Matrix Functions and Matrix Equations

Chun Hua Guo, University of Regina, Canada
Valeria Simoncini, University of Bologna, Italy

A Newton-Galerkin-ADI Method for Large-Scale Algebraic Riccati Equations

Peter Benner, TU Chemnitz, Germany
benner@mathematik.tu-chemnitz.de
Thu 16:45, Room Pacinotti

Solving large-scale algebraic Riccati equations (AREs) is one of the central tasks in solving optimal control problems for linear and, using receding-horizon techniques, also nonlinear instationary partial differential equations. Large-scale AREs also occur in several model reduction methods for dynamical systems. Due to sparsity and large dimensions of the resulting coefficient matrices, standard eigensolver-based methods for AREs are not applicable in this context. In the recent two decades, several approaches for such large-scale AREs have been suggested. They mainly fall into two categories:

1. *Galerkin-projection*: the ARE is projected onto a low-dimensional subspace, e.g., a suitable Krylov subspace, then the small scale ARE is solved using a standard solver and the solution is prolongated to full-scale;
2. *Newton's method:* exploit sparsity in the resulting linear system of equations (= a Lyapunov equation) to be solved in each step.

Here, we will present the hybrid method suggested in [1]. It is based on exploiting the advantages of both ideas. Numerical experiments confirm the high efficiency of this new method and demonstrate its applicability to the aforementioned application areas.

[1] P. Benner and J. Saak, A Galerkin-Newton-ADI Method for Solving Large-Scale Algebraic Riccati Equations. Preprint SPP1253-090, DFG Priority Programme 1253 "Optimization with Partial Differential Equations", January 2010.

Joint work with Jens Saak (TU Chemnitz)

Computation of matrix functions arising in the analysis of complex networks

Michele Benzi, Emory University, Atlanta, GA, USA
benzi@mathcs.emory.edu
Thu 15:00, Room Pacinotti

Quantitative methods of network analysis naturally lead to large-scale computations for functions of matrices associated with sparse graphs. This talk will describe some of the main quantities of interest in network analysis as introduced by Estrada, Hatano, D. Higham and others. We combine decay bounds [2,3] and Gaussian quadrature rules [4] to derive a priori bounds and efficient numerical methods for estimating the quantities of interest. Numerical experiments using small-world, range-free, and Erdös-Renyi graphs will be used to illustrate the algorithms. This talk is based in part on the report [1].

[1] M. Benzi and P. Boito, Quadrature Rule-Based Bounds for Functions of Adjacency Matrices, Technical Report TR-2009-031, Department of Mathematics and Computer Science, Emory University, January 2010.
[2] M. Benzi and G. H. Golub, Bounds for the entries of matrix functions with applications to preconditioning, BIT, 29 (1999), pp. 417–438.
[3] M. Benzi and N. Razouk, Decay bounds and $O(n)$ algorithms for approximating functions of sparse matrices, ETNA, 28 (2007), pp. 16–39.
[4] G. Meurant and G. H. Golub, Matrices, Moments and Quadrature with Applications. Princeton University Press, Princeton, NJ, 2010.

Joint work with Paola Boito (Emory University and CERFACS)

On the numerical solution of the matrix equation $\sum_{i=1}^{k} \log(XA_i^{-1}) = 0$

D.A. Bini, University of Pisa, Italy
bini@dm.unipi.it
Mon 15:00, Room Pacinotti

Let $A_i,\ i = 1, \ldots, k$ be real symmetric positive definite $n \times n$ matrices. It is known that the minimum of the function $\sum_{i=1}^{k} d(X, A_i)^2$ for $d(X, Y) = ||X^{-1/2} Y X^{-1/2}||_F$ is attained at a matrix X which solves the equation $\sum_{i=1}^{k} \log(XA_i^{-1}) = 0$. This solution X is called the Karcher mean of the matrices $A_1, \ldots, A_k$.

We introduce the iteration

$$X_{\nu+1} = X_\nu \exp(\theta \sum_{i=1}^{k} \log(XA_i))$$

and its first order approximation

$$X_{\nu+1} = X_\nu + \theta X_\nu \sum_{i=1}^{k} \log(XA_i)$$

for approximating the Karcher mean.

We provide a convergence analysis with a dynamical determination of the optimal parameter θ and show that under certain conditions, convergence is locally quadratic with the optimal choice of θ. We provide a way for the choice of an initial approximation which greatly speeds up the convergence. Numerical experiments which validate our analysis are reported.

Joint work with B. Iannazzo (University of Perugia)

On different classes of Lyapunov equations

Tobias Damm, University of Kaiserslautern, Germany
damm@mathematik.uni-kl.de
Mon 15:25, Room Pacinotti

Lyapunov equations are fundamental e.g. in stability analysis or model order reduction reduction. As is well-known, different forms of Lyapunov operators occur for different classes of systems such as linear stochastic systems, linear delay equations or bilinear systems. We give a short review of these matrix equations and report on some new results and numerical methods.

Inertia and Rank Characterizations of the Expressions $A - BXB^* - CYC^*$ and $A - BXC^* \pm CX^*B^*$

Delin Chu, National University of Singapore, Singapore
matchudl@nus.edu.sg
Mon 15:50, Room Pacinotti

In this paper we consider the admissible inertias and ranks of the expressions $A - BXB^* - CYC^*$ and $A - BXC^* \pm CX^*B^*$ with unknowns X and Y in the four cases when these expressions are : (i) complex self-adjoint,(ii) complex skew-adjoint,

(iii) real symmetric, (iv) real skew symmetric. We also provide a construction for X and Y to achieve the desired inertia/rank, that uses only unitary/orthogonal transformation thus leading to a numerically reliable construction. Consequently, necessary and sufficient solvability conditions for the matrix equations

$$A - BXB^* - CYC^* = 0,$$

and

$$A - BXC^* \pm CX^*B^* = 0$$

are provided.

Joint work with (Y.S.Hung (Department of Electrical and Electronic Engineering. The University of Hong Kong, Hong Kong) and Hugo J. Woerdeman (Department of Mathematics, Drexel University, Philadelphia, USA.))

Hierarchical and Multigrid Methods for Matrix and Tensor Equations

L. Grasedyck, RWTH Aachen, Germany
lgr@mis.mpg.de

Mon 11:00, Room Pacinotti

Hierarchical and Multigrid methods are among the most efficient methods for the solution of large-scale systems that stem, e.g. from the discretization of partial differential equations (PDE). In this talk we will review the generalization of these methods to the solution of matrix equations [1], [2], and equations that possess a tensor structure [3]. The standard hierarchical and multigrid methods can perfectly be combined with low rank (matrix) and low tensor rank representations. The benefit is that the solution is computable in almost optimal complexity with respect to the amount of data needed for the representation of the solution. As an example we consider a PDE posed in a product domain $\Omega \times \Omega, \Omega \subset \mathbb{R}^d$ and discretized with N^d basis functions for the domain Ω. Under separability assumptions on the right-hand side the system is solved in low rank form in $\mathcal{O}(N^d)$ complexity (instead of $\mathcal{O}(N^{2d})$ required for the full solution). For a PDE on the product domain $\underbrace{\Omega \times \cdots \times \Omega}_{D \text{ times}}$ one can even solve the system in low tensor rank form in $\mathcal{O}(N^d)$ complexity (instead of $\mathcal{O}(N^{Dd})$ required for the full solution). The state of the art will be shortly summarized.

[1] L. Grasedyck, W. Hackbusch, A Multigrid Method to Solve Large Scale Sylvester Equations, SIMAX 29, pp. 870-894, 2007.
[2] L. Grasedyck, Nonlinear multigrid for the solution of large scale Riccati equations in low-rank and H-matrix format, Num.lin.alg.appl. 15, pp. 779-807, 2008.
[3] L. Grasedyck, Hierarchical Singular Value Decomposition of Tensors, Technical Report 27/2009, Max Planck Institute for Mathematics in the Sciences, Leipzig, `www.mis.mpg.de`.

Krylov-enhanced parallel integrators for linear problems

S. Güttel, University of Geneva, Switzerland
Stefan.Guettel@unige.ch

Thu 15:50, Room Pacinotti

The parareal algorithm is a numerical method to integrate evolution problems on parallel computers. The main components of this algorithm are a coarse integrator, which quickly propagates information on a coarse partition of the time interval, and a fine integrator, which solves the evolution problems more accurately on each subinterval. The performance of this algorithm is well understood for diffusive problems, but it can also have spectacular performance when applied to certain non-linear problems. In [2] the authors proposed a Krylov-enhanced version of the parareal algorithm, which for linear problems is equivalent to the modified PITA algorithm described in [1]. Both of these algorithms can be successful for 2nd order ODE's. Refining the analysis in [2], we study the convergence of the Krylov-enhanced parareal algorithm and consider the particularly interesting special case when the coarse integrator is a polynomial or rational Krylov-based exponential or trigonometric integrator.

[1] C. Farhat, J. Cortial, C. Dastillung & H. Bavestrello, Time-parallel implicit integrators for the near-real-time prediction of linear structural dynamic responses. Internat. J. Numer. Methods Engrg. 67 (2006), pp. 697–724.
[2] M. Gander & M. Petcu, Analysis of a Krylov subspace enhanced parareal algorithm for linear problems. ESAIM: Proc. 25 (2008), pp. 114–129.

Rational Approximation to Trigonometric Operators

M. Hochbruck, Karlsruhe Institute of Technology, Germany
marlis.hochbruck@kit.edu

Thu 15:25, Room Pacinotti

We will discuss the approximation of trigonometric operator functions that arise in the numerical solution of wave equations by trigonometric integrators. It is well known that Krylov subspace methods for matrix functions without exponential decay show superlinear convergence behavior if the number of steps is larger than the norm of the operator. Thus, Krylov approximations may fail to converge for unbounded operators. In this talk, a rational Krylov subspace method is proposed which converges not only for finite element or finite difference approximations to differential operators but even for abstract, unbounded operators. In contrast to standard Krylov methods, the convergence will be independent of the norm of the operator and thus of its spatial discretization. We will discuss efficient implementations for finite element discretizations and illustrate our analysis with numerical experiments.

[1] V. Grimm und M. Hochbruck Rational approximation to trigonometric operators BIT, vol. 48, no. 2, pp. 215229 (2008)

Joint work with V. Grimm (Karlsruhe Institute of Technology)

A binary powering Schur algorithm for computing primary matrix roots

B. Iannazzo , Università di Perugia, Italy
bruno.iannazzo@dmi.unipg.it

Thu 17:35, Room Pacinotti

Let p be a positive integer. A primary pth root of a square matrix A is a solution of the matrix equation $X^p - A = 0$ which can be written as a polynomial of A.

If A has no nonpositive real eigenvalues then there exists only one primary pth root whose eigenvalues lie in the sector $\mathcal{S}_p = \{z \in \mathbb{C} \setminus \{0\} : |\arg(z)| < \pi/p\}$, which is called principal pth root and denoted by $A^{1/p}$.

The main numerical problem is to compute $(A^{1/p})^r$, for $0 < r < p$ integer. This problem is encountered in certain applications, among which financial models, and in the numerical computation of other matrix functions [2].

We present an algorithm for computing primary roots of a nonsingular matrix A. The algorithm is based on the Schur

decomposition of A. In particular, if A has no nonpositive real eigenvalues, it computes $A^{1/p}$ using only real arithmetics.

The algorithm has an order of complexity lower than the customary Schur based algorithm, namely the Smith algorithm [3], and it is a valid alternative to the algorithms based on rational matrix iterations.

[1] F. Greco and B. Iannazzo, *A binary powering Schur algorithm for computing primary matrix roots*, Numer. Algorithms, 2010.
[2] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, USA, 2008.
[3] M. I. Smith, *A Schur algorithm for computing matrix pth roots*, SIAM J. Matrix Anal. Appl., 2003.

Joint work with F. Greco (Università di Perugia)

Error estimates for two rational Krylov subspace methods to solve the Lyapunov equation with a rank one right-hand side

L. Knizhnerman, Mathematical Modelling Department of Central Geophysical Expedition, Moscow, Russia
mmd@cge.ru
Mon 12:15, Room Pacinotti

The Extended Krylov Subspace Method has recently arisen as a competitive method for solving large-scale Lyapunov equations. Using the theoretical framework of orthogonal rational functions (Faber–Dzhrbashyan series, Blaschke products), in this talk we report on a general a priori error estimate when the known term has rank one, i.e., the equation has the form

$$AX + XA^* + bb^* = 0, \qquad A, X \in \mathbf{R}^{N\times N}, \qquad b \in \mathbf{R}^N,$$

with a positively definite known matrix A.

We also apply the same technique to analyze the behavior of the Rational Krylov Subspace Method, applied to the same problem, with a priori chosen shifts (EKSM corresponds to cyclically repeated shifts 0 and ∞).

Special cases, such as symmetric coefficient matrix, are also treated.

Numerical experiments confirm the proved theoretical assertions.

Joint work with V. Druskin (Schlumberger–Doll Research, Cambridge, USA), V. Simoncini (University of Bologna, Italy), M. Zaslavsky (Schlumberger–Doll Research, Cambridge, USA)

Filters connecting quadratic systems

Peter Lancaster, University of Calgary, Canada.
lancaste@ucalgary.ca
Mon 16:45, Room Pacinotti

The diagonalization of quadratic systems $L(\lambda) = M\lambda^2 + D\lambda + K$ is a fundamental problem in many applications. These systems may have real or complex matrix coefficients, with or without symmetries. Diagonalization by the application of strict equivalence or congruence transformations directly to $L(\lambda)$ is well-understood but is possible for only a very restrictive class of systems. Diagonalization by applying structure preserving transformations to a *linearization* of $L(\lambda)$ has also been developed recently, and is possible for a wider class of systems.

Here, we describe the possibility of finding *linear* systems of the form $F(\lambda) := F_1\lambda + F_0$ for which

$$\tilde{F}(\lambda)L(\lambda) = \tilde{L}(\lambda)F(\lambda)$$

and $\tilde{L}(\lambda)$ is *diagonal*. We call thesefunctions *linear filters*.

We show how filters can be constructed using familiar structures of "standard pairs" and "structure preserving transformations".

Joint work with S.D.Garvey, (University of Nottingham, UK), A.Popov, (University of Nottingham, UK), U.Prells, (University of Nottingham, UK), I.Zaballa, (Euskal Herriko Unibersitatea, Spain).

Stabilizing complex symmetric solution of the equation $X + A^\top X^{-1} A = Q$ arising in nano research

Wen-Wei Lin, National Chiao Tung University, Taiwan
wwlin@math.nctu.edu.tw
Mon 17:10, Room Pacinotti

We study the existence and characteristic of the stabilizing complex symmetric solution X_s for the matrix equation $X + A^\top X^{-1} A = Q$ arising in nano research. In stead of using the deep theory of linear operators we give a new proof on the existence of X_s by using only basic knowledge of linear algebra. Furthermore, we show that the imaginary part of X_s is positive semi-definite with rank$=m/2$, where m is the number of simple unimodular eigenvalues of the rational matrix-valued function $\psi(\lambda) \equiv Q + \lambda A + \lambda^{-1} A^\top$. We also present a doubling algorithm for computing the desired solution X_s efficiently and reliably.

Joint work with Chun-Hua Guo (University of Regina, Canada) and Yueh-Cheng Kuo (National University of Kaohsiung, Taiwan).

Algorithms for nonnegative quadratic vector equations

F. Poloni, Scuola Normale Superiore, Pisa, Italy
f.poloni@sns.it
Mon 17:35, Room Pacinotti

We investigate a vector equation having the form

$$Mx = a + b(x, x), \tag{1}$$

where $a, x \in \mathbb{R}^n_{\geq 0}$, M is an $n\times n$ M-matrix and $b : \mathbb{R}^n_{\geq 0} \times \mathbb{R}^n_{\geq 0} \to \mathbb{R}^n_{\geq 0}$ is a bilinear map. The equation (1) appears in the study of Markovian binary trees [Bean, Kontoleon Taylor, *Ann. Oper. Res.* '08; Hautphenne, Latouche, Remiche, *LAA* '08].

We propose a new functional iteration (and a corresponding Newton method) for its solution, based on the computation of the Perron vector of a special matrix. The most interesting property of these methods is that their convergence behaviour does not degrade when the equation is close to null recurrent, in contrast to the traditional algorithms. This means that they are particularly effective on the most "difficult" problems.

Moreover, we may weaken the hypotheses of the original probabilistic equation in order to obtain a general framework for systems of quadratic equations with nonnegativity constraints, encompassing nonsymmetric algebraic Riccati equations [Guo, Laub, *SIMAX* '00], Lu's simple equation [Lu L.-Z., *SIMAX* '05], and several quadratic equations in queuing theory and probability [Bini, Latouche, Meini, *LAA* '02 and '03]. This allows us to give a unified treatment of the numerical methods for their solution. In some cases, this unification leads to new algorithms or more general proofs.

It is still an open problem whether it is possible to extend the new Perron vector-based iterations to this larger family of quadratic equations.

Joint work with D. A. Bini (University of Pisa), B. Meini (University of Pisa)

Lur'e Equations and Even Matrix Pencils

T. Reis, TU Berlin / TU Hamburg-Harburg (Germany)
reis@math.tu-berlin.de

Mon 11:25, Room Pacinotti

Lur'e equations are a generalization of algebraic Riccati equations and they arise in linear-quadratic optimal control problem which are singular in the input. It is well-known that there is a one-to-one correspondence between the solutions of Riccati equations and Lagrangian eigenspaces of a certain Hamiltonian matrix. The aim of this talk is to generalize this concept to Lur'e equations. We are led to the consideration of deflating subspaces of even matrix pencils.

Dimension reduction for damping optimization of linear vibrating systems

Ninoslav Truhar, University of Osijek, Croatia
ntruhar@mathos.hr

Mon 11:50, Room Pacinotti

Consider a damped linear vibrational system described by the differential equation

$$M\ddot{x} + D\dot{x} + Kx = 0\,, \qquad x(0) = x_0, \quad \dot{x}(0) = \dot{x}_0\,,$$

where M, D, K are mass, damping and stiffness matrix, respectively.

A very important question arises in considerations of such systems: *for given mass and stiffness determine the damping matrix so as to insure an optimal evanescence.*

It can be shown that this optimization problem is equivalent to the following minimization problem:

$$\mathrm{trace}(X) = \min\,,$$

where X is solution of the following Lyapunov equation:

$$AX + XA^T = -GG^T\,,$$

here A is $2n \times 2n$ matrix obtained from M, D and K, and G is matrix with full column rank, and $\mathrm{rank(G)} \ll n$.

Finding the optimal D such that the trace of X is minimal is a very demanding problem, caused by the large number of trace calculations, which are required for bigger matrix dimensions. We propose a dimension reduction to accelerate the optimization process and we present corresponding error bound for the approximation of the solution of Lyapunov equation obtained by this reduction. We will show a new estimates for the eigenvalue decay of the solution X which include the influence of the right-hand side G on the eigenvalue decay rate of the solution. Also, we will present an efficient algorithm for the minimization of $\mathrm{trace}(X)$ using a low rank Cholesky ADI method based on a new set of ADI parameters.

Joint work with Peter Benner, Chemnitz University of Technology, Germany and Zoran Tomljanović, University of Osijek, Croatia

---

Algorithms for matrix functions

K. Ziętak, Wrocław University of Technology, Poland
krystyna.zietak@pwr.wroc.pl

Thu 17:10, Room Pacinotti

The matrix sector function, introduced by Shieh, Tsay and Wang, is a generalization of the matrix sign function. For a positive integer p and a matrix $A \in \mathbb{C}^{n\times n}$, having no eigenvalues with argument $(2k+1)\pi/p$ for $k = 0, 1, \ldots, p-1$, the matrix sector function is defined by $\mathrm{sect}_p(A) = A(\sqrt[p]{A^p})^{-1}$, where $\sqrt[p]{X}$ denotes the principal pth root of X. For $p = 2$ the matrix sector function is the matrix sign function.

We derive and investigate a family of iterations for the sector function, based on the Padé approximants of a certain hypergeometric function. This generalizes a result of Kenney and Laub [3] for the sign function and yields a whole family of iterative methods for computing the matrix pth root.

We prove that the principal Padé iterations for the matrix sector function are structure preserving. It generalizes the result of Higham, Mackey, Mackey, Tisseur [1] for the principal Padé iterations for the matrix sign function (see also Iannazzo [2]).

We also focus on the coupled Padé iterations for computing the matrix pth root. The talk is based on [4] and some current investigations.

[1] N.J. Higham, D.S. Mackey, N. Mackey, F. Tisseur, Computing the polar decomposition and the matrix sign decomposition in matrix groups, *SIAM J. Matrix Anal. Appl.* 25 (2004), 1178–1192.
[2] B. Iannazzo, A family of rational iterations and its application to the computation of the matrix pth root, *SIAM J. Matrix Anal. Appl.*, 30 (2008), 1445–1462.
[3] Ch.S. Kenney, A.J. Laub, Rational iterative methods for the matrix sign function, *SIAM J. Matrix Anal. Appl.* 12 (1991), 273–291.
[4] B. Laszkiewicz, K. Ziętak, A Padé family of iterations for the matrix sector function and the matrix pth root, *Numer. Lin. Alg. Appl.* 16 (2009), 951–970.
