Hoon Ju* Young-Jin Kim, Hyun-Min Kim

Department of Mathematics, Pusan National University, Republic of Korea

The diagonal update method can be used in the Bernoulli's method to solve a quadratic matrix equation $AX^2 + BX + C = 0$, and it has better results on iteration number and time than the pure Bernoulli's method [1]. In this talk, we suggest the optimal constant which extends the sufficient condition to use the diagonal update method and guarantees the monotone convergence. Moreover, with some numerical experiments, we also compare the number of iterations defined by the generalized diagonal update method and the pure Bernoulli's method. Furthermore, we show that this generalized diagonal update method is useful to solve matrix equations with the form $AX^2 + \epsilon BX + C = 0$.

In detail, we consider the following quadratic matrix equation

$$Q_1(X) = AX^2 + BX + C = 0 \tag{1}$$

where

$A \in \mathbb{R}^{n\times n}$ is a diagonal matrix with positive diagonal elements,
$B \in \mathbb{R}^{n\times n}$ is a nonsingular M-matrix,
$C \in \mathbb{R}^{n\times n}$ is an M-matrix such that $B^{-1}C$ is nonnegative.

The equation (1) was motivated by a quadratic eigenvalue problem arising from an overdamped vibrating system [4]. In order to improve the pure Bernoulli's method in [2] and the diagonal update method in [1], we suggest the optimal constant $\gamma^* := \min\{\text{real}(\text{eig}(B - C)), 2\}$ and the generalized diagonal update skills:

$$\mathcal{G}_\gamma(X) = -(B + X - (\gamma - 1)\delta_X I)^{-1}(C + (\gamma - 1)\delta_X X), \tag{2}$$
$$\mathcal{H}_\gamma(X) = -(B - \gamma\delta_X I)^{-1}(X^2 + \gamma\delta_X X + C), \tag{3}$$

where $\delta_X = \min\{1, \min\{|\text{diag}(X)|\}\}$ and $1 \le \gamma < \gamma^*$. When $B - C - I$ is a nonsingular M-matrix, we can prove that both Berboulli's iterations defined by (2) and (3) with $X_0 = 0$ converge to the primary solvent X^*, by using some properties of M-matrices which are in [3].

Furthermore, we consider a numerical example with the following $m \times m$ coefficient matrices:

$$A = I,\ B = \epsilon \begin{bmatrix} 20 & -10 & & & & \\ -10 & 30 & -10 & & & \\ & -10 & 30 & -10 & & \\ & & -10 & \ddots & \ddots & \\ & & & \ddots & 30 & -10 \\ & & & & -10 & 20 \end{bmatrix}, C = \begin{bmatrix} 15 & -5 & & & & \\ -5 & 15 & -5 & & & \\ & -5 & 15 & -5 & & \\ & & -5 & \ddots & \ddots & \\ & & & \ddots & 15 & -5 \\ & & & & -5 & 15 \end{bmatrix}$$

which are motivated by [1] and [5], where $\epsilon \in \mathbb{R}$. We used the following algorithms:

$$\begin{cases} X_0 = 0,\ \delta_X = \min\{1, \min\{|\mathrm{diag}(X)|\}\},\ \gamma = \min\{\mathrm{real}(\mathrm{eig}(B-C)), 2\} - 0.0001 \\ X_{i+1} = -(B + X_i)^{-1}C, \qquad \text{(BI1)} \\ X_{i+1} = -(B + X_i - (\gamma-1)\delta_{X_i} I)^{-1}(C + (\gamma-1)\delta_{X_i} X_i). \qquad \text{(BI1-OC)} \end{cases}$$

$$\begin{cases} X_0 = 0,\quad \delta_X = \min\{1, \min\{|\mathrm{diag}(X)|\}\},\ \gamma = \min\{\mathrm{real}(\mathrm{eig}(B-C)), 2\} - 0.0001 \\ X_{i+1} = -B^{-1}(X_i^2 + C), \qquad \text{(BI2)} \\ X_{i+1} = -(B - \gamma\delta_{X_i} I)^{-1}(X_i^2 + \gamma\delta_{X_i} X_i + C). \qquad \text{(BI2-OC)} \end{cases}$$

When $\epsilon = 0.95$, we can not use the original diagonal update method in [1], because $B - C - 2I$ is not a nonsingular M-matrix. Fortunately, if we use the generalized diagonal update method with $\gamma = 1.8683$, we have the good results as follows:

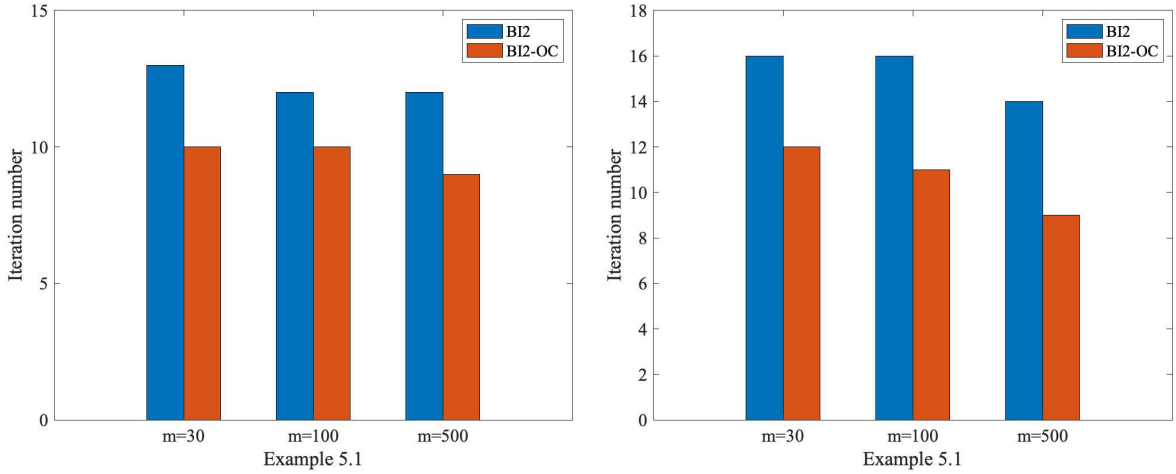


Figure 1: Comparison of iteration number with the methods BI1, BI1-OC (left), and BI2, BI2-OC (right)

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

On Newton Method for the Minimal Positive Solution of a System of Multi-Variable Nonlinear Matrix Equations

Contributing Author *Taehyeong Kim* Hyun-Min Kim

Department of Mathematics, Pusan National University, Busan 46241, Republic of Korea

In this study, we consider the minimal positive solution of the following system of the multi-variable nonlinear matrix equations that can be expressed in the form

$$\begin{cases} A_{1,n}X_1^n + A_{1,n-1}X_2^{n-1} + \cdots + A_{1,2}X_{n-1}^2 + A_{1,1}X_n + A_{1,0} = 0, \\ A_{2,n}X_2^n + A_{2,n-1}X_3^{n-1} + \cdots + A_{2,2}X_n^2 + A_{2,1}X_1 + A_{2,0} = 0, \\ \qquad\qquad\qquad\qquad \vdots \\ A_{n,n}X_n^n + A_{n,n-1}X_1^{n-1} + \cdots + A_{2,2}X_{n-2}^2 + A_{n,1}X_{n-1} + A_{n,0} = 0 \end{cases} \tag{1}$$

where $X_i \in \mathbb{R}^{p\times p}$ are unknown matrices, $A_{i,j} \in \mathbb{R}^{p\times p}$ for $i = 1, 2, \ldots, n$ and $j = 0, 1, \ldots, n$. We give the following assumptions on the coefficient matrices of the system (1):

For $i = 1, 2, \ldots, n$ and $j = 2, 3, \ldots, n$,
$A_{i,j}$ is a positive matrix or a nonnegative irreducible matrix,
$-A_{i,1}$ is nonsingular M-matrix,
$A_{i,0}$ is a positive matrix.

For $j = 0, 1, \ldots, n$, set the coefficient matrices A_j, unknown matrix Y and the permutation matrix P, then the system (1) can be equivalently reformulated as

$$\begin{aligned} F(Y) &= A_nY^n + A_{n-1}P^\top Y^{n-1}P + \cdots + A_1(P^\top)^{n-1}YP^{n-1} + A_0 \\ &= \sum_{j=0}^{n} A_j(P^\top)^{n-j}Y^jP^{n-j} = 0. \end{aligned} \tag{2}$$

Newton's iteration for solving equation (2) can be stated as

$$\begin{cases} D_{Y_i}(H_i) = -F(Y_i), \\ Y_{i+1} = Y_i + H_i. \end{cases} \quad i = 1, 2, \dots$$
$$\text{where } D_Y(H) = \sum_{p=1}^{n} \left(A_p \left(P^\top \right)^{n-p} \left(\sum_{q=0}^{p-1} Y^q H Y^{p-q-1} \right) (P)^{n-p} \right) \tag{3}$$

Note that the matrices in equation (3) is of size $np \times np$, which implies that the computation cost of the classical Newton's iteration is expensive if n and p are large. To reduce the computation cost, we propose the modified Newton's iteration as follows:

$$\begin{cases} A_{1,n}\Gamma_n^{(1,i)}(H_{1,i}) + A_{1,n-1}\Gamma_{n-1}^{(2,i)}(H_{2,i}) + \cdots + A_{1,1}H_{n,i} = -F_1(X_{1,i}, \dots, X_{n,i}), \\ A_{2,n}\Gamma_n^{(2,i)}(H_{2,i}) + A_{2,n-1}\Gamma_{n-1}^{(3,i)}(H_{3,i}) + \cdots + A_{2,1}H_{1,i} = -F_2(X_{1,i}, \dots, X_{n,i}), \\ \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ A_{n,n}\Gamma_n^{(n,i)}(H_{n,i}) + A_{n,n-1}\Gamma_{n-1}^{(1,i)}(H_{1,i}) + \cdots + A_{n,1}H_{n-1,i} = -F_n(X_{1,i}, \dots, X_{n,i}), \\ X_{1,i+1} = X_{1,i} + H_{1,i}, \\ X_{2,i+1} = X_{2,i} + H_{2,i}, \\ \qquad \vdots \qquad\qquad\qquad\qquad\qquad\qquad i = 1, 2, \dots \\ X_{n,i+1} = X_{n,i} + H_{n,i}, \\ \text{where } \Gamma_k^{(j,i)}(H) = \sum_{p=1}^{k} X_{j,i}^{p-1} H X_{j,i}^{k-p} \text{ for } j = 1, 2, \dots, n. \end{cases} \tag{4}$$

Set $X_{j,0} = 0$ for $j = 1, 2, \dots, n$, we prove that the sequences $\{X_{j,i}\}$ generated by (4) converge to the minimal positive solution of system (1). And some numerical experiments are given to show the efficiency of the modified Newton's iteration in calculation time and memory.

References

[1] G.J. Davis, *Algorithm 598: An algorithm to compute solvents of the matrix equation* $AX^2 + BX + C = 0$, ACM Trans. Math. Software, 9:246-254, (1983).

[2] N.J. Higham , H.-M. Kim, *Solving a duadratic matrix equation by Newton's method with exact line searches*, SIAM J. Matrix. Anal. Appl., 23:303-316, (2006).

[3] R.A. Horn , C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press,1994.

[4] H.-M. Kim, *Convergence of Newton's method for solving a class of quadratic matrix equations*, Honam Mathematic J., 30:399-409, (2008).

[5] J. Meng , S.-H. Seo , H.-M. Kim, *Condition numbers and backward error of a matrix polynomial equation arising in stochastic models*, J. Sci. Comput., 76:759-776, (2018).

[6] J.-H. Seo , H.-M. Kim, *Convergence of pure and relaxed Newton methods for solving a matrix polynomial equation arising in stochastic models*, Linear Algebra and its Applications, 440:34-49, (2014).

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

A matrix-oriented POD-DEIM algorithm applied to semilinear differential matrix equations

Gerhard Kirsten Valeria Simoncini

Università di Bologna

We are interested in numerically approximating the solution $\boldsymbol{U}(t) \in \mathcal{S}$ to the following semilinear matrix differential equation

$$\dot{\boldsymbol{U}}(t) = \boldsymbol{A}\boldsymbol{U}(t) + \boldsymbol{U}(t)\boldsymbol{B} + \mathcal{F}(\boldsymbol{U}, t), \quad \boldsymbol{U}(0) = \boldsymbol{U}_0, \tag{1}$$

where $\boldsymbol{A} \in \mathbb{R}^{n_x \times n_x}, \boldsymbol{B} \in \mathbb{R}^{n_y \times n_y}$, and $t \in [0, T_f] = \mathcal{T} \subset \mathbb{R}$, equipped with appropriate boundary conditions. The function $\mathcal{F} : \mathcal{S} \times \mathcal{T} \to \mathbb{R}^{n_x \times n_y}$ is a sufficiently regular nonlinear function that can be evaluated elementwise, and $\mathcal{S}$ is a functional space containing the sought after solution.

The problem (1) arises for instance in the discretization of two-dimensional partial differential equations of the form

$$u_t = \ell(u) + f(u, t), \quad u = u(x, y, t) \quad \text{with } (x, y) \in \Omega \subset \mathbb{R}^2,\ t \in \mathcal{T}, \tag{2}$$

and given initial condition $u(x, y, 0) = u_0(x, y)$, for certain choices of the physical domain Ω. The differential operator ℓ is linear in u, typically a second order operator in the space variables, while $f : S \times \mathcal{T} \to \mathbb{R}$ is a nonlinear function, where S is an appropriate space with $u \in S$.

In this talk we present a matrix-oriented model order reduction strategy for the problem (1) that leads to a semilinear *matrix* differential equation with the same structure as (1), but of significantly reduced dimension. More precisely, we determine an approximation to $\boldsymbol{U}(t)$ of the type

$$\boldsymbol{V}_{\ell,U} \boldsymbol{Y}_k(t) \boldsymbol{W}_{r,U}^\top, \quad t \in [0, T_f], \tag{3}$$

where $\boldsymbol{V}_{\ell,U} \in \mathbb{R}^{n \times k_1}$ and $\boldsymbol{W}_{r,U} \in \mathbb{R}^{n \times k_2}$ are matrices to be determined, independent of time. Here $k_1, k_2 \ll n$ and we let $k = (k_1, k_2)$. The function $\boldsymbol{Y}_k(t)$ is determined as the numerical solution to the following *reduced* semilinear matrix differential problem

$$\begin{aligned} \dot{\boldsymbol{Y}}_k(t) &= \boldsymbol{A}_k \boldsymbol{Y}_k(t) + \boldsymbol{Y}_k(t) \boldsymbol{B}_k + \widehat{\mathcal{F}_k(\boldsymbol{Y}_k, t)} \\ \boldsymbol{Y}_k(0) &= \boldsymbol{Y}_k^{(0)} := \boldsymbol{V}_{\ell,U}^\top \boldsymbol{U}_0 \boldsymbol{W}_{r,U}, \end{aligned} \tag{4}$$

with $\boldsymbol{A}_k = \boldsymbol{V}_{\ell,U}^\top \boldsymbol{A}\boldsymbol{V}_{\ell,U}$, $\boldsymbol{B}_k = \boldsymbol{W}_{r,U}^\top \boldsymbol{B}\boldsymbol{W}_{r,U}$, and $\widehat{\mathcal{F}_k(\boldsymbol{Y}_k, t)}$ is a matrix-oriented approximation to

$$\mathcal{F}_k(\boldsymbol{Y}_k, t) = \boldsymbol{V}_{\ell,U}^\top \mathcal{F}(\boldsymbol{V}_{\ell,U}\boldsymbol{Y}_k\boldsymbol{W}_{r,U}^\top, t)\boldsymbol{W}_{r,U}. \tag{5}$$

By stacking the columns of the matrix $\boldsymbol{U}(t)$ one after the other into a long vector, a collection of existing approaches typically map the matrix-valued problem (1) to a vector-valued dynamical system, for which order reduction is a well-established procedure. Among various methods, the Proper Orthogonal Decomposition (POD) [5, 2] methodology has been widely employed. The overall effectiveness of the POD procedure is largely influenced by the capability of evaluating the nonlinear term within the reduced space, motivating a considerable amount of work towards this estimation. One very successful approach is the Discrete Empirical Interpolation Method (DEIM) [3], which is based on the Empirical Interpolation Method (EIM) originally introduced in [1].

However, a shortcoming of these vectorization procedures is the massive computational and storage demand in the offline phase. Even in the online phase, several vectors of length $N = n_x n_y$ need to be stored to lift the low-dimensional functions back to the full dimension. Here we address precisely this shortcoming, focusing on POD for dimension reduction and DEIM for interpolation of the nonlinear function.

To this end, we devise a matrix-oriented POD approach tailored towards the construction of the matrix reduced problem formulation (4). An adaptive procedure is also developed to limit the number of snapshots contributing to the generation of the approximation spaces. The reduction of the nonlinear term is then performed by means of a fully matricial interpolation using left and right projections onto two distinct reduction spaces, giving rise to a new two-sided version of DEIM. By maintaining a matrix-oriented reduction, we are able to employ first order exponential integrators at negligible costs. Numerical experiments on benchmark problems illustrate the effectiveness of the new setting [4].

References

[1] M. Barrault, Y. Maday, N. C. Nguyen and A. T. Patera, *An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations*, C. R. Math. Acad. Sci. Paris, (2004) pp. 667–672.

[2] P. Benner, V. Mehrmann and D. Sorensen, *Dimension reduction of large-scale systems*, Springer-Verlag, Berlin/Heidelberg, Germany, 2005.

[3] S. Chaturantabut and D. C. Sorensen, *Nonlinear model reduction via discrete empirical interpolation*, SIAM J. Sci. Comput., (2010) pp. 2737–2764.

[4] G. Kirsten and V. Simoncini, *A matrix-oriented POD-DEIM algorithm applied to semilinear differential matrix equations*, arXiv preprint arXiv:2006.13289, (2021).

[5] K. Kunisch and S. Volkwein, *Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition*, J. Optim. theory Appl., (1999) pp. 345–371.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Low-rank tensor uniformization for tumor progression modeling

Peter Georg Lars Grasedyck *Maren Klever* Rudolf Schill
Rainer Spang Tilo Wettig

Institut für Geometrie und Praktische Mathemaik, RWTH Aachen University

Inspired by current research in tumor progression modeling, we consider continuous-time Markov chains that describe interacting processes. In general a Markov chain is defined by its state space S, its initial probability distribution $\mathbf{p}(0) \in \mathbb{R}^S$ and its transition rate matrix $\mathbf{Q} \in \mathbb{R}^{S\times S}$. In tumor progression models, a tumor is identified by its genotype where d mutations may have occurred or not. The Markov chain starts with the absence of all mutations and then gradually accumulates mutations. Its state space S describes the set of possible tumors and thus, grows exponentially in the number of mutations d with $|S| = 2^d$. Modeling the transition rate matrix $\mathbf{Q}$ leads to an optimization problem based on given tumor data. Since the age of a tumor at its diagnosis is typically unknown, this optimization requires the so-called *time-marginal distribution* $\mathbf{p} \in \mathbb{R}^S$ which is defined as

$$\mathbf{p} := \int_0^\infty \exp(-t) \cdot \exp\left(\mathbf{Q}t\right) \mathbf{p}(0)\ \mathrm{d}t \qquad \text{where} \quad t \sim \mathrm{Exp}[1].$$

The time-marginal distribution is obtained as the solution of the linear system

$$(\mathbf{Id} - \mathbf{Q})\mathbf{p} = \mathbf{p}(0). \tag{1}$$

However, the size of this system renders classical solvers infeasible.

In this talk, we introduce the concept of *Mutual Hazard Networks* [3] which allows for a low-rank tensor representation of the transition rate matrix $\mathbf{Q}$. Using low-rank tensor techniques reduces the computational complexity from exponential to linear provided the distribution $\mathbf{p}$ exhibits a low-rank structure. Previously known iterative methods also allow for low-rank approximations of the solution for (1) but are unable to guarantee that its entries sum up to one, i.e,

$$\sum_{x\in S} \mathbf{p}_x = 1, \tag{2}$$

as required for a probability distribution. We present a convergent iterative method based on the concept of uniformization [2]. This strategy allows us to use low-rank tensor formats and, at the same time, to satisfy condition (2). Numerical experiments illustrate that the time-marginal distribution is well approximated with low rank.

References

[1] P. Georg, L. Grasedyck, M. Klever, R. Schill, R. Spang, T. Wettig, *Low-rank tensor methods for Markov chains with applications to tumor progression models*, arXiv:2006.08135, (2020).

[2] W. K. Grassmann, *Transient solutions in markovian queueing systems*, Computers & Operations Research, (1977), pp. 47-53.

[3] R. Schill, S. Solbrig, T. Wettig, R. Spang, *Modelling cancer progression using Mutual Hazard Networks*, Bioinformatics, (2019), pp. 241-249.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

A Successive Divide and Conquer Algorithm for the Unilateral Quadratic Matrix Equation

Peter Benner *Martin Köhler*

Max Planck Institute for Dynamics of Complex Technical Systems,
Sandtorstr. 1, D-39106 Magdeburg, Germany

The Unilateral Quadratic Matrix Equation (UQME)

$$AX^2 + BX + C = 0 \tag{1}$$

is a rarely mentioned equation in numerical linear algebra. Only a few publications and algorithms cover the solution of this equation. Beside a direct solved based on the underlying quadratic eigenvalue problem, only the Bernoulli Iteration and Newton's Method are known [2, 3] for general matrices A, B, and C.

In our contribution, we develop an algorithm using the underlying eigenvalue problem

$$Fy = \begin{bmatrix} 0 & I \\ -C & -B \end{bmatrix} y = \lambda \begin{bmatrix} I & 0 \\ 0 & A \end{bmatrix} y = \lambda G y, \quad y = \begin{bmatrix} v \\ \lambda v \end{bmatrix}, \tag{2}$$

and the connection that if X is a solution of the UQME (1), it can be written as

$$X = Z_{21} Z_{11}^{-1}. \tag{3}$$

Thereby, the matrices Z_{21} and Z_{11} originate from the generalized Schur decomposition of the matrix pair (F, G):

$$(Q^H F Z, Q^H G Z) = \left(\begin{bmatrix} T_{11} & T_{12} \\ & T_{22} \end{bmatrix}, \begin{bmatrix} S_{11} & S_{12} \\ & S_{22} \end{bmatrix} \right) = (T, S) \tag{4}$$

with

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \text{ and } Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}, \tag{5}$$

where each block is of size $m \times m$. Since there is no predefined order of the eigenvalues on the diagonals of (T, S), a variety of solutions X exists. Typically, the *minimal* solution, which corresponds to the m smallest eigenvalues, or the *dominant* solution, derived using the deflating subspace related to the m largest eigenvalues, are of interest. Obtaining these solutions with the computation of the generalized Schur decomposition (4) and a

subsequent eigenvalue reordering is a time-consuming procedure due to the complexity and the behavior of the QZ algorithm. Obviously, it is sufficient to compute the block decomposition as shown in Equation (4), *without* having (T_{11}, S_{11}) and (T_{22}, S_{22}) in (quasi) upper triangular form. Only the zero block in the lower left of (4) needs to be of size $m \times m$. Furthermore, if we assume that we want to compute the minimal solution, the blocks of (T, S) have to fulfill the following condition:

$$|\lambda| < |\mu| \quad \forall \lambda \in \Lambda\left(T_{11}, S_{11}\right), \mu \in \Lambda\left(T_{22}, S_{22}\right). \tag{6}$$

Our Successive Divide and Conquer algorithm (SDC) computes a sequence of unitary transformations (Q_k, Z_k) with

$$Q = Q_1 \cdots Q_p \quad \text{and} \quad Z = Z_1 \cdots Z_p \tag{7}$$

until the block partitioning (4)/(5) fulfills the condition (6) and (T_{11}, S_{11}) are of size $m \times m$. Thereby, the sequence of transformation matrices (Q_k, Z_k) is computed from deflating subspaces, which are computed with the help of the matrix disc function [1] and a dedicated scaling technique [4]. In this way, the algorithm successively transforms the matrix pair (F, G) as long as necessary.

If the UQME (1) is given with real data and the desired solution should be real as well, a minimal solution does not exist if there a complex conjugate eigenvalue pair would be split when computing the block decomposition (4). In order to overcome this issue, we relax this to computing the quasi minimal solution, which differs from the minimal solution in one missing real eigenvalue.

The numerical experiments show that our algorithm is a fast replacement for solving the problem with the QZ algorithm. Even with a pure naive MATLAB implementation of our algorithm, we obtain a speed up of 5.5 to 10 compared to the QZ algorithm. The comparison to the Bernoulli iteration strongly depends on its convergence behavior. Especially, if the gap between the spectra of (T_{11}, S_{11}) and (T_{22}, S_{22}) in (4) gets very small, our algorithms comes up with a much faster solution. Here, we obtain a speed up of 1.25 with a naive MATLAB implementation of the SDC algorithm. Finally, we show an example for the quasi minimal solution, where the Bernoulli iteration does not converge at all.

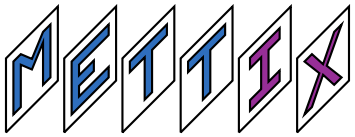
References

[1] P. Benner. *Contributions to the Numerical Solution of Algebraic Riccati Equations and Related Eigenvalue Problems*. Dissertation, Fakultät für Mathematik, TU Chemnitz–Zwickau, 09107 Chemnitz (Germany), February 1997.

[2] N. J. Higham and H.-M. Kim. Numerical analysis of a quadratic matrix equation. *IMA J. Numer. Anal.*, 20(4):499–519, 2000.

[3] N. J. Higham and H.-M. Kim. Solving a quadratic matrix equation by Newton's method with exact line searches. *SIAM J. Matrix Anal. Appl.*, 23(2):303–316, 1 2001.

[4] M. Köhler. *Approximate Solution of Non-Symmetric Generalized Eigenvalue Problems and Linear Matrix Equation on HPC Platforms*. Dissertation, Department of Mathematics, Otto-von-Guericke University, Magdeburg, Germany, 2021.



Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Inexactness in low-rank methods for large matrix equations

Patrick Kürschner

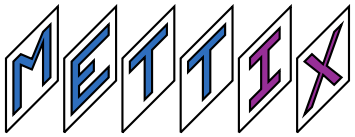
Leipzig University of Applied Sciences (HTWK Leipzig),
Centre for Mathematics and Natural Sciences ,
Gustav-Freytag-Straße 42a, 04277 Leipzig

One established strategy to solve large-scale algebraic matrix equations is to compute a low-rank approximations of the solution by iterative methods. Often, these methods work internally with rational Krylov subspaces which necessitates to solve a large and sparse linear system of equations inside each iteration step. In this presentation we will look at inexact variants of rational Krylov subspace and low-rank ADI methods [1], where "inexact" refers to the scenario when the occurring *inner* linear systems are solved inexactly by iterative methods such as, e.g., preconditioned Krylov subspace methods. Our main focus are estimates for the required accuracy regarding the inner linear systems which will dictate when the *inner* Krylov subspace methods can be terminated, thus potentially saving some computational effort without endangering the functionality of the *outer* low-rank method. If time permits, we will also look at inexact matrix-valued low-rank Krylov methods [2] for general linear matrix equations, where inexactness occurs in the form of low-rank truncations and similar considerations as before can be done.

References

[1] P. Kürschner, M. Freitag, *Inexact methods for the low rank solution to large scale Lyapunov equations*, BIT Numerical Mathematics. 60:1221–1259, 2020

[2] D. Palitta, P. Kürschner, *On the convergence of Krylov methods with low-rank truncations*, Numerical Algorithms, 2021



Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Recent Progresses on Highly Entrywise Accurate Methods For Matrix Equations

Ren-Cang Li

Department of Mathematics, University of Texas at Arlington, Texas, USA

In 2002, Alfa, Xue, and Ye showed that the inverse of a nonsingular M-matrix can be determined to highly relative entrywise accuracy by a triplet representation of the M-matrix, and devised the so-called GTH-like algorithm, a variant of Gaussian elimination, to deliver a numerical inverse with comparable entrywise relative accuracy. The breakthrough form the foundation of later developments in numerical solutions of the M-matrix algebraic Riccati equation (MARE) and the Quasi-Birth-and-Death (QBD) equation with guaranteed high relative entrywise accuracy. In this talk, we will survey those developments, including recent ones on the shifted M-matrix algebraic Riccati equation and the structured M-matrix algebraic Riccati equation.

References

[1] A. S. Alfa, J. Xue, and Q. Ye. Accurate computation of the smallest eigenvalue of a diagonally dominant M-matrix. *Math. Comp.*, 71:217–236, 2002.

[2] A. S. Alfa, J. Xue, and Q. Ye. Entrywise perturbation theory for diagonally dominant M-matrices with applications. *Numer. Math.*, 90(3):401–414, 2002.

[3] Guiding Gu, Li Wang, and Ren-Cang Li. Highly accurate latouche-ramaswami logarithmic reduction algorithm for Quasi-Birth-and-Death process. *J. Math. Study*, 2022. to appear.

[4] Changli Liu, Wei-Guo Wang, Jungong Xue, and Ren-Cang Li. Accurate numerical solution for structured M-matrix algebraic Riccati equations. *J. Comput. Appl. Math.*, 396(113614), 2021.

[5] Changli Liu, Jungong Xue, and Ren-Cang Li. Accurate numerical solution for shifted M-matrix algebraic Riccati equations. *J. Sci. Comput.*, 84(15), 2020.

[6] G. T. Nguyen and F. Poloni. Componentwise accurate fluid queue computations using doubling algorithms. *Numer. Math.*, 130(4):763–792, 2015.

[7] Jungong Xue and Ren-Cang Li. Highly accurate doubling algorithms for M-matrix algebraic Riccati equations. *Numer. Math.*, 135(3):733–767, 2017.

[8] Qiang Ye. On Latouche-Ramaswami's logarithmic reduction algorithm for quasi-birth-and-death processes. *Stoch. Models*, 18:449–467, 2002.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Hierarchical adaptive low-rank format with applications to discretized PDEs

Daniel Kressner *Stefano Massei* Leonardo Robol

Technische Universiteit Eindhoven (TU/e)

Low-rank based data compression can sometimes lead to a dramatic acceleration of numerical simulations. A striking example is the solution of time dependent elliptic PDEs of the form:

$$\begin{cases} \frac{\partial u}{\partial t} = Lu + f(t, u, \nabla u) & (x, y) \in \Omega, \quad t \in [0, T_{\max}] \\ u(x, y, 0) = u_0(x, y) \end{cases} \tag{1}$$

where $\Omega \subset \mathbb{R}^2$ is a rectangular domain, L is a linear differential operator, f is nonlinear and (1) is coupled with appropriate boundary conditions in space. When the source term and the solution are smooth, their (structured) discretizations lead to matrices that allow for excellent low-rank approximations. Under suitable assumptions on the differential operator, one can recast the corresponding discretized PDE as a matrix equation [1, 2]. In turn, this yields the possibility to facilitate efficient algorithms for matrix equations with low-rank right-hand side [3, 4]. However, in many situations of interest the smoothness property is not present in the whole domain. A typical instance are solutions that feature singularities along curves, while being highly regular elsewhere. This renders a global low-rank approximation ineffective. During the last decades, there has been significant effort in developing hierarchical low-rank formats that apply low-rank approximation only locally [5]. These formats recursively partition the matrix into blocks that are either represented as a low-rank matrix or are sufficiently small to be stored as a dense matrix. These techniques are usually applied in the context of operators with a discretization known to feature low-rank off-diagonal blocks, such as integral operators with singular kernel.

We propose a new format that automatically adapts the choice of the hierarchical partitioning and the location of the low-rank blocks without requiring the use of an admissibility criterion. The admissibility is decided on the fly by the success or failure of low-rank approximation techniques. We call this format *Hierarchical Adaptive Low-Rank (HALR)* matrices.

We discuss techniques for the efficient adaptation of the structure in case of moving singularities, with the aim of tracking the time-evolution of the solution of (1); the numerical tests demonstrate that the proposed techniques can effectively detect changes in the structure, and ensure the desired level of accuracy. We develop efficient Lyapunov

and Sylvester solvers for matrix equations with HALR right-hand-side and *Hierarchically off-diagonal low-rank* (HODLR) coefficients. Several numerical experiments demonstrate both the effectiveness and the flexibility of the approach.

References

[1] D. Palitta, V. Simoncini, *Matrix-equation-based strategies for convection–diffusion equations*, BIT Numerical Mathematics, (2016) pp. 751-776.

[2] A. Townsend, S. Olver, *The automatic solution of partial differential equations using a global spectral method*, Journal of Computational Physics, (2015) pp. 106-123.

[3] P. Benner, R. C. Li, N. Truhar, *On the ADI method for Sylvester equations*, Journal of Computational and Applied Mathematics, (2009) pp. 1035-1045.

[4] V. Druskin, L. Knizhnerman, V. Simoncini, *Analysis of the rational Krylov subspace and ADI methods for solving the Lyapunov equation*, SIAM Journal on Numerical Analysis, (2011) pp. 1875-1898.

[5] W. Hackbusch, *Hierarchical matrices: algorithms and analysis*, Springer, Heidelberg, 2015.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Model reduction and utilization of the reduced basis for aggregation kinetic equations

Ivan Timokhin *Sergey A. Matveev* Alexander Smirnov Eugene Tyrtyshnikov

Marchuk Institute of Numerical Mathematics RAS, Gubkin st 8, 119333, Moscow

In this work we apply Proper Orthogonal Decomposition (POD) for model reduction in application to kinetic equations of irreversible aggregation process with source of monomers:

$$\frac{dn_k}{dt} = J_k + \frac{1}{2}\sum_{i+j=k} C_{i,j} n_i n_j - n_k \sum_{j=1}^{N} C_{j\,k} n_j, \qquad k = \overline{1,N} \tag{1}$$

We assume the existence of an orthonormal basis, gathered as columns of a matrix $V \in \mathbb{R}^{N\times R}$, such that

$$\left\| n(t) - VV^T n(t) \right\| \ll \left\| n(t) \right\|, \tag{2}$$

where $n(t)$ is the solution of the original system and further introduce the reduced solution

$$x(t) \equiv V^T n(t), \qquad x(t) \in \mathbb{R}^R \tag{3}$$

which can be evaluated within $O(R^3)$ operations instead of $O(N \log N)$ required for the original system. In case $R \ll N$ evaluation of the reduced solution requires much less operations and depends only on the dimensionality of the basis. Final reconstruction of the original solution can be done as

$$n(t) \approx \tilde{n}(t) = V\tilde{x}(t). \tag{4}$$

Thus, the main problem is choice of algorithm for construction of the target basis allowing to perform the reduction of dimensionality. In our work we show that POD exploiting the method of "snapshots" allows to obtain the basis and in our experiments we show that $R \ll N$.

All in all, we also show that utilization of such basis allows to perform a speedup of compuations for aggregation kinetic equations without significant loss of accuracy of the solutions. At the same time, we also demonstrate problematic sides of the chosen approach – the control of the precision of the reduced solution seems to be not an easy task due to the nonlinearity of both method and model and requires additional studies in future.

This work has been supported by Russian Science Foundation Project (21-71-10072).

References

[1] I. V. Timokhin, E. E. Tyrtyshnikov, A. P. Smirnov, *Model reduction in Smoluchowski-type equations*, arXiv preprint arXiv:2008.09439 (2020)

[2] I. V. Timokhin, E. E. Tyrtyshnikov, A. P. Smirnov, *Method for Reduced Basis Discovery in Nonstationary Problems*, Doklady Mathematics, Vol. 103 No. 2 (2021)

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Solution of the matrix equation for geodesics associated with the Riemannian metric on the space of positive-definite matrices based on the power potential

Nadia Chouaieb Bruno Iannazzo *Maher Moakher*

LAMSIN, ENIT, University of Tunis El Manar, B.P. 37, 1002, Tunis, Tunisia

In this work, we give a closed-form expression for the nonlinear second-order differential equation of geodesic curves associated with the Riemannian metric given by the Hessian of the power potential function on $\mathcal{P}_n$, the space of positive-definite matrices of order n. The β-power potential function on $\mathcal{P}_n$ is [2]

$$\Phi_\beta(X) = \frac{1-(\det X)^\beta}{\beta}, \quad \beta \neq 0. \tag{1}$$

It generalizes the logarithmic potential in the sense that $\lim_{\beta\to 0} \Phi_\beta(X) = -\ln\det(X)$.

For $\beta < \frac{1}{n}$, the Hessian of (1) is positive definite, and hence it provides at each point $X \in \mathcal{P}_n$ a one-parameter family of Riemannian metrics on $\mathcal{P}(n)$ given by

$$g_{\beta,X}(U,V) := (\det X^\beta)\big(\mathrm{tr}(X^{-1}UX^{-1}V) - \beta\,\mathrm{tr}(X^{-1}U)\,\mathrm{tr}(X^{-1}V)\big), \tag{2}$$

where U and V are points of the tangent space to $\mathcal{P}_n$ at X, identified as usual with the space of symmetric matrices of order n.

A geodesic curve $\{X(t),\ t \in [0,1]\}$ with respect to the Riemannian metric (2) satisfies the second-order matrix differential equation

$$\frac{d}{dt}\left(\frac{\partial g_{\beta,X}(X',X')}{\partial X'}\right) - \frac{\partial g_{\beta,X}(X',X')}{\partial X} = O_n. \tag{3}$$

Theorem 1 ([1]). *Let $X : [0,1] \to \mathcal{P}(n)$ be a smooth geodesic on $\mathcal{P}(n)$ equipped with the Riemannian metric (2). Then, by introducing the matrix function $G(t) = X^{-1}(t)X'(t)$, the second-order ODE (3) can be written as the decoupled first-order system for X and G:*

$$G' = \frac{\beta}{2(1-n\beta)}\left(\mathrm{tr}(G^2) - \beta\,\mathrm{tr}^2(G)\right) I - \beta\,\mathrm{tr}(G)G, \tag{4a}$$

$$X' = XG. \tag{4b}$$

It is worthy to note that (4.a) is a nonlinear (quadratic) ODE for $G(t)$. Once $G(t)$ is obtained, the linear first-order ODE (4.b) can be solved for $X(t)$.

We show that, under some conditions on β, there exists a unique geodesic curve for the metric (2) joining two positive-definite matrices A and B and we provide an explicit expression for this geodesic.

Before we state our main result, let us define the following measure of linear independence between two symmetric positive definite matrices A and B

$$\gamma_\beta(A,B) := \frac{|\beta|\delta(\det(A)^{-1/n}A, \det(B)^{-1/n}B)}{2\sqrt{1/n-\beta}}, \tag{5}$$

where $\delta(\cdot,\cdot)$ is the Riemannian distance on $\mathcal{P}_n$ given for any two matrices $M, N \in \mathcal{P}_n$ by $\delta(M,N) := \left(\sum_{i=1}^n \ln^2 \lambda_i\right)^{1/2}$, with $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $M^{-1}N$.

Theorem 2. *If $A, B \in \mathcal{P}(n)$ are linearly independent, set $d := \delta(\det(A)^{-1/n}A, \det(B)^{-1/n}B)$ and*

$$\beta_1 := -\pi\frac{\sqrt{\pi^2n^2+4nd^2}+\pi n}{2nd^2}, \quad \beta_2 := \pi\frac{\sqrt{\pi^2n^2+4nd^2}-\pi n}{2nd^2}.$$

Then, for $\beta \in (\beta_1, 0) \cup (0, \beta_2)$, there exists a unique geodesic joining A and B given by

$$G_\beta(A,B,t) = \eta(t)A(A^{-1}B)^{\alpha(t)}, \quad t \in [0,1], \tag{6}$$

where

$$\alpha(t) = \frac{1}{\gamma}\arctan\Big(\frac{t\sigma\sin\gamma}{1-t+t\sigma\cos\gamma}\Big), \quad \eta(t) = \Big(\frac{(1-t)^2+2t(1-t)\sigma\cos\gamma+t^2\sigma^2}{\sigma^{2\alpha(t)}}\Big)^{\frac{1}{n\beta}},$$

with $\sigma = \det(A^{-1}B)^{\beta/2}$ and $\gamma := \gamma_\beta(A,B)$.

The geodesic curve (6) has an exponential part, similar to that of the geometric mean, but with exponent $\alpha(t)$; and a scalar power part, $\eta(t)$, which reduces to the weighted $\frac{2}{n\beta}$-power mean when $\gamma = 0$.

When β goes to 0 then (6) becomes the matrix geometric mean

$$\lim_{\beta\to 0} G_\beta(A,B,t) = G_0(A,B,t) := A(A^{-1}B)^t.$$

Furthermore, if A and B are linearly dependent matrices in $\mathcal{P}(n)$, then (6) reduces to the matrix $\frac{n\beta}{2}$-power mean

$$G_\beta(A,B,t) = \big((1-t)A^{\frac{n\beta}{2}} + tB^{\frac{n\beta}{2}}\big)^{\frac{2}{n\beta}}.$$

References

[1] N. Chouaieb, B. Iannazzo, M. Moakher, *Geometries on the cone of positive-definite matrices derived from the power potential and their relation to the power means*, to appear in Linear Algebra and its Applications, (2021).

[2] A. Ohara, N. Suda, S. Amari, *Dualistic differential geometry of positive definite matrices and its applications to related problems*, Linear Algebra and its Applications, **247** (1996) pp. 31–53.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

On the low-rank approximations in the Chebyshev norm

Stanislav Morozov Nikolai Zamarashkin Eugene Tyrtyshnikov

Marchuk Institute of Numerical Mathematics
of the Russian Academy of Sciences, Moscow

Low-rank matrices and tensors are ubiquitous in science. A lot of methods have been developed using low-rank structures in computational mathematics [1], computational fluid dynamics [2], movie preferences [3] and automated machine learning [4]. To date, most of the methods have been developed to build effective low-rank approximations in the Frobenius norm. In the matrix case, the quality of such approximations depends on the decrease rate of the singular values of the matrix. However, recent results show that low-rank approximations of matrices in other norms can be effective even without decreasing singular values. One fundamental result was proved in [5]:

Theorem. Let $X \in \mathbb{R}^{m\times n}$ with $m \geq n$ and $0 < \varepsilon < 1$. Then, with

$$r = \lceil 72 \log (2n+1)/\varepsilon^2 \rceil \tag{1}$$

we have

$$\inf_{\text{rank}Y \leq r} \|X - Y\|_C \leq \varepsilon \|X\|_2, \quad \text{where} \quad \|X\|_C = \max_{i,j} |X_{ij}| \tag{2}$$

This work is devoted to low-rank approximations in the Chebyshev, that is, elementwise norm (2). For simplicity of presentation, the results are given in the matrix case, although some of them can be generalized to tensors. Let a matrix $A \in \mathbb{R}^{m\times n}$ be given. We call

$$\mu = \inf_{U\in\mathbb{R}^{m\times r}, V\in\mathbb{R}^{n\times r}} \left\|A - UV^T\right\|_C, \tag{3}$$

the value of the best Chebyshev approximation of rank r. To date, the problem of constructing low-rank approximations in the Chebyshev norm has been little studied. One of the few works devoted to this topic is [6], in which a method for finding local minima for rank 1 approximations has been proposed. The authors construct an approximation using the alternance method. Let $U^{(0)}$ be given. In the case of rank 1, it is easy to find a matrix (in fact, a vector) $V^{(0)}$, being the solution of the problem

$$\mu = \min_{V\in\mathbb{R}^{n\times r}} \left\|A - U^{(0)}V^T\right\|_C \tag{4}$$

Then the authors find the matrix (vector) $U^{(1)}$ that is optimal for $V^{(0)}$ and repeat the procedure.

The results presented in [6] can be generalized to the case of an arbitrary rank. The key step here is the possibility of solving the problem of the form (4), which becomes challenging for $r > 1$. This problem can be reduced to solving several independent problems of the form

$$\mu = \inf_{v \in \mathbb{R}^r} \|a - Uv\|_\infty, \quad a \in \mathbb{R}^m, \quad U \in \mathbb{R}^{m \times r} \tag{5}$$

Suppose that in the matrix $U \in \mathbb{R}^{m \times r}$ all submatrices of size $r \times r$ are non-singular. Let I_k be a subset of k indices from 1 to m, $I_k = \{i_1, i_2, \ldots, i_k\}$. Let us denote a_{I_k} the subvector of the vector a with elements from I_k, and U_{I_k} the submatrix of the matrix U with rows from I_k. It can be shown that there exists a subset I_{r+1} of $r + 1$ indices such that the solution of the problem

$$\widehat{\mu} = \inf_{v \in \mathbb{R}^r} \left\|a_{I_{r+1}} - U_{I_{r+1}} v\right\|_\infty \tag{6}$$

coincides with the solution of the problem (5). Such a subset I_{r+1} is called characteristic. The problem of size $r + 1$ can be solved exactly in $O(r^4)$ operations ([7]). The above arguments allow us to find the optimal solution to the problem (5) (and therefore (4)) by iterating over all subsets of $r + 1$ indices. However, more efficient methods for finding the characteristic set and solving the problem (4) can also be constructed based on generalizations of the Remez algorithm. Such algorithm does not require iterating over all subsets and in practice work in polynomial time.

This work has been supported by Russian Science Foundation Project (21-71-10072).

References

[1] M. Bebendorf, *Hierarchical Matrices: A Means to Efficiently Solve Elliptic Boundary Value Problems*, Lect. Notes Comput. Sci. Eng. 63, (2008)

[2] J. A. Tropp, A. Yurtsever, M. Udell, and V. Cevher, *More Practical Sketching Algorithms for Low-Rank Matrix Approximation*, Tech. rep. 2018-01, California Institute of Technology, (2018)

[3] R. M. Bell and Y. Koren, *Lessons from the Netflix prize challenge*, ACM SIGKDD Explorations Newsletter, 9, (2007)

[4] C. Yang, Y. Akimoto, D. W. Kim, and M. Udell, *OBOE: Collaborative filtering for automl initialization*, in NIPS Workshop on Automated Machine Learning, (2018)

[5] M. Udell, A. Townsend, *Why are big data matrices approximately low rank?*, SIAM Journal on Mathematics of Data Science, (2019)

[6] V. A. Daugavet, *The uniform approximation of a function of two variables that is given in tabular form by a product of functions of one variable*, U.S.S.R. Comput. Math. Math. Phys., (1971, in Russian)

[7] V. K. Dzyadyk, *On the approximation of functions on sets consisting of a finite number of points*, The theory of approximation of functions and its applications, (1974, in Russian)

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Stein-based preconditioners for weak-constraint 4D-Var

Jemima Tabeart *Davide Palitta*

Dipartimento di Matematica, Universitá di Bologna

State-of-the-art algorithms for data assimilation are very sophisticated schemes which try to predict the most likely state of a dynamical system by combining information from observations and prior models. The linearized weak-constraint four-dimensional variational assimilation problem (4D-Var) can be reformulated as a saddle point problem that needs to be preconditioned to ensure fast convergence in terms of number of iterations.

In this talk we illustrate novel preconditioning operators which involve the solution of certain Stein matrix equations. In addition to achieving better computational performance, the latter machinery allows us to derive tighter bounds for the eigenvalue distribution of the preconditioned problem.

A panel of diverse numerical results displays the effectiveness of the proposed methodology compared to current state-of-the-art algorithms.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Improved variants of the Hutch++ algorithm for trace estimation

David Persson Alice Cortinovis Daniel Kressner

Institute of Mathematics, EPF Lausanne, 1015 Lausanne

This work is concerned with two improved variants of the Hutch++ algorithm [2] for estimating the trace $\mathrm{tr}(\mathbf{A})$ of a square matrix $\mathbf{A}$, implicitly given through matrix-vector products. Hutch++ combines randomized low-rank approximation in a first phase with stochastic trace estimation in a second phase. In turn, Hutch++ only requires $O\left(\varepsilon^{-1}\right)$ matrix-vector products to approximate $\mathrm{tr}(\mathbf{A})$ within a relative error ε with high probability. This compares favorably with the $O\left(\varepsilon^{-2}\right)$ matrix-vector products needed when using stochastic trace estimation alone. In Hutch++, the number of matrix-vector products is fixed a priori and distributed in a prescribed fashion among the two phases. In this work, we derive an adaptive variant of Hutch++, which outputs an estimate of $\mathrm{tr}(\mathbf{A})$ that is within some prescribed error tolerance with a prescribed success probability, while splitting the matrix-vector products in a near-optimal way among the two phases. For the special case of symmetric positive semi-definite $\mathbf{A}$, we present another variant of Hutch++, called Nyström++, which utilizes the so called Nyström approximation [1] and requires only one pass over the matrix $\mathbf{A}$, as compared to two passes with Hutch++. We prove that the theoretical results on Hutch++ extend to Nyström++. Numerical experiments demonstrate the effectiveness of our two new algorithms.

References

[1] A. Gittens, M. W. Mahoney, *Revisiting the Nyström approximation for improved large-scale machine learning*, J. Mach. Learn., (2016), pp. 3977–4041.

[2] R. A. Meyer, C. Musco, C. Musco, D. P. Woodruff, *Hutch++: Optimal stochastic trace estimation*, In *Symposium on Simplicity in Algorithms (SOSA)*, SIAM, (2021), pp. 142–155.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Matrix Completion with Sparse Measurement Errors

Sergey Petrov

Marchuk Institue of Numerical Mathematics RAS, 119333, Moscow

Matrix completion generally refers to the problem of finding a matrix based on the knowledge of a small fraction of it's elements, under the assumption that the target matrix has a low rank. Certain approaches, like SVP algorithm [1], have been developed for this problem, notably, with geometric convergence bounds under the incoherence hypothesis, and work has been done on the complexity reduction of the said SVP algorithm [2].

In this work, the problem of completion of matrices of small ranks is considered in a special setting, where each element of the matrix may be erroneous with probability $\rho_e = \mathcal{O}\left(\frac{1}{n}\right)$.

Although such a perturbation is extremely sparse on a given mask of known elements, it is not incoherent and algorithms such as SVP method most likely will not work. A new iterative method is proposed that is insensitive to rare observation errors. The method provides the low rank matrix and defines a set containing the erroneous matrix elements. The cardinality of the erroneous set is only a finite number of times greater than the cardinality of a true set of errors.

The method maintains a geometric convergence rate, which is supported by numerical experiments on artificial data. The approach is also applicable to the problem of approximating a given matrix by the sum of a sparse matrix and a matrix of low rank.

This work has been supported by Russian Science Foundation Project (21-71-10072).

References

[1] R. Meka, P. Jain, I.S. Dhillon. Guaranteed rank minimization via singular value projection //arXiv preprint arXiv:0909.5457. – 2009.

[2] O.S. Lebedeva, A.I. Osinsky, S.V. Petrov. Low-Rank Approximation Algorithms for Matrix Completion with Random Sampling. Computational Mathematics and Mathematical Physics (2021), 61(5), pp. 799-815.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Mixed precision recursive block diagonalization for bivariate functions of matrices

Stefano Massei *Leonardo Robol*

Department of Mathematics, University of Pisa, Italy

Matrix functions, denoted by $f(A)$, appear naturally in several applications, such as the analysis of complex networks and control theory. The efficient evaluation of matrix functions has been subject of a thourough study in recent years [1].

This concept extends extends quite naturally to the bivariate setting. Given two square matrices $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$ and a complex-valued function $f(x, y)$, the *bivariate matrix function* $f\{A, B^T\}$ [3] is a linear operator on $\mathbb{C}^{m\times n}$. As in the univariate case, the definition of $f\{A, B^T\}$ can be given, equivalently, in terms of (bivariate) Hermite interpolation, power series expansion and contour integration. In the latter formulation, let Λ_A and Λ_B be the spectra of A and B, respectively, and let $f(x, y)$ be analytic in an open neighborhood of $\Lambda_A \times \Lambda_B$; then, $f\{A, B^T\}$ is defined as

$$\begin{aligned} f\{A, B^T\} : \mathbb{C}^{m\times n} &\longrightarrow \mathbb{C}^{m\times n} \\ C &\longrightarrow f\{A, B^T\}(C) := \oint_{\Gamma_A}\oint_{\Gamma_B} f(x, y)(xI - A)^{-1}C(yI - B)^{-1}\, dxdy, \end{aligned}$$

with Γ_A, Γ_B closed contours enclosing Λ_A and Λ_B, respectively.

Several matrix equations and related numerical problems can be expressed as the evaluation of bivariate matrix functions: solving a Sylvester equations corresponds to evaluating $1/(x + y)$; computing the Frechét derivative of $f(z)$ at A can be rephrased as evaluating the divided difference at the matrices A and A^T, applied to the desired direction. Similarly, any matrix function of $I \otimes A + B^T \otimes I$ can be recast as a bivariate matrix function of the form $f(x, y) = h(x + y)$.

For univariate matrix functions, the Schur-Parlett algorithm provides a robust tool for evaluating $f(A)$ for a generic function. The approach does not extend easily to the bivariate setting.

We propose a numerically reliable method for computing $f\{A, B^T\}(C)$ for a general function $f(x, y)$ without requiring that A and/or B can be diagonalized with a well conditioned similarity transformation. In complete analogy to the univariate scenario, our procedure computes the Schur decompositions $A = Q_A T_A Q_A^*$ and $B = Q_B T_B Q_B^*$, so that the task boils down to evaluate the bivariate function for triangular coefficients:

$$f\{A, B^T\}(C) = Q_A f\{T_A, T_B^T\}(\widetilde{C})Q_B^*, \qquad \widetilde{C} := Q_A^* C Q_B.$$

A generalized block recurrence is applied to retrieve $f\{T_A, T_B^T\}(\widetilde{C})$; the recursion requires to compute f on pairs of diagonal blocks of T_A and T_B^T and to solve Sylvester equations involving either diagonal blocks of T_A or of T_B. In view of the latter operation, we need to reorder the Schur forms of A and B such that distinct diagonal blocks have sufficiently separated eigenvalues. Finally, we evaluate f on the smallest diagonal blocks of T_A and T_B^T, the so-called *atomic blocks*, via a truncated bivariate Taylor expansion or, in the spirit of [2], with a randomized approximate diagonalization technique combined with high precision arithmetic. The procedure can be interpreted as an implicit (recursive) block-diagonalization strategy, where the eigenvectors matrices are not formed explicitly.

References

[1] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, 2008, .

[2] N. J. Higham and X. Liu, *A multiprecision derivative-free Schur–Parlett algorithm for computing matrix functions*, (2020).

[3] D. Kressner, *Bivariate matrix functions*, Operators and matrices, 8 (2014), pp. 449–466, .

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

A distance formula for tuples of doubly commuting matrices

Priyanka Grover *Sushil Singla*

Department of Mathematics, Shiv Nadar University, Delhi NCR, India.

For a tuple of operators $\boldsymbol{A} = (A_1, \ldots, A_d)$, $\operatorname{dist}(\boldsymbol{A}, \mathbb{C}^d\, \boldsymbol{I})$ is defined as $\min_{\boldsymbol{z}\in\mathbb{C}^d} \| \boldsymbol{A} - \boldsymbol{z}\, \boldsymbol{I} \|$ and $\operatorname{var}_x(\boldsymbol{A})$ as $\| \boldsymbol{A}\, x\|^2 - \sum_{j=1}^d \left|\langle x|A_j x\rangle\right|^2$. Ming [3] showed that if $\boldsymbol{A}$ is a tuple of commuting normal operators on a Hilbert space $\mathcal{H}$, then

$$\sup_{\|x\|=1} \operatorname{var}_x(\boldsymbol{A}) = R_{\boldsymbol{A}}^2, \tag{1}$$

where $R_{\boldsymbol{A}}$ is the radius of the smallest disc containing the Taylor spectrum of $\boldsymbol{A}$. We have $R_{\boldsymbol{A}} = \operatorname{dist}(\boldsymbol{A}, \mathbb{C}^d \boldsymbol{I})$.

We will state the idea of the proof of the following. For tuples of doubly commuting matrices, we have

$$\operatorname{dist}(\boldsymbol{A}, \mathbb{C}^d\, \boldsymbol{I})^2 = \sup_{\|x\|=1} \operatorname{var}_x(\boldsymbol{A}).$$

The main facts we will use for the proof will be the normal form for a collection of doubly commuting matrices proved in the main theorem of [1] and the idea of the proof of Proposition 4 of [2].

References

[1] V. Bolotnikov, L. Rodman, *Normal forms and joint numerical ranges of doubly commuting matrices*, Linear Algebra Appl. 301 (1999) 187–194.

[2] A. T. Dash, *Tensor products and joint numerical range*, Proc. Amer. Math. Soc. 40 (1973) 521–526.

[3] F. Ming, *Garske's inequality for an n-tuple of operators*, Integral Equations Operator Theory 14 (1991), 787–793.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Solving PDEs on hypercubes using Chebyshev interpolation

Daniel Kressner *Christoph Strössner*

Institute of Mathematics, EPF Lausanne, Switzerland.

In this talk, we extend our work [2] on the approximation of trivariate functions in functional low-rank formats by combining tensorized Chebyshev interpolation and a low-rank approximation of the coefficient tensor. In the spirit of Chebfun [1], we want to perform numerical computations with these approximations. The application of a differential operator to an approximation in this format, can be evaluated by directly modifying the coefficient tensor. Following the ideas of Chebop2 [3] for two-dimensional domains, we develop a spectral method to solve PDEs on three-dimensional hypercubes by inverting the differential operator in combination with boundary conditions. This leads to a tensor equation, which for certain PDEs has a structure generalizing generalized Sylvester equations from matrices to tensors of order 3.

References

[1] T. A. Driscoll, N. Hale, L. N. Trefethen, *Chebfun Guide*, Pafnuty Publications, Oxford, 2014.

[2] S. Dolgov, D. Kressner, C. Strössner, *Functional Tucker approximation using Chebyshev interpolation*, arxiv e-prints, (2020) p. arXiv:2007.16126.

[3] A. Townsend, S. Olver, *The automatic solution of partial differential equations using a global spectral method*, J. Comput. Phys., 299 (2015) pp. 106-123.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Non-intrusive model order reduction for cross-diffusion systems

Bülent Karasözen *Murat Uzunca* Gülden Mülayim

Department of Mathematics, Sinop University, Sinop-Turkey

Non-intrusive reduced-order models (ROMs) are developed for parametrized cross-diffusion systems [1]

$$\begin{aligned} u_t^\mu &= d_u \nabla^2 u^\mu + d_{vu} \nabla^2 v^\mu + f(u^\mu, v^\mu; \mu), \quad (x,t) \in \Omega \times (0,T], \\ v_t^\mu &= d_{uv} \nabla^2 u^\mu + d_v \nabla^2 v^\mu + g(u^\mu, v^\mu; \mu), \quad (x,t) \in \Omega \times (0,T], \end{aligned} \tag{1}$$

where $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ is the spatial domain, and $T > 0$ is the target time. The parameter dependent variables $u^\mu = u(x,t;\mu)$ and $v^\mu = v(x,t;\mu)$ represent chemical concentrations or population densities, ∇^2 is the Laplace operator, and $f(u^\mu, v^\mu; \mu)$, $g(u^\mu, v^\mu; \mu)$ are the nonlinear reaction terms. Exploiting the Kronecker structure in the finite-difference discretization of (1), the resulting system of ordinary differential equations (ODEs) are integrated in time with the semi-implicit Euler method in matrix or tensor form [2].

The ROMs are constructed using a two-level approach. In the first level, for each parameter value from a sample set of parameters, using higher-order singular value decomposition (HOSVD) [3, 4], the space-time coefficients of the ROMs are computed from the snapshot tensor data, which correspond to the core tensor of the truncated Tucker decomposition of the tensor. In the second level, applying standard singular value decomposition (SVD) to the matrix containing the reduced space-time coefficients for each parameter, the ROM basis is obtained in the training phase. In the test phase, reduced-order solutions are constructed for new parameter values using radial bases functions (RBF). The computational efficiency and accuracy of the ROMs are illustrated to predict the patterns of two examples of cross-diffusion systems of the form (1), the 2D Schnackenberg and 3D Brusselator equations.

References

[1] A. Madzvamuse, H.S. Ndakwo, R. Barreira, *Cross-diffusion-driven instability for reaction-diffusion systems: analysis and simulations*, Journal of Mathematical Biology, (2015), pp. 709-743.

[2] M.C. D'Autilia, I. Sgura, V. Simoncini, *Matrix-oriented discretization methods for reaction-diffusion PDEs: Comparisons and applications*, Computers & Mathematics with Applications, (2020), pp. 2067-2085.

[3] L. De Lathauwer, B. De Moor,J. Vandewalle, *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl., (2000), pp. 253–1278.

[4] N. Vannieuwenhoven, R. Vandebril, K. Meerbergen, *A new truncation strategy for the higher-order singular value decomposition*, SIAM J. Sci. Comput., (2012), pp. A1027–A1052.

Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Operator-dependent prolongation and restriction for parameter-dependent multigrid methods using low-rank tensor formats

Lars Grasedyck *Tim A. Werthmann*

Institut für Geometrie und Praktische Mathematik, RWTH Aachen University

We discuss the solution of parameter-dependent linear systems, i.e., $A(p)u(p) = f$, using parameter-dependent multigrid methods. Such a system arises, e.g., from a discretization of a PDE:

$$\begin{aligned} -\nabla \cdot (\sigma(x,p)\nabla u(x,p)) &= f(x) \qquad \text{for } x \in \Omega, \\ u(x,p) &= 0 \qquad \text{for } x \in \partial\Omega. \end{aligned} \tag{1}$$

In case of discontinuous $\sigma(x,p)$, e.g., if the parameters are jumping or random, $\nabla u(x,p)$ is discontinuous, too, cf. [1]. Therefore using standard linear interpolation for the prolongation and restriction, as in [2], is inaccurate and the convergence rate of a multigrid method declines.

In this talk, we motivate how to deal with these discontinuous $\sigma(x,p)$ in a parameter-dependent multigrid method.

To do so, we will recapitulate the convergence theory of parameter-dependent multigrid methods which we proved in [2]. This theory holds for arbitrary parameter-dependent problems. To achieve a data-sparse representation of the parameter-dependent linear system we recapitulate low-rank tensor formats. Our main question is then: How to deal with discontinuous $\sigma(x,p)$ in (1)?

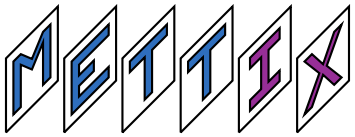
We motivate the derivation of an operator-dependent prolongation and restriction based on block Gaussian elimination, cf. [3]. Numerical experiments using these operator-dependent prolongations and restrictions illustrate a fast convergence of low-rank tensor multigrid methods for discontinuous $\sigma(x,p)$.

References

[1] P. M. De Zeeuw, *Matrix-dependent prolongations and restrictions in a blackbox multigrid solver*, J. Comput. Appl. Math., (1990), pp. 1–27.

[2] L. Grasedyck, M. Klever, C. Löbbert, T. A. Werthmann, *A parameter-dependent smoother for the multigrid method*, arXiv:2008.00927, (2020).

[3] A. Krechel, K. Stüben, *Operator Dependent Interpolation in Algebraic Multigrid*, in Lecture Notes in Computational Science and Engineering, Springer Berlin Heidelberg, 1998.



Matrix Equations and Tensor Techniques IX
Perugia, September 9-10, 2021

Multiway data regression using a spline-type tensor decomposition

Lieven De Lathauwer *Raphaël Widdershoven* Nithin Govindarajan

Nico Vervliet Martijn Boussé

Stadius, KU Leuven, Belgium, Leuven

An important research trend is the transition from vector and matrix based mathematical engineering to generalizations that make use of higher-order tensors. Tensor-algorithms have revolutionised high-dimensional computations. In machine learning or more specifically supervised learning, the potential still largely has to be unleashed [5]. In supervised learning, one searches for an approximation of a nonlinear map. This is done for example with neural networks or support vector machines. Polynomials and splines can also be used, thanks to tensor decompositions that allow for a compact representation of multivariate polynomials and splines [1].

Previously, a similar model was studied specifically with polynomials [4]. These are easy to represent and allow for a uniform approximation of continuous functions, but fall short when nonlinearities are present in the approximated map. In this research [6] splines are used. These give the same advantages, but also allow the approximation of isolated nonlinearities [3]. The extra difficulty is the placement of the spline-knots.

The contribution of this research is a general solution method to approximate an unknown map, based on noisy data [2]. Different implementations have been compared. For classical multivariate splines, one has an exponential time and memory complexity in function of the number of input-variables. With this model it is reduced to a linear complexity. Finally, a few heuristic methods for placing the spline-knots are compared.

References

[1] G. Beylkin, J. Garcke, M. J. Mohlenkamp, *Multivariate Regression and Machine Learning with Sums of Separable Functions*, SIAM, (2009) pp. 1840-1857.

[2] M. Boussé, N. Sidiropoulos, L. De Lathauwer, *NLS Algorithm for Kronecker-Structured Linear Systems with a CPD Constrained Solution*, EUSIPCO, (2019) pp. 1-5.

[3] N. Govindarajan, N. Vervliet, L. De Lathauwer, *Regression and classification with spline-based separable expansions*, Technical Report 21-67, ESAT-STADIUS, KU Leuven (Leuven, Belgium), 2021

[4] S. Hendrikx, M. Boussé, N. Vervliet, L. De Lathauwer, *Algebraic and optimization based algorithms for multivariate regression using symmetric tensor decomposition*,

[5] N. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. Papalexakis, C. Faloutsos, *Tensor Decomposition for Signal Processing and Machine Learning*, IEEE, (2017) pp. 3551-3582.

[6] R. Widdershoven, L. De Lathauwer, N. Govindarajan, M. Boussé, N. Vervliet, *Spline-type tensorontbindingen voor gesuperviseerd leren*, Leuven : KU Leuven. Faculteit Ingenieurswetenschappen, Leuven, 2021.