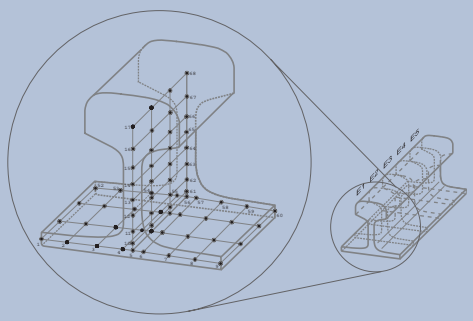
GAMM-SIAM

Conference on Applied Linear Algebra

organized in cooperation with ILAS

HEINRICH HEINE UNIVERSITÄT DÜSSELDORF

BERGISCHE UNIVERSITÄT WUPPERTAL



Düsseldorf, July 24-27, 2006

Joint GAMM–SIAM Conference on Applied Linear Algebra

organized in cooperation with ILAS

July 24–27, 2006, Düsseldorf, Germany

Book of Abstracts

A variant of the Orthomin(k) method for solving linear systems

Kuniyoshi Abe
Faculty of Economics and Information, Gifu Shotoku University
1-38, Nakauzura, Gifu, 500-8288 Japan
abe@gifu.shotoku.ac.jp

Shao-Liang Zhang
Department of Computational Science and Engineering, Nagoya University
Furo-cho, Nagoya, 464-8601 Japan
zhang@na.cse.nagoya-u.ac.jp

By Krylov subspace methods, we are solving a large sparse linear system $A\boldsymbol{x} = \boldsymbol{b}$, where A and $\boldsymbol{b}$ stand for an n-by-n (not necessarily symmetric) matrix and an n-vector, respectively. The Bi-Conjugate Gradient (Bi-CG) method and a number of *product-type Krylov subspace methods* are a well-known generic Krylov subspace method for this problem. Moreover the Generalized Minimal Residual (GMRES), the Generalized Conjugate Residual (GCR) and ORTHOMIN(k) methods have been known as a Krylov subspace method based on the minimum residual approach.
When applying Bi-CG and product-type Krylov subspace methods to singular linear systems, the residual norms temporarily decrease but diverge without attaining the minimum residual $\min_{\boldsymbol{b}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$. On the other hand, GMRES, GCR and ORTHOMIN(k) are preferably applied to the problem, since their residual norms decrease monotonously and are expected to converge to the minimum. Furthermore, for singular linear systems, these methods are known theoretically to attain the minimum residual under a certain condition. However, in the actual computation with finite precision arithmetic, the residual of ORTHOMIN(k) is often observed to be reduced further than the theoretically expected level. Therefore, we propose a variant of ORTHOMIN(k), which is mathematically equivalent to the original ORTHOMIN(k) method, but uses recurrence formulas that are different from those of ORTHOMIN(k); they contain alternative expressions for the auxiliary vector and the recurrence coefficients. Our implementation has the same computational costs as ORTHOMIN(k).
In numerical experiments on nonsingular systems our implementation is not necessarily less affected by rounding errors than ORTHOMIN(k), but numerical results on singular systems show that our implementation is more accurate and less affected by rounding errors than ORTHOMIN(k).

Implicit Riemannian trust-region method for symmetric generalized eigenproblems

Pierre-Antoine Absil
Université catholique de Louvain and University of Cambridge
Peterhouse, Cambridge CB2 1RD, UK
http://www.inma.ucl.ac.be/∼absil/

Christopher G. Baker, Kyle A. Gallivan
School of Computational Science, Florida State University
Tallahassee, FL 32306-4120, USA
http://www.scs.fsu.edu/{∼cbaker,∼gallivan}

We have recently proposed a manifold-based trust-region method to compute extreme eigenpairs of symmetric matrix pencils [3]. The general optimization method had been shown [1, 2] to enjoy strong global and local convergence properties, which are inherited by its "brute force" application to the extreme eigenproblem. However, the trust-region mechanism, which restricts the norm of the update in an adaptive way, has some inherent inefficiencies: when the trust-region is too large, valuable time may be spent computing an update that ends up being rejected; when the trust-region is too small, computational effort may be wasted computing updates that are exceedingly conservative. It may take the adaptive process several iterates to bring the trust-region radius to an appropriate size. In this work, we show how the particular structure of the eigenproblem can be exploited to remedy this drawback. We present an analysis that provides us knowledge of the model fidelity at every step of the inner iteration, making the standard, explicit trust-region mechanism obsolete, while preserving the convergence properties. We also show how the stopping criterion of the outer iteration can be monitored in the inner iteration, in order to prevent the final call to the inner iteration from performing more work than is needed.

References

[1] P.-A. Absil, C. G. Baker, and K. A. Gallivan, *Trust-region methods on Riemannian manifolds with applications in numerical linear algebra*, in Proceedings of the 16th International Symposium on Mathematical Theory of Networks and Systems (MTNS2004), Leuven, Belgium, 5–9 July 2004, 2004.

[2] ———, *Trust-region methods on Riemannian manifolds*. http://www.inma.ucl.ac.be/˜absil/, submitted, 2005.

[3] ———, *A truncated-CG style method for symmetric generalized eigenvalue problems*. J. Comput. Appl. Math., to appear, 2006.

Algebraic multigrid for multilevel Toeplitz and matrix algebra linear systems: proof of convergence and optimality

Antonio Aricò
Dipartimento di Matematica "Leonida Tonelli", Università di Pisa
Largo Bruno Pontecorvo 5, 56127 Pisa, Italy
arico@mail.dm.unipi.it

Marco Donatelli, Stefano Serra-Capizzano
Dipartimento di Fisica e Matematica, Università dell'Insubria - Sede di Como
Via Valleggio 11, 22100 Como, Italy
{marco.donatelli,stefano.serrac}@uninsubria.it

Algebraic multigrid (AMG) has been introduced and investigated from late '80s on (see [8]). Algorithms to solve Toeplitz, tau and circulant systems have been introduced in early '90s, for hermitian semipositive matrices. In such algorithms smoothers are relaxed Richardson, the approach is Galerkin and the restrictors are chosen according to the matrix coefficients. All the matrix structures can be related to a couple of parameters, that are the dimension and the generating function (also called symbol).
The goal was to deal with ill-conditioned problems of course, with number of levels $d \in \mathbb{N} \setminus \{0\}$ ($d > 1$ leads to multilevel structures), where the symbol is 2π periodic (and vanishes in at least a point, with even order $2k$), and to show AMG optimality i.e. that the convergence rate does not depend on the matrix size but only on the symbol (for similar results in a PDE setting see [3]).
Optimality results for $d = 1, 2$ problems has been shown by Serra-Capizzano and Fiorentino for two-grid methods (TGM) [6, 7] applied to tau and Toeplitz matrices in the ill-conditioned simmetric case; Chan et oth. [5] showed AMG convergence (but not optimality) in $d = 1$ Toeplitz case but with laplacian-like conditioning.
We remark TGM optimality results in [4], that are a consequence of a convergence theorem due to Ruge and Stüben [8]. That theory, when combined with Perron-Frobenius theorem to prove uniform symbol convergence, can be applied to get optimality for AMG with any number of subgrids [2, 1].
Here an AMG algorithm is presented (substantially equal to the one originately reported in [6], but with stronger restrictor/interpolator) as well as the proof of optimality in ill-conditioned algebra case. Then it is shown how to extend the proof to the multilevel matrix algebra case.
Numerical examples are presented and discussed to solve $(d = 2, 3)$-dimensional problems, which leads to multilevel matrices with a 2π periodic symbol that vanishes in a point x_0 and is positive in $[-\pi, \pi)^d \setminus \{x_0\}$.

References

[1] A. Aricò and M. Donatelli, *A V-cycle multigrid for multilevel matrix algebras: proof of optimality*. Under revision for Numer. Math.

[2] A. Aricò, M. Donatelli, and S. S. Capizzano, *V-cycle optimal convergence for certain (multilevel) structured linear systems*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 186–214.

[3] A. Brandt, *Guide to multigrid development*, in Multigrid Methods, W. Hackbusch and U. Trottenberg, eds., vol. 960 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1982, pp. 220–312.

[4] S. S. Capizzano, *Convergence analysis of two-grid methods for elliptic Toeplitz and PDEs matrix-sequences*, Numer. Math., 92 (2002), pp. 433–465.

[5] R. H. Chan, Q.-S. Chang, and H.-W. Sun, *Multigrid method for ill-conditioned symmetric Toeplitz systems*, SIAM J. Sci. Comput., 19 (1998), pp. 516–529.

[6] G. Fiorentino and S. Serra, *Multigrid methods for Toeplitz matrices*, Calcolo, 28 (1991), pp. 283–305.

[7] ———, *Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions*, SIAM J. Sci. Comput., 17 (1996), pp. 1068–1081.

[8] J. W. Ruge and K. Stüben, *Algebraic multigrid*, in Multigrid Methods, S. McCormick, ed., Frontiers in Appl. Math., SIAM, Philadelphia, PA, 1987, ch. 4, pp. 73–130.

Stopping criteria in the context of finite element discretization

Mario Arioli
Rutherford Appleton Laboratory, UK
Chilton, Didcot, Oxfordshire, OX11 0QX, UK
m.arioli@rl.ac.uk

We combine linear algebra techniques with finite element techniques to obtain a reliable stopping criterion for Krylov method based algorithms. The Conjugate Gradient method has for a long time been successfully used in the solution of the symmetric and positive definite systems obtained from the finite-element approximation of self-adjoint elliptic partial differential equations. Taking into account recent results [3, 4, 5] which make it possible to approximate the energy norm of the error during the conjugate gradient iterative process, in [1] we introduce a stopping criterion based on an energy norm and a dual space norm linked to the continuous problem. Moreover, we show that the use of efficient preconditioners does not require us to change the energy norm used by the stopping criterion.
In [2], we extend the previous results on stopping criteria to the case of nonsymmetric positive-definite problems. We show that the residual measured in the norm induced by the symmetric part of the inverse of the system matrix is relevant to measuring convergence in a finite element context. We then provide alternative ways of calculating or estimating this quantity.
Finally, we extend the results to matrices having a block partitioned structure. Here, the aim is to extract useful information regarding the accuracy of the computed function from which we would approximate the true solution of the original partial differential equation when mixed finite element methods are used.

References

[1] M. Arioli. *A stopping criterion for the conjugate gradient algorithm in a finite element method framework*, Numer. Math. 97 (2004), pp 1-24.

[2] M. Arioli, D. Loghin, and A. Wathen. *Stopping criteria for iterations in finite-element methods*, Numer. Math. 99 (2005), pp 381-410.

[3] G. Meurant, *Numerical experiments in computing bounds for the norm of the error in the preconditioned conjugate gradient algorithm*, Numerical Algorithms, 22 (1999), pp. 353–365.

[4] Z. Strakoš and P. Tichý, *On error estimation by conjugate gradient method and why it works in finite precision computations*, Electronic Transactions on Numerical Analysis, 13 (2002), pp. 56–80.

[5] Z. Strakoš and P. Tichý, *'Error estimation in preconditioned conjugate gradients*, to appear in BIT.

$\mathcal{H}$-matrix techniques for high-frequency Helmholtz BEM

Lehel Banjai
Institute of Mathematics, University of Zurich
Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
lehelb@math.unizh.ch

In this talk, we discuss the application of $\mathcal{H}$-matrix techniques to the solution of Helmholtz problems with large wave number κ. We consider the Brakhage-Werner integral formulation of the problem, discretised by the Galerkin boundary element method (BEM). The dense $n \times n$ Galerkin matrix arising from this approach is represented by a sum of an $\mathcal{H}$-matrix and an $\mathcal{H}^2$-matrix.
We use a well-known multipole expansion to construct the $\mathcal{H}^2$-matrix [4]. The numerical instability problems of using this multipole expansion have been well documented. We propose a new algorithm to deal with this issue. The part of the matrix that causes difficulty is approximated by an $\mathcal{H}$-matrix, which can be constructed in a stable manner by the adaptive cross approximation algorithm (ACA)[1]. Furthermore, we use existing algebraic recompression methods [3] to reduce storage and complexity of arithmetical operations of the $\mathcal{H}$-matrix. The LU-decomposition of such a recompressed $\mathcal{H}$-matrix can then be computed efficiently [2, 3]. Once the LU-decomposition is available, the $\mathcal{H}$-matrix can also be used as an effective preconditioner. We concentrate on the BEM with piecewise constant basis functions and prove that for a given accuracy $\epsilon > 0$, the proposed algorithm has complexity $O(\kappa \log \kappa \log n + n \log \kappa \log \frac{1}{\epsilon})$ for the construction of the $\mathcal{H}$ and $\mathcal{H}^2$ matrices and for the matrix-vector product. The linear systems are solved by a preconditioned iterative method. We present numerical results for scattering problems in two dimensions.

References

[1] M. Bebendorf, *Approximation of boundary element matrices*, Numer. Math., 86 (2000), pp. 565–589.

[2] ———, *Hierarchical LU decomposition based preconditioners for BEM*, Technical report 28/2004. Leipzig, Germany: Max Planck Institute for Mathematics in the Sciences, (2004).

[3] L. Grasedyck, *Adaptive recompression of $\mathcal{H}$-matrices for BEM*, To appear in Computing, (2004).

[4] V. Rokhlin, *Rapid solution of integral equations of scattering theory in two dimensions*, J. Comput. Phys., 86 (1990), pp. 414–439.

A note on the error analysis of classical Gram–Schmidt

Jesse Barlow
Department of Computer Science and Engineering, The Pennsylvania State University
111 IST Building, University Park, PA 16802–6822, USA
barlow@cse.psu.edu

Alicja Smoktunowicz
Faculty of Mathematics and Information Science, Warsaw University of Technology
Pl. Politechniki 1, Warsaw, 00-661 Poland
smok@mini.pw.edu.pl

Julien Langou
The University of Tennessee, Department of Computer Science
1122 Volunteer Blvd., Knoxville, TN 37996-3450, USA
langou@cs.utk.edu

An error analysis result is given for classical Gram–Schmidt factorization of a full rank matrix A into $A = QR$ where Q is left orthogonal (has orthonormal columns) and R is upper triangular. The work presented here shows that the computed R satisfies $R^T R = A^T A + E$ where E is an appropriately small backward error, but only if the diagonals of R are computed in a manner similar to Cholesky factorization of the normal equations matrix.
A similar result is stated in [Giraud at al, Numer. Math. 101(1):87–100,2005]. However, for that result to hold, the diagonals of R must be computed in the manner recommended in this work.

Overlooked numerical aspects in the history of continued fractions

Friedrich L. Bauer
Institut für Informatik, Technische Universität München
Nördliche Villenstr. 19, D-82288 Kottgeisering, Germany

The mathematical treatment of continued fractions that started after 1600 first dealt with practical arithmetic topics. This explains that only in 1770 Johannes Heinrich Lambert found a transformation of the Gregory power series for arctan into an infinite continued fraction by applying the Euclidian algorithm to the power series in a polynomial ring. From the numerical point of view the Lambert transformation is of great importance, since it leads to convergence acceleration.
While Euler's transformation (1739) turns the sequence of partial sums of the power series into a sequence of partial quotients of the continued fraction (transforming the badly convergent Gregory series into a continued fraction that converges equally badly), the continued fraction developed by the Lambert transformation shows linear convergence.
Euler's transformation was well recognized (Perron: 'Besonders häufig wird Euler's transformation auf Potenzreihen angewandt'), and the advantages of the Lambert transformation were ignored in the literature. The Lambert transformtion is not even explicitly mentioned in Oskar Perron's excellent book of 1913, but only in a form that dates back to Gauß (1813) and is based there on a transformation of a quotient of a pair of hypergeometric functions.
Furthermore, which is hardly ever mentioned in the literature, the Lambert arctan continued fraction is 'limit periodic', and its 'limit continued fraction' allows the analysis of the convergence behaviour in quantitative form. Also other interesting continued fractions that result from hypergeometric power series via the Lambert transformation are 'limit periodic'; for these as well linear convergence is the consequence.
Numerical experiments (by coutesy of Christoph Haenel) will demonstrate this.

Model reduction for large-scale systems based on hierarchical matrix arithmetic

Ulrike Baur
Institut für Mathematik, Technische Universität Berlin
Strasse des 17.Juni 136 D-10623 Berlin, Germany
baur@math.tu-berlin.de

We consider linear time-invariant stable systems of the following form

$$\Sigma : \begin{cases} \dot{x}(t) &= Ax(t) + Bu(t), & x(0) = x_0 \\ y(t) &= Cx(t) & t \geq 0 \end{cases}$$

with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $C \in \mathbb{R}^{p\times n}$ arising, e.g., from the discretization of parabolic PDEs. We will assume that the system Σ is large-scale, having order $n = \mathcal{O}(10^5)$ with $n \gg m, p$. Balanced truncation and related methods are the most commonly used model reduction techniques used in control theory. These methods require the solution of large-scale matrix equations and are therefore of limited use for very large-scale problems. We consider approaches based on data-sparse matrix approximations, hierarchical matrix formats, and the corresponding formatted arithmetic in order to obtain efficient model reduction methods having linear-polylogarithmic complexity.

Inexact interpolation for model reduction

Christopher Beattie
Department of Mathematics,
Virginia Polytechnic Institute and State University,
Blacksburg, Virginia 24061-0123, USA
beattie@vt.edu

Rational Krylov subspaces (which equivalently force transfer function interpolation) are capable of providing H_∞-optimal approximation in model reduction for large scale linear dynamical systems, although large linear systems of equations must be solved to generate the required subspace bases. In practice, iterative solvers become necessary and early termination becomes prudent but this then raises the possibility of undesirable degradation of the approximating subspace. Choice of termination criteria and effect of preconditioning evidently can then influence the quality of the final reduced order models. General bounds on the H_∞ error associated with a reduced order model are discussed that provide a new tool to understand both features.

GMRES error estimates in terms of the numerical range

Bernhard Beckermann
Laboratoire Painlevé, UMR 8524 USTL-CNRS, Université de Lille 1
UFR de Mathématiques, Bât. M2, F-59655 Villeneuve d'Ascq, France
bbecker@math.univ-lille1.fr

A popular method for solving non-hermitian systems $Ax = b$ is given by GMRES [5]. Provided that A has a positive definite hermitian part $(A + A^*)/2$, Elman [3] pointed out the following upper bound for the norm of the kth residual r_k of GMRES for every $k \geq 0$

$$\frac{||r_k||}{||r_0||} \leq \sin^k(\beta), \qquad \text{where } \cos(\beta) := \frac{\lambda_{\min}((A + A^*)/2)}{||A||} \tag{1}$$

with $\beta \in [0, \pi/2)$, where by $\|\cdot\|$ we denote the Euclidean vector norm. Estimate (1) is obtained by iterating the inequality for $k = 1$ corresponding to a one-dimensional minimization problem, which should allow for some improvements.
Following the discussion in [4], we propose similar bounds of the form

$$\frac{||r_k||}{||r_0||} \leq (2 + \gamma)\gamma^k \tag{2}$$

based on the field of values (or numerical range)

$$W(A) := \{(Ay, y) : y \in \mathbb{C}^n, ||y|| = 1\}$$

which by the Hausdorff Theorem is convex. It is shown in [2] (see also [1]) that one may chose $\gamma = 2\sin(\frac{\beta}{4-2\beta/\pi}) < \sin(\beta)$ in (2), with β as in (1). More generally [1], inequality (2) remains valid in case $0 \notin W(A)$ by taking the so-called asymptotic convergence factor $\gamma = 1/|\phi(0)|$, where ϕ is the Riemann conformal map mapping the exterior of $W(A)$ onto the exterior of the unit disk, and $\phi(\infty) = \infty$.

References

[1] B. Beckermann, *Image numérique, gmres et polynômes de faber*, C. R. Acad. Sci. Paris, Ser. I, 340 (2005), pp. 855–860.

[2] B. Beckermann, S. Goreinov, and E. Tyrtyshnikov, *Some remarks on the elman estimate for gmres*, SIAM J. Matrix Anal. Appl., 27 (2006), pp. 772–778.

[3] S. Eisenstat, H. Elman, and M. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.

[4] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, Frontiers in Applied Mathematics 17, SIAM, 1997.

[5] Y. Saad and M. Schultz, *Gmres: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Comput., 7 (1986), pp. 856–869.

An efficient reconstruction method for discrete band-limited signals

Hamid R. Behmard
Western Oregon University
345 N. Monmouth Avenue, Monmouth Oregon 97361
behmarh@wou.edu

A prevailing question in signal analysis is the following: given the samples $x(t_n), n \in \mathbb{Z}$, of a signal $x(t) \in L_2(\mathbb{R})$, how to recover x. If the signal x does not contain high frequencies, i.e., if x is band-limited, and if the sampling set is dense enough, then x can be recovered uniquely. The classical Shannon-Whittaker-Kotel'nikov sampling theorem permits reconstruction of a bandlimited signal from its values on a set of equidistant points on the real line $\mathbb{R}$[6].
It has been extended to higher dimensions and different groups setting; see the reviews [7, 6]. One of these groups setting is the finite cyclic group $\mathbb{Z}_N$. In digital signal processing, a digital signal x of length N is a finite sequence $x(0), x(1), \cdots, x(N-1) \in \mathbb{C}^N$, where the set of indices is identified with $\mathbb{Z}_N$. Periodic sampling, introduced by Kohlenberg [8] is well established and considers sampling sets which are unions of cosets of one subgroup. In [1], we considered sampling theory for sampling sets which are unions of shifted lattices. These sampling sets were not necessarily periodic. We note that periodic and nonperiodic samplings fall under irregular or nonuniform classification of sampling theory [3, 5]. Since the discrete problem of irregular sampling is a finite-dimensional and linear one, a reconstruction of signal is equivalent to solving a large system of linear equations or to inverting a large matrix. In the case of nonperiodic sampling, a signal x could be reconstructed via a recursive reconstruction algorithm from its samples provided the sampling sets and the support of the Fourier transform of x satisfy certain compatibility conditions; see the manuscripts [1, 9]. These compatibility conditions were relaxed in the recent manuscript [2]. One application of these relaxed compatibility conditions is sampling sets that are combination of periodic and nonperiodic sampling subgroups. In this talk, we describe the mathematics of this hybrid sampling set. We also demonstrate its implementation to speech processing.

References

[1] H. Behmard and A. Faridani, *Sampling of band limited functions on unions of shifted lattices*, J. Fourier Anal. Apple. 8 (2002), ppm. 43-58.

[2] H. Behmard, A. Faridani and D. Walnut, *Construction of Sampling Theorems for Unions of Shifted Lattices*, Sampling Theory in Signal and Image Processing, to appear.

[3] J.J. Benedetto, *Irregular sampling and frames.* In C.K. Chui (ed.), Wavelets - A Tutorial in Theory and Applications, Academic Press, 1992, pp. 445-507.

[4] A. Faridani, *A generalized sampling theorem for locally compact abelian groups*, Math. Comp. 63(1994), pp. 307-327.

[5] H.G. Feichtinger and K. Gröchenig, *Theory and practice of irregular sampling*, Wavelets: Mathematics and Applications, (J.J. Benedetto and M.W. Frazier eds.), CRC Press, 1993, pp. 305-363.

[6] J.R. Higgins, *Five short stories about the cardinal series*, Bull. Amer. Math. Soc., 12(1985), pp. 45-89.

[7] A. J. Jerri, *The Shannon sampling theorem - its various extensions and applications: a tutorial review*, Proc. IEEE, 65(1977), pp. 1565-1596.

[8] A. Kohlenberg, *Exact interpolation of bandlimited functions*, J. Appl. Phys., 24(1953), pp. 1432-1436.

[9] D. Walnut, *Nonperiodic Sampling of Bandlimited Functions on Unions of Rectangular Lattices*, J. Fourier Anal. Appl., 2(1996), pp. 435-452,

Early research on iterative methods: the Italian contribution

Michele Benzi
Department of Mathematics and Computer Science, Emory University
400 Dowman Drive, Atlanta, GA 30322 USA
benzi@mathcs.emory.edu

In this talk I will review some early work on iterative methods for solving linear systems by Italian mathematicians during the 1930s, with particular attention to the contributions of Lamberto Cesari (1910–1990) and Gianfranco Cimmino (1908–1989). I will also provide some background information on Italian applied mathematics and especially on Mauro Picone's *Istituto Nazionale per le Applicazioni del Calcolo*, where most of this early numerical work took place. Finally, I will illustrate the influence of Cimmino's work on modern and current research.
This talk is based on my paper "Gianfranco Cimmino's Contributions to Numerical Mathematics," *Atti del Seminario di Analisi Matematica*, Dipartimento di Matematica dell'Università di Bologna. Volume Speciale: Ciclo di Conferenze in Memoria di Gianfranco Cimmino, Marzo-Aprile 2004, Tecnoprint, Bologna (2005), pp. 87–109.

An augmented Lagrangian approach for the Oseen problem

Michele Benzi
Department of Mathematics and Computer Science, Emory University
400 Dowman Drive, Atlanta, GA 30322 USA
benzi@mathcs.emory.edu

Maxim Olshanskii
Faculty of Mechanics and Mathematics, Moscow State University
Moscow 119899, Russia
Maxim.Olshanskii@mtu-net.ru

We describe an effective solver for the discrete Oseen problem based on an augmented Lagrangian formulation of the corresponding saddle point system. The proposed method is a block triangular preconditioner used with a Krylov subspace iteration (BiCGStab). The crucial ingredient is a novel multigrid approach for the (1,1) block, which extends a technique introduced by Schöberl for elasticity problems to nonsymmetric problems. Our analysis indicates that this approach results in fast convergence, independent of the mesh size and largely insensitive to the viscosity. We present experimental evidence for both isoP2-P0 and isoP2-P1 finite elements in support of our conclusions. We also show results of a comparison with a state-of-the-art coupled multigrid solver, showing the competitiveness of our approach.

Inexact inverse iteration applied to a Jordan block

Jörg Berns-Müller
Fachbereich Informatik und Mathematik, Johann Wolfgang Goethe Universität Frankfurt
Postfach 11 19 32, 60054 Frankfurt am Main, Germany
berns@math.uni-frankfurt.de

Inexact inverse iteration is inverse iteration where the arising shifted linear systems are not solved by direct methods, like using an LU-decomposition, but are approximated by using an iterative technique. In contrast to inverse iteration, which is a well known and well studied technique, the convergence of inexact inverse iteration is not well established. Recent works on the convergence of inexact inverse iteration restrict to diagonalisable eigenvalue problems. While defects in non targeted eigenvalues have little effect the situation is rather complicated if the targeted eigenvalue is defective.
In this talk we consider the convergence of inexact inverse iteration when the targeted eigenvalue corresponds to a Jordan block. We provide a clear cut and simple result for the case of a fixed shift implementation of inexact inverse iteration. Further we look at variable shift methods like the Rayleigh quotient iteration.

On classification of spatial matrices

Maxim Bershadsky
Ben-Gurion University of the Negev, Department of Mathematics
Department of Mathematics,Ben Gurion University of the Negev, P.O.B. 653,Be'er Sheva 84105,ISRAEL
maximb@mail.sapir.ac.il

Genrich Belitskii
Ben-Gurion University of the Negev, Department of Mathematics
Department of Mathematics,Ben Gurion University of the Negev, P.O.B. 653,Be'er Sheva 84105,ISRAEL
genrich@cs.bgu.ac.il

Vladimir Sergeichuck,
Institute of Mathematics, Terechenkivska 3, Kiev, Ukraine
sergeich@imath.kiev.ua

Some general propositions on classification of spatial matrices are stated. The probem is proved to be superwild although the similar plane problem has a simple and well-known solution - the canonical form of Smith. The spatial matrices have different canonical forms and their classification depends on the field. Such a classification has been presented for some particular cases. Applications are presented.

The GSVD and domain decomposition methods for planar eigenvalue problems

Timo Betcke
Institut Computational Mathematics, TU Braunschweig
Pockelsstr. 14, 38106 Braunschweig, Germany
t.betcke@tu-bs.de

In 1983 Descloux and Tolley proposed a domain decomposition method for the eigenvalue problem

$$-\Delta u = \lambda u \quad \text{in } \Omega \tag{1a}$$

$$u = 0 \quad \text{on } \partial\Omega, \tag{1b}$$

where $\Omega \subset \mathbb{R}^2$ is a polygonal domain [2]. In each subdomain they used Fourier-Bessel functions that automatically satisfy the eigenvalue equation (1a) to approximate eigenfunctions. By imposing continuity conditions on the interfaces between the subdomains they derived the parameter dependent generalized eigenvalue problem

$$F(\lambda)x(\lambda) = \mu(\lambda)G(\lambda)x(\lambda), \tag{2}$$

where $F(\lambda)$ is symmetric positive semi-definite and $G(\lambda)$ is symmetric positive definite. The minima of the smallest eigenvalue $\mu_1(\lambda)$ are then approximations to the eigenvalues of (1).
Although their method converges exponentially for an increasing number of Fourier-Bessel functions in each subdomain the accuracy to which eigenvalues of (1) can be obtained by minimizing $\mu_1(\lambda)$ is limited to the squareroot of machine precision. This was pointed out by Driscoll in 1997. He solved this problem by explicitly computing the derivative $\mu_1'(\lambda)$ and solving $\mu_1'(\lambda) = 0$ [3].
In this talk we present another solution to the limited accuracy problem of the method of Descloux and Tolley. Instead of generalized eigenvalues we use the generalized singular value decomposition (GSVD) to formulate the method and minimize the smallest generalized singular value $\sigma_1(\lambda)$ of a certain pencil $\{A(\lambda), B(\lambda)\}$. This formulation also solves the limited accuracy problem and can deliver eigenvalue approximations close to machine precision. Furthermore, while the methods of Descloux, Tolley and Driscoll need evaluations of boundary and domain integrals this is fully avoided in our formulation. Our method can also be interpreted as an extension of the subspace angle method proposed in [1] to domain decomposition methods.
For complicated domains the matrices $A(\lambda)$ and $B(\lambda)$ can become large and sparse. Since we are only interested in the smallest generalized singular values of $\{A(\lambda), B(\lambda)\}$ this motivates the use of projection methods for the GSVD.
We present several numerical examples including accurate eigenvalue computations on multiply connected domains.

References

[1] T. Betcke and L. N. Trefethen, *Reviving the method of particular solutions*, SIAM Review, 47 (2005), pp. 469–491.

[2] J. Descloux and M. Tolley, *An accurate algorithm for computing the eigenvalues of a polygonal membrane*, Comput. Methods Appl. Mech. Engrg., 39 (1983), pp. 37–53.

[3] T. A. Driscoll, *Eigenmodes of isospectral drums*, SIAM Review, 39 (1997), pp. 1–17.

Algebraic multigrid for symmetric indefinite systems

Matthias Bollhoefer
Institut für Mathematik, MA 4-5, TU Berlin
Str. des 17. Juni 136, D-10623 Berlin, Germany
bolle@math.tu-berlin.de

Marcus Grote
Departement Mathematik, Universität Basel
Rheinsprung 21, CH-4051 Basel, Switzerland
marcus.grote@unibas.ch

Olaf Schenk
Departement Informatik, Universität Basel
Klingelbergstr. 50, CH-4056 Basel, Switzerland
olaf.schenk@unibas.ch

We will discuss an algebraic multigrid method that is devoted for the solution of large sparse symmetric indefinite systems. In particular we address the situation when the system is highly indefinite. The cases arise e.g. from a discretization of a Helmholtz equation for high wave numbers or the Anderson model of localization.
The method is mainly based on three major ingredients:

1. symmetric maximum weight matchings to increase the block diagonal dominance of the system,
2. inverse–based pivoting to drive the coarsening process and finally
3. filtering techniques to handle frequencies near zero eigenvalues.

We will illustrate the resulting multilevel method of this approach for selected numerical examples.

References

[1] M. Bollhöfer and Y. Saad, *Multilevel preconditioners constructed from inverse–based ILUs*, SIAM Journal on Scientific Computing, (to appear).

[2] M. Hagemann and O. Schenk, *Weighted matchings for the preconditioning of symmetric indefinite linear systems*, Technical Report CS-2004-005, Department of Computer Science, University of Basel, 2004.

[3] O. Schenk, M. Bollhöfer, and R. A. Römer, *On large scale diagonalization techniques for the anderson model of localization*, SIAM Journal on Scientific Computing, (to appear).

Hierarchically refined grid for the efficient calculation of boundary conditions for the simulation in infinite domains

Matthias Bolten
Zentralinstitut für Angewandte Mathematik, Forschungszentrum Jülich
D-52425 Jülich, Germany
m.bolten@fz-juelich.de

If mesh-based approaches are used for the simulation of infinite space, most often Dirichlet boundary conditions are imposed at the border of a finite simulation box and the solution is computed on the grid inside of the simulation box, e.g. in [2].
Examples for the need of the simulation of such infinite domains can be found in applications from engineering [3] or from molecular dynamics simulation [4]. Considering e.g. a biomolecule in vacuum, the Dirichlet conditions have to be known on the border of the simulation box in order to compute the electrostatic potential in the interior domain. It is clear that all inner points, or (in case of discrete particle simulations) charged particles contribute to the boundary potential which results in a complexity of $\mathcal{O}(N^{5/3})$ to determine the Dirichlet conditions in case that a uniform grid is used. This is unacceptable, especially with the existance of efficient solvers like multigrid methods. In particle simulation codes one can use a multipole expansion [4], but this will not work for all applications.
Another approach is presented here: A hierarchically refined grid is used to reduce the number of boundary points to a fixed number, so the resulting complexity is $\mathcal{O}(N)$, only. With increasing grid size, the number of the necessary additional unknowns grows much slower than the number of inner grid points by construction. The resulting linear system can be solved with standard solvers.
This technique is used to calculate the electrostatic potential of a system of N charges. In order to avoid a discontinuous source function due to point charges, a splitting into short- and long-range contributions to the electrostatic potential is used, where the long range part has a smooth source function. This part is calculated by solving the Poisson equation, which therefore has a well behaved smooth solution. A geometric multigrid method, which is based on Brandt's Multi-level Adaptive Technique [1], is used as solver.
The complexity of the whole resulting method is $\mathcal{O}(N)$ and thus an optimal solver is found for the Poisson equation in infinite space.
The technique can be easily adapted to solve other PDEs in infinite domains as well, as long as the boundary values can be precomputed somehow.

References

[1] A. Brandt, *Multi-level adaptive technique (mlat) for fast numerical solution to boundary value problems*, in Proceedings of the Third International Conference on Numerical Methods in Fluid Mechanics, H. Cabammes and R. Temam, eds., vol. 1 of Lecture Notes in Physics, Berlin, Heidelberg, New York, July 1972, Springer-Verlag, pp. 82–89.

[2] O. Buneman, *Simulation of infinite space on a finite computer*, in Proceedings of the Fourth Conference on the Numerical Simulation of Plasmas, J. P. Boris and R. A. Shanny, eds., Washington, D.C., November 1970, Naval Research Laboratory, pp. 642–649.

[3] R. H. Burkhart, *Asymptotic expansion of the free-space green's function for the discrete 3-d poisson equation*, SIAM J. Sci. Comput., 18 (1997), pp. 1142–1162.

[4] G. Sutmann and B. Steffen, *A particle-particle particle-multigrid algorithm for long range interactions in molecular systems*, Comp. Phys. Comm., 169 (2005), pp. 343–346.

Solving elliptic finite element systems in nearly linear time with support preconditioners

Erik Boman
Sandia National Labs
Albuquerque, NM 87185-1111
egboman@sandia.gov

Bruce Hendrickson
Sandia National Labs
Albuquerque, NM 87185-1111
bah@cs.sandia.gov

Stephen Vavasis
Department of Computer Science, Cornell University
4130 Upson Hall, Ithaca, NY 14853, USA
vavasis@cs.cornell.edu

We show in this talk how support preconditioners can be applied to a class of linear systems arising from use of the finite element method to solve linear elliptic problems. Finite element discretizations of elliptic PDEs give rise to large sparse linear systems of equations. A topic of great interest is preconditioners for iterative solution of such systems. Our contribution in this paper has two parts: First, we show how SPD matrices of a particular form $K = A^T \bar{D}^{1/2} \bar{J}^T \bar{J} \bar{D}^{1/2} A$ can be well approximated by a symmetric diagonally dominant, M-matrix. The quality of approximation depends only on the condition of J. Since significant theory has been developed for symmetric, diagonally dominant M-matrices, we know they can be solved efficiently using preconditioned iterative methods. Second, we show that the stiffness matrix K from a certain class of finite elements for elliptic problems has the structure mentioned above. In this context, J has a condition number bounded in terms of mesh quality, intra-element jumps in conductivity, and other mesh-related properties but is independent of problem size, shape of the domain (except for cusps), variation in element size, and interelement variation in conductivity field. These two contributions put together imply the existence of a nearly linear-time iterative method for solving elliptic finite element problems.

On the convergence of exact and inexact multiplicative Schwarz methods for singular M-matrices and Markov chains

Stefan J. Borovac
Fachbereich C – Mathematik und Naturwissenschaften, Bergische Universität Wuppertal
Gauß-Str. 20, D-42097 Wuppertal, Germany
borovac@math.uni-wuppertal.de

Consider the system of linear equations

$$x = Bx + b, \quad B \in \mathbb{R}^{n\times n}, B \geq 0, \rho(B) = 1, \tag{1}$$

where $b \in \mathbb{R}^n$ is in the range of $I - B$ and the null space of $I - B$ has dimension 1.
It is known that multiplicative Schwarz methods applied to (1) need not converge against the desired solution. We present a graph based method that allows one to find a partition of the underlying system and an order of the subspace corrections such that convergence will be achieved with a reliable freedom of overlap. The method is based on a reordering of the variables (state space reordering) and will be applicable to an irreducible B and also to slightly more general systems.
Since the graph of the above B poses a minimal structure which guarantees the dimension of $I - B$ to be 1, our method will preserve that structure and automatically leads to convergent Schwarz iterations for both exact and (two-stage) inexact methods.
If relaxed inexact Schwarz methods are used it will be shown that a state space reordering is not necessary and convergence will be achieved with some minor additional requirements.
Furthermore, some approaches for generalisations will be discussed.

On Krylov subspace matrix function evaluations in time integration schemes for space discretized Maxwell's equations

Mike A. Botchev
University of Twente, Department of Applied Mathematics
P.O. Box 217, NL-7500 AE Enschede, the Neherlands
mbotchev@na-net.ornl.gov

Space discretizations of time dependent Maxwell's equations typically result in a large system of ordinary differential equations (ODE's) of a special matrix structure. Recently, several special time integration schemes involving matrix functions [2, 3] have been succesfully applied for this class of ODE systems [1]. In this talk, we discuss aspects of efficient computation of the matrix functions for these problems using Krylov subspace projections.
This talk is based on author's joint work with Davit Harutyunyan, Nadezda Orlovskaya and Jaap van der Vegt.

References

[1] M. A. Botchev, D. Harutyunyan, and J. J. W. van der Vegt, *The Gautschi time stepping scheme for edge finite element discretizations of the maxwell equations*, J. Comput. Phys., (2006). To appear.

[2] V. L. Druskin and L. A. Knizhnerman, *Krylov subspace approximations of eigenpairs and matrix functions in exact and computer arithmetic*, Numer. Lin. Alg. Appl., 2 (1995), pp. 205–217.

[3] M. Hochbruck, C. Lubich, and H. Selhofer, *Exponential integrators for large systems of differential equations*, SIAM J. Sci. Comput., 19 (1998), pp. 1552–1574.

A generalized conjugate direction method for nonsymmetric linear systems

Edouard R. Boudinov
Fortis Bank
Rokin 55, 1012KK, Amsterdam, The Netherlands
edouard.boudinov@nl.fortisbank.com

Arkadiy I. Manevich
Dept. Computational Mechanics and Strength of Structures, Dniepropetrovsk National University
Naukova str. 13, 49050, Dniepropetrovsk, Ukraine
aman@atlantis.dp.ua

A new method for solving nonsymmetric large and/or ill-conditioned linear algebraic systems of equations $\mathbf{A}x = b, \quad x, b \in \Re^N$ is proposed, which unifies the advantages of the conjugate direction (CG) and GMRES methods. The method is based on the generalized ("one-sided") notion of conjugacy proposed in a few works, in particular, by G.W. Stewart: vectors $\{d_i\}$ are named conjugate if $(d_i, \mathbf{A}d_k) = 0$ for $i < k$ (in general case $(d_i, \mathbf{A}d_k) \neq 0$ for $i \geq k$)
In each iteration the step is done along the last found conjugate vector d_k, but this vector is constructed using all preceding conjugate vectors $\{d_i\},\ i < k$:

$$x_{k+1} = x_k + \alpha_k d_k, \quad \alpha_k = \frac{(r_k, d_k)}{(d_k, \mathbf{A}d_k)}, \quad d_k = r_k^0 + \sum_{i=1}^{k-1} \beta_i^{(k)} d_i \tag{1}$$

(r_k^0 is the normalized residual vector). It is shown that conditions of "one-sided" conjugacy result in a simple two-term recurrent formula for $\beta_i^{(k)}$ (instead of the direct solution of the corresponding triangular set of equations with respect to $\beta_i^{(k)}$ $(i = 1, \ldots, k-1)$, which would lead to "long" recurrent relationships expressing each $\beta_i^{(k)}$ through all preceding values of $\beta_j^{(k)},\ j = 1, \ldots, i-1$)).
A principal peculiarity of the method is that the residuals are obtained in the Arnoldi orthogonalization process:

$$r_{k+1} = -\alpha_k \left(\mathbf{A}r_k^0 - \sum_{i=1}^{k} \gamma_{k,i} r_i^0 \right), \quad \gamma_{k,i} = (\mathbf{A}r_k^0, r_i^0), \quad r_k^0 = \frac{r_k}{\|r_k\|} \tag{2}$$

In finite precision arithmetic this formula is found to be much more accurate than the usual updating formula $r_{k+1} = r_k - \alpha_k \mathbf{A}d_k$. In addition, one can carry out an additional orthogonalization cycle (*reorthogonalization*), in the case of very ill-conditioned problem.
The vectors r_k^0 form an orthonormal basis set in Krylov subspaces, which does not degenerate and is used for the most calculations. Note that we avoid solving an auxiliary linear problem with the Hessenberg matrix in each iteration for computing the step, unlike GMRES. By the cost of storing the entire basis $\{r_k^0\}$ such an approach ensures a rather stable computational scheme with high accuracy, in distinction from methods using only one or two last basis vectors.
The algorithm has been realized in JAVA programming language and has been tested on a variety of large-scale and/or ill-conditioned linear algebraic problems. The performance of the method has been compared with that of the known efficient algorithms (CG, Bi-CG, CGS, Bi-CGSTAB, GMRES). The proposed method is found to be very stable and efficient for large scale and/or ill-conditioned problems.

The generalized eigenvalue problem for nonsquare pencils

Gregory Boutry
Catholic University of Lille, Faculté Libre des Sciences et Technologies
41 rue du port 59046 LILLE
gregory.boutry@icl-lille.fr

Michael Elad
Computer Science Department, The Technion
Israel Inst. of Techology, Haifa, 32000 Israel
elad@cs.technion.ac.il

Gene H. Golub
Department of Computer Science
Stanford University, Stanford CA. 94305-9025 USA
golub@stanford.edu

Peyman Milanfar
Department of Electrical Engineering,
University of California, Santa Cruz CA. 95064 USA
milanfar@ee.ucsc.edu

This work focuses on nonsquare matrix pencils $A - \lambda B$ where $A, B \in \mathcal{M}^{m \times n}$ and $m > n$. Traditional methods for solving such non-square generalized eigenvalue problems $(A - \lambda B)\underline{v} = \underline{0}$ are expected to lead to no solutions in most cases. In this paper we propose a different treatment: We search for the minimal perturbation to the pair (A, B) such that these solutions are indeed possible. This talk proposes insight into the characteristics of the described problems along with practical numerical algorithms towards their solution. We also present a simplifying factorization for such non-square pencils, and some relations to the notion of pseudospectra.

References

[1] E. Elmroth, P. Johansson, and B. Kågström, *Bounds for the distance between nearby jordan and kronecker structures in a closure hierarchy.*, Numericals Methods and Algorithms XIV, Zapiski Nauchnykh Seminarov, 268 (2000), pp. 24–48.

[2] T. G. Wright and L. N. Trefethen, *Pseudospectra of rectangular matrices*, IMA Journal Numer. Anal., 22 (2002), pp. 501–519.

Model order reduction with Krylov-based rational interpolation

Angelika Bunse-Gerstner
Zentrum für Technomathematik, Fachbereich 3, Universität Bremen,
Postfach 330 440, 28334 Bremen, Germany
Bunse-Gerstner@math.uni-Bremen.de

Modelling complex physical phenomena often leads to dynamical systems which are too large to be handled in further computational processes. Model reduction methods can then serve as a tool to reduce the complexity by an automatic reduction of the order of the system while approximately preserving its behavior. Already rather well developed are techniques for linear time invariant dynamical systems of the form

$$\begin{aligned} \mathbf{E}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{aligned}$$

with transfer function $\mathbf{H}(s) = \mathbf{C}\left(s\mathbf{E} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D}$, where $\mathbf{A}, \mathbf{E} \in \mathbb{R}^{n\times n}, \mathbf{B} \in \mathbb{R}^{n\times m}, \mathbf{C} \in \mathbb{R}^{p\times n}, \mathbf{D} \in \mathbb{R}^{p\times m}$ are constant coefficient matrices, with $n \gg m, p$ and $\mathbf{u}, \mathbf{y}$ are vector-valued input and output functions, respectively.
For very large sytems Krylov-based projection methods are often favorable. They compute the reduced system of dimension $k \ll n$ as

$$\begin{aligned} \mathbf{W}^T\mathbf{E}\mathbf{V}\dot{\mathbf{x}}(t) &= \mathbf{W}^T\mathbf{A}\mathbf{V}\mathbf{z}(t) + \mathbf{W}^T\mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{V}\mathbf{z}(t) + \mathbf{D}\mathbf{u}(t), \end{aligned}$$

where $\mathbf{V}$ and $\mathbf{W}$ are matrices of dimension $n \times k$. Their column vectors are bases of rational Krylov subspaces which are chosen such that the transfer function of the reduced system interpolates the transfer function $\mathbf{H}(s)$ and perhaps some of its derivatives in given points $s_1, \ldots s_\ell$ in the complex plane. This can be viewed as a (generalized) moment matching or multi-point Padé technique. If $\mathbf{H}(s)$ is a matrix valued function, the interpolation condition might also be relaxed to interpolating $\mathbf{H}(s)$ only in certain (tangential) directions.
The interpolation points and tangential direction should be chosen, such that the transfer function of the reduced systems is a reasonable overall approximation to the original transfer function, such that in fact the input-output behavior of the reduced and original system is similar. But which choice of the interpolation points and directions will give us such approximations? Moreover, if the original systems have additional structure - they may e.g. be stable, passive or stem from systems with derivatives of higher order - we may need or want to preserve this structure in the reduced system. How can this be achieved? In this talk we show results, experiments and connections to Sylvester equations and certain eigenvalue problems with which we approach answers to these questions.

Comparing Leja and Krylov approximations of large scale matrix exponentials

Marco Caliari
Department of Pure and Applied Mathematics, University of Padua
Via G. Belzoni 7, 35131 Padova, Italy
mcaliari@math.unipd.it

We have implemented a numerical code (ReLPM, Real Leja Points Method) for polynomial interpolation of the matrix exponential propagators $\exp(\Delta t A)\,\mathbf{v}$ and $\varphi(\Delta t A)\,\mathbf{v}$, $\varphi(z) = (\exp(z) - 1)/z$. The ReLPM code is tested and compared with Krylov-based routines, on large scale sparse matrices arising from the spatial discretization, by Finite Differences or Finite Elements, of 2D and 3D advection-diffusion equations.
We have also implemented an efficient parallel version of the ReLPM showing good scaling behavior even when using more than thousand processors.

Joint work with L. Bergamaschi, A. Martínez and M. Vianello.

New ideas for computing complex eigenvalues of an asymmetric matrix applied to metastable states

Tucker Carrington
Departement de chimie, Universite de Montreal
Case postale 6128, succursale Centre-ville, Montreal (Quebec) H3C 3J7, Canada
Tucker.Carrington@umontreal.ca

Jean Christophe Tremblay
Departement de chimie, Universite de Montreal
Case postale 6128, succursale Centre-ville, Montreal (Quebec) H3C 3J7, Canada
jc.tremblay@umontreal.ca

Lifetimes of metastable states may be determined by computing complex eigenvalues of a complex-symmetric matrix. One way to do this is to use the complex-symmetric version of Cullum and Willoughby's non-orthogonalised Lanczos procedure. It is also possible use shift and invert complex-symmetric Lanczos to compute eigenvalues window by window. Both of these approaches require doing complex matrix-vector products. In this talk I shall outline a new method for obtaining complex eigenvalues from real matrix-vector products. The complex-symmetric eigenvalue problem is written as a quadratic real eigenvalue problem which is linearized. We attempted to solve the linearized (asymmetric) problem using the symmetric indefinite Lanczos method but discovered that the symmetric indefinite Lanczos method works poorly for the purpose of computing many eigenpairs. Instead, we use the standard two-sided Lanczos algorithm (exploiting the symmetry of the eigenvalue problem to reduce the number of matrix-vector products by a factor of two) to compute approximate eigenvectors. It is not possible to get very accurate eigenvectors. We use groups of approximate eigenvectors as a basis and compute extremely accurate complex eigenvalues.

Continuation BSOR-LG method in multi-component Bose–Einstein condensates

Shu-Ming Chang
Department of Mathematics, National Tsing Hua University
101 Kuang Fu Road Sec. 2, 30013 Hsinchu, Taiwan
schang@mx.nthu.edu.tw

We develop a continuation block successive over-relaxation Lanczos-Galerkin (BSOR-LG) method for the computation of positive bound states of multi-component Bose-Einstein Condensates (BEC). It is well-known [2, 4, 6] that coupled Gross-Pitaevskii equations (CGPEs), also called coupled nonlinear Schrödinger equations, can be used as a mathematical model to describe a multi-component BEC in different hyperfine spin states on the corresponding condensate wave functions. In ultracold dilute Bose gas, different hyperfine spin states may repel each other and form separate nodal domains, such a phenomenon, called phase separation of a multiple mixture of BEC, has been extensively investigated by experimental and theoretical physicists [3, 4, 6].
First we discretize the time-independent CGPEs to a nonlinear algebraic eigenvalue problem (NAEP) and develop a structured continuation method based on the classical continuation method [1, 5] for the computation of possibly all positive bound states of a multi-component BEC. Second, in order to utilize the sparsity and the block structure of the associated NAEP, we propose a continuation method combined with the BSOR iteration [9, p. 594-596] and the Lanczos-Galerkin projection method [7, 8] for tracing the solution curve of the NAEP. Third, we prove that the primal stalk of the solution curve of the NAEP coincides with the unique global minimizer of a single-component BEC which is represented by an initial value ODE. Furthermore, we prove that the solution curve of the NAEP will encounter a first bifurcation point at a finite value of the repulsive scattering length. For a multi-component BEC, we prove that m identical ground/bound states will bifurcate into m different ground/bound states at a finite repulsive inter-component scattering length. Numerical results show that various positive bound states of a two/three-component BEC are solved efficiently and reliably by the continuation BSOR-LG method.

References

[1] E. L. Allgother and K. Gerog, *Numerical path following*, North-Holland, Amsterdam, 1997.

[2] W. Z. Bao, *Ground states and dynamics of multi-component Bose-Einstein condensates*, SIAM Multiscale Modeling & Simulation, 2 (2004), pp. 210–236.

[3] S. M. Chang, C. S. Lin, T. C. Lin, and W. W. Lin, *Segregated nodal domains of two-dimensional multi-species Bose-Einstein condensates*, Physica D, 196 (2004), pp. 341–361.

[4] D. S. Hall, M. R. Matthews, J. R. Ensher, C. E. Wieman, and E. A. Cornell, *Dynamics of component separation in a binary mixture of Bose-Einstein condensates*, Phys. Rev. Lett., 81 (1998), pp. 1539–1542.

[5] H. B. Keller, *Lectures on Numerical Methods in Bifurcation Problems*, Springer-Verlag, Berlin, 1987.

[6] C. J. Myatt, E. A. Burt, R. W. Ghrist, E. A. Cornell, and C. E. Wieman, *Production of two overlapping Bose-Einstein condensates by sympathetic cooling*, Phys. Rev. Lett., 78 (1997), pp. 586–589.

[7] B. N. Parlett, *A new look at the Lanczos algorithm for solving symmetric systems of linear equations*, Lin. Alg. Appl., 29 (1980), pp. 323–346.

[8] Y. Saad, *On the Lanczos method for solving symmetric linear systems with several right-hand-sides*, Math. Comput., 48 (1987), pp. 651–662.

[9] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, Berlin, 2 ed., 1991.

Numerical methods for integer least squares problems

Xiao-Wen Chang
School of Computer Science, McGill University
3480 University Street, Montreal, Quebec, Canada H3A 2A7
chang@cs.mcgill.ca

Integer least squares (ILS) problems may arise from many applications such as communications, cryptograph, global positioning systems, etc, see, e.g., [1], [6], and [7]. Unlike real LS problems, ILS problems are NP-hard. Solving an ILS problem usually consists of two phases: reduction and search. The goal of the reduction is to make the search process more efficient. In this talk, we first briefly introduce a typical algorithm for ordinary ILS problems and then present our recent work on various ILS problems, which may subject to different constraints, see [2], [3], [4] and [5]. Some simulation results will be given to show our algorithms can be significantly faster than the current ones in the literature.

References

[1] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, *Closest point search in lattices*, IEEE Trans. Inform. Theory, 48 (2002), pp. 2201–2214.

[2] X.-W. Chang, G. Golub, and Q. Han, *Solving elliposoid-constrained integer least squares problems*. In preparation.

[3] X.-W. Chang and Q. Han, *Solving box-constrained integer least squares problems*. Submitted, 2005.

[4] X.-W. Chang and X. Yang, *Solving underdetermined integer least squares problems with a class of constraints*. Submitted, 2006.

[5] X.-W. Chang, X. Yang, and T. Zhou, *MLAMBDA: A modified LAMBDA method for integer least-squares estimation*, J. of Geodesy, 79 (2005), pp. 525–565.

[6] M. O. Damen, H. El Gamal, and G. Caire, *On maximum-likelihood detection and the search for the closest lattice point*, IEEE Trans. Inform. Theory, 49 (2003), pp. 2389–2402.

[7] P. Teunissen, *The least-squares ambiguity decorrelation adjustment: a method for fast GPS ambiguity estitmation*, J. of Geodesy, 70 (1995), pp. 65–82.

When is the Laplacian spectrum of a weighted digraph real?

Pavel Chebotarev
Institute of Control Sciences of the Russian Academy of Sciences
65 Profsoyuznaya Str., 117997 Moscow, Russia
pavel4e@gmail.com

Rafig Agaev
Institute of Control Sciences of the Russian Academy of Sciences
65 Profsoyuznaya Str., 117997 Moscow, Russia
arpo@ipu.ru

The Laplacian matrix of a weighted digraph Γ is the matrix $L = L(\Gamma) = (\ell_{ij})$, where ℓ_{ij} equals minus the weight of arc (i, j) in Γ when $j \neq i$, and the diagonal entries are such that the row sums are zero. Some spectral properties of L have been studied in [1]. The question in the title is important in the analysis of vehicle formations [2, 3], because the stability of formations can be checked using a Nyquist-like criterion derived from the spectral properties of the digraph Laplacian.

It is easy to see that the Laplacian spectrum is real for two simple classes of weighted digraphs: (i) digraphs without cycles; (ii) symmetric digraphs for which the weights of the (i, j) and (j, i) arcs are equal for every i and $j \neq i$. It can be said that, for digraphs of type (ii), all cycles are countervailed by contradirectional cycles. So a natural informal hypothesis is that the spectrum is not completely real if and only if the weighted digraph is "essentially cyclic". Thus, the problem is to give a relevant definition to the "essential cyclicity". Consider a few examples.

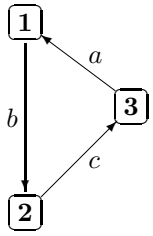


Fig. 1.

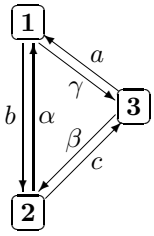


Fig. 2.

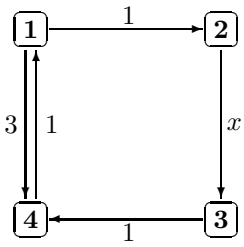


Fig. 3.

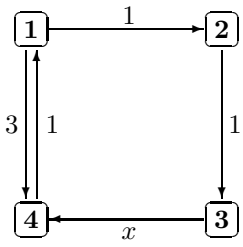


Fig. 4.

Suppose that for the digraph on Fig. 1, $a \geq b$ and $a \geq c$. It turns out that the Laplacian spectrum is not completely real iff $\sqrt{a} < \sqrt{b} + \sqrt{c}$. Thus, the single cycle with arc weights a, b, and c is "essential" iff $\sqrt{a}$, $\sqrt{b}$, and $\sqrt{c}$ satisfy the strict triangle inequality.

For the digraph on Fig. 2, let us call the cycle (a, b, c) "Latin" and the cycle (α, γ, β) "Greek". We say that vertex 1 is Latin if $a > \alpha$ and Greek if $\alpha > a$. Similarly, vertex 2 (vertex 3) is Latin if $b > \beta$ (resp., $c > \gamma$) and Greek if $\beta > b$ (resp., $\gamma > c$). In these terms, the Laplacian spectrum of the digraph is not completely real iff [vertices 1, 2, and 3 are Latin and $\sqrt{a-\alpha}$, $\sqrt{b-\beta}$, and $\sqrt{c-\gamma}$ satisfy the strict triangle inequality] or [vertices 1, 2, and 3 are Greek and $\sqrt{\alpha-a}$, $\sqrt{\beta-b}$, and $\sqrt{\gamma-c}$ satisfy the strict triangle inequality]. Unfortunately, there is no such simple condition of "essential cyclicity" in the general case. For example, for the digraphs shown in Fig. 3 and Fig. 4, the Laplacian spectrum is not completely real iff x is strictly between x_1 and x_2, where $x_{1,2} = \left(37 - Q \pm \sqrt{290 - 36z - 504/z + 3454/Q}\right)/12$, $z = 0.5\sqrt[3]{671 + 65\sqrt{65}}$, and $Q = \sqrt{36z + 145 + 504/z}$. Approximately, $0.266 < x < 2.441$. In the paper, we consider some classes of "essentially cyclic" weighted digraphs.

References

[1] R. Agaev and P. Chebotarev, *On the spectra of nonsymmetric Laplacian matrices*, Linear Algebra and Its Applications, 399 (2005), pp. 157–168.

[2] J. A. Fax and R. M. Murray, *Graph Laplacians and stabilization of vehicle formations*, in Proceedings of the 15th IFAC Congress, Barcelona, Spain, July 2002.

[3] J. J. P. Veerman, G. Lafferriere, J. Caughman, and A. Williams, *Flocks and formations*. To appear in J. Stat. Phys.

The chaotic numbers of complete multipartite graphs

Nam-Po Chiang
Department of Applied Mathematics, Tatung University
40 Chungshan North Rd. 3rd Sec., Taipei, Taiwan
npchian@ttu.edu.tw

Let $G = (V, E)$ be a connected graph and let ϕ be a permutation of V. The total relative displacement of the permutation ϕ of G is

$$\delta_\phi(G) = \sum_{\{x,y\} \subset V} |d(x, y) - d(\phi(x), \phi(y))|,$$

where $d(x, y)$ means the distance between x and y in G. The maximum value of $\delta_\phi(G)$ among all permutations in a graph G is called the chaotic number of G and the permutation which attains to the chaotic number is called a chaotic mapping of G.
In this talk, we consider the chaotic numbers of complete multipartite graphs. We find an algorithm running to find the chaotic numbers along a linear algebra approaching.

Higher-order discrete evolution systems: analytic properties involving several matrices and numerical eigenvalue methods

Julio Claeyssen
Instituto de Matematica-Universidade Fedetal do Rio Grande do Sul
P.O.Box 10673, 90.001-900 Porto Alegre-RS,Brasil
julio@mat.ufrgs.br

Teresa Tsukazan
Instituto de Matematica-Universidade Fedetal do Rio Grande do Sul
P.O.Box 10673, 90.001-900 Porto Alegre-RS,Brasil
teresa@mat.ufrgs.br

In this work we shall formulate directly and exploit a matrix basis that is generated by a fundamental discrete initial-value Green matrix function of a linear evolution system of arbitrary order with square matrix coefficients. The formulation in terms of a fundamental matrix function is done without using the standard approach of reducing a higher-order system into a first-order equation involving companion matrices.

We shall derive extensions of the Cayley-Hamilton theorem for several matrices and of iterative methods for the higher-order algebraic eigenvalue problem. Time forcing solutions are decomposed in such a way that the initial values of particular responses induce free responses. This is of particular interest in numerical integration as well as signal processing. By focussing our attention on the fundamental matrix function, that can be given in a closed form, we can also characterize the asymptotic behaviour of solutions.

The eigenanalysis is considered in connection with the power method and the Krylov method. The validity of this later method relates to controllability matrix associated with higher-order systems. We also derive an iterative formula, in terms of the fundamental matrix basis, for the powers of a block companion matrix that allows to obtain a translation property for the fundamental matrix function.

Matricial form of finite differences problems: a new approach

Claire David
Université Pierre et Marie Curie-Paris 6
Laboratoire de Modélisation en Mécanique, UMR CNRS 7607
Boîte courrier $n^0$162, 4 place Jussieu, 75252 Paris cedex 05, France
david@lmm.jussieu.fr

The aim of our work is to develop general optimization methods for finite difference schemes used to approximate linear differential equations. The specific case of the transport equation

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0 \tag{1}$$

with the initial condition $u(x,0) = u_0(x)$ is exposed.
To begin with, finite differences schemes are grouped into classes corresponding to specific values of a given set of parameters, which can be handled through symbolic calculus. We show that finite difference problems can be solved matricially, without loops. The theoretical study of the related linear parametric matricial system, that bears the initial condition, yields general results which enables to study the intrinsic properties of a given scheme. In particular, the minimization of the numerical error is taken into account.

Tensor decompositions: an overview

Lieven De Lathauwer
CNRS - Lab. ETIS
6, avenue du Ponceau, BP 44, F95014 Cergy-Pontoise, France
delathau@ensea.be

In data analysis, mathematical modelling, signal processing and scientific computing, multilinear algebraic techniques are becoming increasingly popular. The algebra of higher-order tensors is more complex but also richer than the algebra of matrices and vectors. Fundamental concepts, such as rank, can be generalized in different ways. Basic decompositions, like the Singular Value Decomposition, have several higher-order counterparts. In this talk we give an overview of the state-of-the-art.

Stability of group-theoretic algorithms for fast matrix multiplication

James Demmel
Computer Science Division
University of California-Berkeley
Berkeley, CA 94720, USA
demmel@cs.berkeley.edu

Ioana Dumitriu, Olga Holtz, Robert Kleinberg
University of California-Berkeley
Berkeley, CA 94720
{dumitriu@math,holtz@math,rbk@cs}.berkeley.edu

We perform forward error analysis for a large class of recursive matrix multiplication algorithms in the spirit of [1]. We show that new group-theoretic algorithms proposed in [3] and [2] are all included in that larger class and are numerically stable. We also analyze in detail the stability of three specific fast group-theoretic algorithms, including one with the (current record) running time of Coppersmith and Winograd [4].

References

[1] D. Bini and G. Lotti, *Stability of fast algorithms for matrix multiplication*, Numer. Math., 36 (1980/81), pp. 63–72.

[2] H. Cohn, R. Kleinberg, B. Szegedy, and C. Umans, *Group-theoretic algorithms for matrix multiplication*, in Foundations of Computer Science. 46th Annual IEEE Symposium on 23–25 Oct 2005, 2005, pp. 379–388.

[3] H. Cohn and C. Umans, *A group-theoretic approach to matrix multiplication*, in Foundations of Computer Science. 44th Annual IEEE Symposium, 2003, pp. 438–449.

[4] D. Coppersmith and S. Winograd, *Matrix multiplication via arithmetic progressions*, J. Symbolic Comput., 9 (1990), pp. 251–280.

The future of LAPACK and ScaLAPACK

James Demmel
Computer Science Division
UC Berkeley
Berkeley, CA 94720, USA
demmel@cs.berkeley.edu

Sven Hammarling
The Numerical Algorithms Group, Ltd
Wilkinson House, Jordan Hill Road, Oxford, OX2 8DR, UK
sven@nag.co.uk

We are planning new releases of the widely used LAPACK and ScaLAPACK numerical linear algebra libraries.
Based on an on-going user survey (www.netlib.org/lapack-dev) and research by many people, we are proposing the following improvements: Faster algorithms (including better numerical methods, memory hierarchy optimizations, parallelism, and automatic performance tuning to accomodate new architectures), more accurate algorithms (including better numerical methods, and use of extra precision), expanded functionality (including updating and downdating, new eigenproblems, etc. and putting more of LAPACK into ScaLAPACK), and improved ease of use (friendlier interfaces in multiple languages).
After an overview, this two part talk will highlight new more accurate algorithms, faster algorithms for polynomial roots and matrix multiplication, pivoted Cholesky, fast updating of factorizations, and hybrid data formats.
To accomplish these goals we are also relying on better software engineering techniques and contributions from collaborators at many institutions. This is joint work with Jack Dongarra.

Two preconditioners for the saddle point equation

Arie de Niet
Insitute for Mathematics and Computing Science, University of Groningen
P.O.Box 800, 9700 AV Groningen, The Netherlands
a.c.de.niet@math.rug.nl

Fred Wubs
Insitute for Mathematics and Computing Science, University of Groningen
P.O.Box 800, 9700 AV Groningen, The Netherlands
f.w.wubs@math.rug.nl

Commonly used finite-element and finite-difference discretizations of the equations describing an incompressible flow, like for example the Navier-Stokes equations, lead to an equation $Kx = b$, where K is of saddle point type:

$$K = \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix}.$$

In this matrix we have $A \in \mathbb{R}^{n\times n}$, a positive definite matrix, not necessarily symmetric, and $B \in \mathbb{R}^{n\times m}$, with $m \leq n$.
From literature a number of good preconditioners for various types of saddle point matrices is available. We would like to add two alternative preconditioners [1] to the collection, namely the following:

$$\hat{K}_{GD}(\omega) = \begin{pmatrix} A + \omega BB^T & 0 \\ 0 & I/\omega \end{pmatrix} \text{ and } \hat{K}_{AC}(\omega) = \begin{pmatrix} A & B \\ B^T & -I/\omega \end{pmatrix}.$$

The first is based on grad-div stabilization and the second on artificial compressibility. Both techniques are usually applied to the original equation, but we will use it in the preconditioning. Eigenvalue analysis shows that with these preconditioners we can achieve low condition numbers for the preconditioned matrix.
The analysis is supported by numerical experiments. We will compare our preconditioners to SIMPLE(R) [4, 5] and the preconditioners of Elman, Wathen, Silvester, Kay and Loghin [6, 2, 3] for three equations: Stokes in a driven cavity, the Oseen equations with a given wind field and a saddle point problem from ocean circulation. The last problem models also the Coriolis force, which gives a strong skew symmetric contribution to A. The success of the preconditioners is problem dependent. The new ones work well on all three problems; the others show serious convergence deterioration on the ocean saddle point problem, which can be cured for some of them by a suitable adaptation. It turned out, that in both number of iterations and overall computation time the new preconditioners are faster than the preconditioners from literature.

References

[1] A. C. de Niet and F. W. Wubs, *Two preconditioners for the saddle point equation*, tech. rep., Institute for mathematics and computing Science, University of Groningen, http://www.math.rug.nl/~wubs/reports/precSPE.pdf, 2006.

[2] H. C. Elman and D. J. Silvester, *Fast nonsymmetric iterations and preconditioning for Navier-Stokes equations*, SIAM J. Sci. Comput., 17 (1996), pp. 33–46.

[3] D. Kay, D. Loghin, and A. J. Wathen, *A preconditioner for the steady-state Navier-Stokes equations*, SIAM J. Sci. Comput., 24 (2002), pp. 237–256 (electronic).

[4] C. Li and C. Vuik, *Some results on the eigenvalue analysis of a SIMPLER preconditioned matrix*, Report 03-08, Delft University of Technology, Department of Applied Mathematical Analysis, Delft, 2003.

[5] ———, *Eigenvalue analysis of the SIMPLE preconditioning for incompressible flow*, Numer. Linear Algebra Appl., 11 (2004), pp. 511–523.

[6] D. J. Silvester and A. J. Wathen, *Fast iterative solution of stabilised Stokes systems. II. Using general block preconditioners*, SIAM J. Numer. Anal., 31 (1994), pp. 1352–1367.

Convergence issues in recycling Krylov subspaces

Eric de Sturler
Department of Mathematics, Virginia Tech
Department of Mathematics, 460 McBryde, Virginia Tech, Blacksburg, VA 24061-0123
sturler@vt.edu

Michael L. Parks
Sandia National Laboratories
Sandia National Labs, Albuquerque
mlparks@sandia.gov

Many problems in science and engineering require the solution of a long sequence of linear systems, with small changes from one matrix to the next but substantial changes over multiple systems. We are particularly interested in cases where both the matrix and the right hand side change and systems are not available simultaneously. Such sequences arise in time-dependent iterations, nonlinear systems of equations and optimization, (distributed) parameter identification and inverse problems, and many other problems.
In recent papers [2, 1, 3] we have proposed methods to recycle selected subspaces from the Krylov spaces generated for previous linear systems to improve the convergence of subsequent linear systems. In this presentation, we discuss several important convergence issues:

- the convergence of Krylov methods that recycle approximate solution spaces, approximate invariant subspaces, and other relevant spaces,
- the relevant perturbation theory for the spaces mentioned above for sequences of matrices arising in a range of applications,
- how fast our proposed methods learn to adapt to a changing problem.

We provide experimental results for a range of problems from tomography, nonlinear mechanics, design optimization, and statistical mechanics.

References

[1] M. Kilmer and E. de Sturler, *Recycling subspace information for diffuse optical tomography*, SIAM Journal on Scientific Computing (accepted for publication), xx (2006). available from http://www-faculty.cs.uiuc.edu/ sturler.

[2] M. L. Parks, E. de Sturler, G. Mackey, D. D. Johnson, and S. Maiti, *Recycling krylov subspaces for sequences of linear systems*, SIAM Journal on Scientific Computing (accepted with minor revisions), xx (2006). available as Tech. Report UIUCDCS-R-2004-2421, March 2004, from http://www-faculty.cs.uiuc.edu/ sturler.

[3] S. Wang, E. de Sturler, and G. H. Paulino, *Large-scale topology optimization using preconditioned krylov subspace methods with recycling*, International Journal for Numerical Methods in Engineering (submitted), xx (2006). available as Technical Report UIUCDCS-R-2006-2678 from http://www-faculty.cs.uiuc.edu/ sturler.

On the change of canonical structures under low rank perturbations: from Jordan to Weierstrass and Kronecker

Fernando De Terán
Departamento de Matemáticas, Universidad Carlos III de Madrid
Avda. Universidad, 30 28911 Madrid, Spain
fteran@math.uc3m.es

Froilán M. Dopico, Julio Moro
Departamento de Matemáticas, Universidad Carlos III de Madrid
Avda. Universidad, 30 28911 Madrid, Spain
{dopico,jmoro}@math.uc3m.es

In this talk we present an overview of the problem concerning the change of the usual spectral canonical forms of matrices and matrix pencils after additive perturbations with low rank. We begin by introducing the matrix case, which has been a subject of interest during the last few years. We present the results contained in [1], where not only the typical change was described, but also necessary and sufficient conditions were given for this change to hold. Similar techniques as the ones used in that paper have allowed us to obtain the generic change of the Weierstrass canonical form for low rank perturbations of regular matrix pencils [2]. Also, necessary and sufficient conditions are obtained. We finally present recent work relative to the change of the Kronecker structure of singular matrix pencils. We describe also the generic change and provide sufficient conditions under which this change holds. A relevant point is that the generic behavior we find for this singular case has nothing to do with the behavior for regular matrix pencils. Besides, the singular case requires very different mathematical techniques.

References

[1] J. Moro and F. M. Dopico, *Low rank perturbation of Jordan structure*, SIAM J. Matrix Anal. Appl. 25(2) (2003), pp. 495-506.

[2] F. De Terán, F. M. Dopico and J. Moro, *Low rank perturbation of Weierstrass structure*, submitted.

From Shannon to Von Neumann: new distance measures for matrix nearness problems

Inderjit S. Dhillon
Department of Computer Sciences, University of Texas at Austin
1 University Station C0500, Austin, TX 78712-0500, USA
inderjit@cs.utexas.edu

A recurring problem in matrix computations is to find a structured matrix that best approximates a given matrix with respect to some distance measure. For example, it may be known a priori that a certain constraint ought to hold, and yet it fails on account of measurement errors or numerical roundoff. An attractive remedy is to replace the tainted matrix by the nearest matrix that does satisfy the constraint. Matrix nearness problems typically measure the distance between matrices with a norm. The Frobenius and spectral norms are pervasive choices because they are typically analytically tractable. Nevertheless, these norms are not always defensible in applications, where it may be wiser to tailor the distance measure to the context. In this talk, I will discuss two such applications that arise in data analysis. In the first application, the input matrix may be viewed as a two-dimensional joint probability distribution, in which case we use a distance measure motivated by the Shannon entropy of the matrix elements. In the second application, the desired matrix is a positive semi-definite matrix that satisfies specified linear constraints, in which case a distance measure derived from the von Neumann entropy is suitable.

Parameter estimation in Markov jump processes

Evelyn Dittmer
Fachbereich Mathematik + Informatik WE 2, Numerische Mathematik/Scientific Computing,
Freie Universität Berlin
Arnimallee 6, D-14195 Berlin, Germany
dittmer@math.fu-berlin.de

Illia Horenko, Philipp Metzner, Christof Schütte
Fachbereich Mathematik + Informatik WE 2, Numerische Mathematik/Scientific Computing,
Freie Universität Berlin
Arnimallee 6, D-14195 Berlin, Germany
{horenko, metzner, schuette}@math.fu-berlin.de

Hidden Markov models are powerful tools that provide simple descriptions of complex dynamics. To overcome many restrictions in the standard hidden Markov model, we present here a modified version where the output process is governed by a Smoluchowski or Langevin [2] equation and the hidden process is given by a Markov jump process.
In time series where the observations are not recorded at equidistant time points an estimation of the generator is required. The generator estimation is not trivial since not every transition matrix has a generator, i.e. P(t) is not necessarily of the shape $P(t) = \exp(tL)$, where L fulfills the generator constraints. For this problem different approaches exist that we will enclose in the theory of hidden Markov models. The direct estimation of the generator from the likelihood function also is a non-trivial task. Yet as proposed in [1] the maximum likelihood estimator can be determined by the EM algorithm, where in each iteration the generator estimator is computed from the solutions of two ODEs.

References

[1] M. Bladt and M. Sorensen, *Statistical inference for discretely observed markov jump processes*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67 (2005), pp. 395–410.

[2] I. Horenko, E. Dittmer, A. Fischer, and C. Schütte, *Automated model reduction for complex systems exhibiting metastability*, accepted for Multiscale Modeling and Simulation, (2006).

Efficient update of hierarchical matrices for BEM

Jelena Djokić
Scientific Computing – Max-Planck-Institut für Mathematik in den Naturwissenschaften
Inselstr. 22, D-04103 Leipzig, Germany
djokic@mis.mpg.de

Certain PDE's (in the simplest case Laplace's equation with Dirichlet(Neumann) boundary conditions) can be reformulated as integral equations (e.g Fredholm integral equation of the first kind) which possess a unique but, for practical applications, not explicitly computable solution. The main reason for the reformulation is the simplification of the original task. While the PDE should be solved in the domain $\Omega \subset \mathbb{R}^d$, $d = \{2, 3\}$, the integral equation is posed on a hypersurface $\Gamma \subseteq \mathbb{R}^d$. The numerical task is to provide an approximated but useful solution to the integral equation using the properties of the kernel function $g(x, y)$ that defines the integral operator. In almost all practical problems we consider, the kernel function of the integral operator has singularities only on the diagonal, i.e., on the set $\{(x, y) \in \Gamma \times \Gamma \mid x = y\}$.
Very popular and widely used methods of solving integral equations (of the first kind) numerically are discretisation schemes like Ritz-Galerkin or collocations. Those methods solve the integral equation approximately by solving a system of linear equations. The problem that arises is in the matrix of the system of linear equations which is, in the general case, densely populated. In order to overcome this obstacle we can approximate that matrix using methods like multipole, wavelets or hierarchical matrices that will be focus of this talk.
he hierarchical matrix technique (or briefly $\mathcal{H}$-matrix technique) has been developed during the past ten years. The main property of the hierarchical matrices is their data-sparse structure (can be described by few data) and the main advantage is that $\mathcal{H}$-matrix arithmetics can be performed in almost optimal complexity $\mathcal{O}(n \log n)$ for $n \times n$ systems. The $\mathcal{H}$-matrix technique has been efficiently applied in many research fields (FEM, BEM, control theory).
In the first part of our talk we will present the construction of $\mathcal{H}$-matrices and some BEM applications. The second part of the talk will be devoted to the update of hierarchical matrices in the case of adaptive discretisation schemes. The question we answer is whether is possible to update the $\mathcal{H}$-matrix instead of constructing new one if the discretisation scheme is locally refined.

Regularized saddle point problems and implicit-factorization preconditioners

Sue Dollar
University of Reading
Department of Mathematics, Whiteknights, PO Box 220, Berkshire, RG6 6AX, UK.
hilary.dollar@keble.oxon.org

Nick Gould
Rutherford Appleton Laboratory
Numerical Analysis Group, Computational Science & Engineering Department, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire, OX11 0QX, UK.
N.I.M.Gould@rl.ac.uk

Wil Schilders
Philips Research Laboratories
Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands
wil.schilders@philips.com

Andy Wathen
Oxford University Computing Laboratory
Numerical Analysis Group, Wolfson Building, Parks Road, Oxford, OX1 3QD, UK.
Andy.Wathen@comlab.ox.ac.uk

The problem of finding good preconditioners for the numerical solution of a certain important class of indefinite linear systems is considered. These systems are of a regularized saddle point structure

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix},$$

where $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{m\times m}$ are symmetric and $B \in \mathbb{R}^{m\times n}$.
In *Constraint preconditioning for indefinite linear systems,* SIAM J. Matrix Anal. Appl., 21 (2000), Keller, Gould and Wathen analyze the idea of using constraint preconditioners that have a specific 2 by 2 block structure for the case of C being zero. We shall extend this idea by allowing the (2,2) block to be symmetric and positive semi-definite (extended constraint preconditioners): results concerning the spectrum and form of the eigenvectors will be presented.
At first glance, the proposed form of preconditioner appears to be as expensive to apply as solving the original system directly. However, implicit factorization preconditioners can be used to dramatically reduce the number of flops required by the preconditioning step: the second half of the talk will introduce some of these factorizations. Numerical results will confirm that they can be very effective in practice.

Regularization by multigrid-type algorithms

Marco Donatelli
Dipartimento di Fisica e Matematica, Università dell'Insubria - Sede di Como
Via Valleggio 11, 22100 Como, Italy
marco.donatelli@uninsubria.it

We consider the de-blurring problem of noisy and blurred images in the case of space invariant point spread functions (PSFs). The use of appropriate boundary conditions leads to linear systems with structured coefficient matrices related to space invariant operators like Toeplitz, circulants, trigonometric matrix algebras etc. We can obtain an effective and fast solver by combining the optimally convergent algebraic multigrid described in [6, 2] with the Tikhonov regularization (see [4]). A theoretical proof of such optimal behavior in terms of convergence speed and total computational cost is provided in [1, 2]. A completely alternative proposal is to apply the latter algebraic multigrid (which is designed ad hoc for structured matrices) with the low-pass projectors typical of the classical geometrical multigrid employed in a PDEs context. Thus, using an appropriate smoother (any iterative regularizing method like conjugate gradient (CG), conjugate gradient for normal equation (CGNE), Landweber etc.), we obtain an iterative regularizing method (see [5] and especially [3]). The resulting iterative regularizing multigrid method proposed and discussed in [5] is based on: **1)** projection in a subspace where it is easier to distinguish between the signal and the noise, **2)** application of an iterative regularizing method in the projected subspace. The projector is chosen according to [6, 2] in order to maintain the same algebraic structure at each recursion level and having a low-pass filter property, which is very useful in order to reduce the noise effects. In this way, we obtain a better restored image with a flatter restoration error curve and also in less time than the auxiliary method used as smoother. A simple choice of the parameters allows to devise a powerful regularizing method whose main features are the following: **a)** it is used with early stopping like any regularizing iterative method and its cost per iteration is about 1/3 of the cost of the method used as smoother (CG, Landweber, CGNE); **b)** it can be adapted to work with all the boundary conditions used in literature (Dirichlet, periodic, Neumann or reflective or anti-reflective) since the basic algebraic multigrid considered in [2, 1] is an optimally convergent method for any of the involved structures (Toeplitz, circulant, cosine-algebra or sine-algebra) which naturally arise from the chosen boundary conditions; **c)** the minimal relative restoration error with respect to the true image is significantly lower with regard to the method used as smoother and the associated curve of the relative restoration errors with respect to the iterations is "flatter" (therefore the quality of the reconstruction is not critically dependent on the stopping iteration); **d)** when it is applied to the system $A\mathbf{f} = \mathbf{g}$ the minimal relative error is comparable with regard to all the best known techniques for the normal equations $A^TA\mathbf{f} = A^T\mathbf{g}$, but in this case the convergence is much faster. As direct consequence of c), the choice of the exact iteration where to stop is less critical than in other regularizing iterative methods while, as a consequence of d), we can choose multigrid procedures which are extremely more efficient than classical techniques without losing accuracy in the restored image. Several numerical experiments show the effectiveness of our proposals.

Work partly joint with Stefano Serra Capizzano.

References

[1] A. Aricò and M. Donatelli, *A v-cycle multigrid for multilevel matrix algebras: proof of optimality*, Numer. Math., to appear.

[2] A. Aricò, M. Donatelli, and S. Serra-Capizzano, *V-cycle optimal convergence for certain (multilevel) structured linear systems*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 186–214 (electronic).

[3] M. Donatelli, *Image Deconvolution and Multigrid Methods*, PhD Thesis in Applied and Computational Mathematics, Univ. of Milano, 2005.

[4] ———, *A multigrid for image deblurring with Tikhonov regularization*, Numer. Linear Algebra Appl., 12 (2005), pp. 715–729.

[5] M. Donatelli and S. Serra-Capizzano, *On the regularizing power of multigrid-type algorithms*, SIAM J. Sci. Comput., in press.

[6] S. Serra Capizzano, *Convergence analysis of two-grid methods for elliptic Toeplitz and PDEs matrix-sequences*, Numer. Math., 92 (2002), pp. 433–465.

A parametrization of the group of symplectic matrices and its applications

Froilán M. Dopico
Departamento de Matemáticas, Universidad Carlos III de Madrid
Avenida de la Universidad, 30, 28911 Leganés (Madrid), Spain
dopico@math.uc3m.es

Charles R. Johnson
Department of Mathematics, The College of William and Mary
Williamsburg, VA, 23187-8795 USA
crjohnso@math.wm.edu

Symplectic matrices play an important role in applications as Hamiltonian Mechanics or linear control theory for discrete time systems. A matrix S is symplectic if $S^T JS = J$, where $J = [0, I_n; -I_n, 0]$, so they are defined as the solutions of a nonlinear matrix equation. This fact makes difficult to work with them both from a theoretical and numerical point of view. In this talk we present an explicit parametrization of the group of symplectic matrices and we present an algorithm to compute this parametrization in a stable way if we are given the entries of a symplectic matrix S. Some potential applications of this algorithm are discussed. Moreover this parametrization allows us to solve many theoretical questions on symplectic matrices, as for instance, to study the intersection between the set of symplectic matrices and other important sets of matrices (with nonnegative entries, totally positive, M-matrices ... and to study other parametrization problems (how to describe the set of symplectic matrices with given n first columns, ...

Floating point issues in rank revealing QR factorization and its applications

Zlatko Drmač
Department of Mathematics, University of Zagreb
Bijenička 30, 10000 Zagreb, Croatia
drmac@math.hr

Zvonimir Bujanović
Department of Mathematics, University of Zagreb
Bijenička 30, 10000 Zagreb, Croatia
zbujanov@math.hr

The purpose of this talk is to warn that most state of the art implementations (e.g. in LAPACK) of the QR factorization with column pivoting can catastrophically fail in an unusual, unexpected, and not previously reported way.
We analyze the failure, its consequences in applications, and a new implementation of column pivoting which has stronger form of backward stability and thus better forward error bound. It will be also shown how to exploit pivoting in perturbation theory of the QR factorization.
Our particular interest is the use of pivoted QR factorization as preconditioner for fast and accurate Jacobi–type SVD algorithm. In that context, we discuss several floating point issues in SVD computation of triangular matrices from pivoted QR factorization. In particular, we show how perturbation theory can be used to avoid slowdown caused by computation with denormalized numbers.

Feature-preserving higher-order geometric regularization

Marc Droske
Department of Mathematics – University of California, Los Angeles
520 Portola Plaza,
Los Angeles, CA, 90055, USA
droske@math.ucla.edu

Andrea Bertozzi
Department of Mathematics – University of California, Los Angeles
520 Portola Plaza,
Los Angeles, CA, 90055, USA
bertozzi@math.ucla.edu

Regularization techniques for curves and surfaces are an essential ingredient for a wide range of problems with evolving geometries, especially in the field of shape analysis and inverse problems. For example, active contours form a class of variational methods, based on nonlinear PDEs, for image segmentation. Typically these methods introduce a local smoothing of edges due to a length minimization or minimization of a related energy. Such approaches have a tendency to smooth corners, which can be undesirable for tasks that involve identifying man-made objects with sharp corners. We consider a new method in which the local geometry of the curve is incorporated into the dynamics in a nonlinear way. Our method is based on ideas from image denoising and simplification of high contrast images using higher-order PDEs, in which piecewise linear shapes are preserved. Furthermore we will discuss a geometric analogue of the Mumford-Shah functional for surfaces, in which the "free discontinuity set" plays the role of a subset of the surface on which no regularity is enforced, leading to the functional

$$E[\mathcal{M},\Gamma] = \int_{\mathcal{M}} \operatorname{dist}(\cdot,\mathcal{M}_0)^p \,\mathrm{d}A + \frac{\mu}{2}\int_{\mathcal{M}\setminus\Gamma} h^2 \mathrm{d}A + \nu|\Gamma|,$$

where h denotes the mean curvature of the free manifold $\mathcal{M}$ and $\Gamma \subset \mathcal{M}$. This approach leads to a coupled higher-order variational system, which is defined over geometric entities of different codimension. In a wide range of identification problems, the incorporation of a-priori information is a very promising approach to make the resulting method more robust. In a shape-analytic context, the mean curvature is a natural intrinsic geometric descriptor. Even though it does not completely characterize the surface, it turns out to be a useful geometric penalization term of to consider the difference of the curvature of the deformed object with that of the reference object. In the level set framework, this involves the transport of the reference curvature in the Eulerian frame, which has to be coupled to the geometric energy on the current geometric configuration.
The above framework leads to higher order geometric variational problems. We will discuss efficient numerical approaches based on gradient flows with suitable metrics and multi-scale techniques.

References

[1] M. Droske and A. Bertozzi, *Feature preserving higher-order regularization of curves and surfaces*, SIAM MMS, (2005). In preparation.

On continuum limits of discrete inverse problems

Vladimir Druskin
Schlumberger Doll Research
320 Bent st., Cambridge, MA 02141-2025, US
druskin1@slb.com

Liliana Borcea, Fernando Guevara Vasquez
Rice University
Computational and Applied Mathematics, MS 134, Rice University, 6100 Main Street, Houston, TX 77005-1892, USA
borcea@caam.rice.edu

Leonid Knizhnerman
Central Geophysical Expedition
Central Geophysical Expedition, Moscow, Russia
mmd@cge.ru

To evaluate $f(A)b$, Krylov subspace methods approximate A by small structured matrices T, in particular, by the tridiagonal (Jacobi) matrix, that matches some moments of the spectral measure of A. One can consider A as a second order differential operator and T as the matrix of the three-point (second order) finite-difference (FD) operator [Druskin, Knizhnerman, SINUM, **37**, 2, 2000, 403-426], hence the name FD Gaussian rules (FDGR). Originally the FDGR was intended as a tool for the solution of exteriour PDE problems and domain decomposition, where it has been quite succesfull. Here we consider another application. We solve the 1D inverse spectral problem and the 2D inverse impedance problem using network or FD approximations. Until this work it was not even clear how to obtain the solution of the continuum 1D problem as the limit of the FD (Jacobi) inverse eigenvalue problem [Chu, Golub, Inverse Eigenvalue Problem, Oxford University Press, 2005]. We begin with FD approximations of solutions of Sturm–Liouville equations. We discretize the equations on so-called optimal grids generated by the FDGR. The simplest optimal grid is constructed as follows. For a staggered grid with k primary and k dual points, we ask that the finite difference operator ($k \times k$ Jacobi matrix) and the Sturm–Liouville differential operator share the k lowest eigenvalues and the values of the orthonormal eigenfunctions, at one end of the interval. This requirement determines uniquely the entries in the Jacobi matrix, which are grid cell averages of the coefficients in the continuum problem. If these coefficients are known, we can find the grid. The point is that in the inverse problem neither the coefficients nor the grid are known. So, a key question is how to choose the grid. We prove, the optimal grid dependence on the unknown coefficients is weak, more precisely, the optimal grid becomes asymptotically independent of the coefficients (under some regularity conditions) as the number k of data points tends to infinity. As a result, an optimal grid computed for a known coefficient (for example, for a constant) gives coefficients that converge pointwise to the true (unknown) solution [Borcea et al, Comm. Pure Appl. Math., **58**, 9, 1231-1279, 2005]. The analysis is based on a novel, explicit perturbation analysis of Lanczos recursions [Druskin et al, LAA, 396, 2005, 103-125]. Finally, we consider extensions to the 2D inverse problems of the electrical impedance tomography using approximations on planar graphs. If time permits, we will discuss implications of the obtained results to the FD solution of multidimensional PDEs with variable coefficients.

On the convergence of additive Schwarz preconditioned GMRES

Xiuhong Du
Department of Mathematics, College of Science and Technology, Temple University
Philadelphia, PA, 19122, USA
dxhdxh@temple.edu

Daniel B. Szyld
Temple University
Department of Mathematics, College of Science and Technology, Temple University
Philadelphia, PA, 19122, USA
szyld@temple.edu

Additive Schwarz preconditioners, when a coarse grid correction is added, are said to be optimal for certain discretized PDE problems, in the sense that bounds on the convergence of iterative methods are independent on the mesh size. Cai and Zou [Numer. Linear Algebra Appl., 9:379–397, 2002] showed with an example that in the absence of a coarse grid correction the usual GMRES bound has a factor of the order of $1/\sqrt{h}$. In this paper we consider the same one-dimensional example (as well as a two-dimensional counterpart) and show that the behavior of the method is not well represented by the above mentioned bound. We use a different bound for GMRES from [Simoncini and Szyld, SIAM Rev., 47:247–272, 2005] and show that the relevant factor is bounded by $c - O(h)$, so that as $h \to 0$, it approximates a constant. Furthermore, for a sequence of meshes the convergence curves are almost identical and the number of GMRES iterations needed for convergence has a very slow growth.

Preconditioner updates for sequences of sparse, large and nonsymmetric linear systems

Jurjen Duintjer Tebbens
Institute of Computer Science, Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2, 182 07 Libeň, Praha 8 Czech Republic
tebbens@cs.cas.cz

Miroslav Tůma
Institute of Computer Science, Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2, 182 07 Libeň, Praha 8 Czech Republic
tuma@cs.cas.cz

In many applications, such as computational fluid dynamics, structural mechanics or numerical optimization, one may need to solve sequences of linear systems. If the computation of efficient preconditioners for the individual linear systems of the sequence is expensive, updating previous preconditioners can be very beneficial. This has been done for large and sparse systems, among others, by recycling subspaces in the context of Krylov subspace methods [4], with small rank updates in case of applying Quasi-Newton methods [2] or by means of diagonal updates for SPD systems arising from parabolic PDE's [1].
Our contribution extends the approach from [1] to generally nonsymmetric systems. We define efficient updates by considering appropriate approximations of the difference between the matrix of a previous and the current linear system [3]. We envisage to address both these approximations as well as permutation techniques for the whole sequence that can make the updates even more powerful. In addition, we discuss matrix estimation strategies necessary to exploit the updates in a matrix-free environment.
This work is supported by the Program Information Society under project 1ET400300415.

References

[1] M. Benzi and D. Bertaccini, *Approximate inverse preconditioning for shifted linear systems*, BIT Numerical Mathematics, 43 (2003), pp. 231–244.

[2] L. Bergamaschi, R. Bru, A. Martinez, and M. Putti, *Quasi-Newton preconditioners for the inexact Newton method*, in the Abstract book of the International Conference On Preconditioning Techniques, Atlanta, May 19-21, 2005.

[3] J. Duintjer Tebbens and M. Tůma, *Preconditioner updates for solving sequences of large and sparse nonsymmetric linear systems*, submitted to SIAM J. Sci. Comput. in 2006.

[4] M. Parks, E. de Sturler, G. Mackey, D. Johnson, and S. Maiti, *Recycling Krylov subspaces for sequences of linear systems*, Technical Report UIUCDCS-R-2004-2421, University of Illinois, 2004.

Parallel numerical linear algebra systems

Alan Edelman
Massachusetts Institute of Technology, Department of Applied Mathematics
77 Massachusetts Avenue, Cambridge, MA 02139 USA
edelman@math.mit.edu

Numerical linear algebra remains at the heart of mathematical software packages, libraries, and research. The goal of obtaining performance from parallel computers continues in such research and commercial entities as the Scalapack project, any number of parallel "MATLAB"s, grid Mathematica, and Interactive Supercomputing's Star-P. By parallel MATLABs, we point out that there have been research projects that included add-ons to the MathWorks' MATLAB environment and its clones as well as parallel uses of the language independent of the product.
In this talk we consider an academic comparison of the approaches taken, how the race for performance has somewhat slowed progress in the field, and what lessons we have learned from the ISC Star-P parallel system project for use with Mathworks' MATLAB.

Preconditioning iterative regularization methods in Hilbert scales

Herbert Egger
Johann Radon Institute for Computational and Applied Mathematics (RICAM) – Austrian Academy of Sciences
Altenbergerstrsße 69, A-4040 Linz, Austria
herbert.egger@oeaw.ac.at

Andreas Neubauer
Institut für Industriemathematik, Johannes Kepler Universität Linz
Altenbergerstraße 69, A-4040 Linz, Austria
neubauer@indmath.uni-linz.ac.at

We study inverse problems of the form

$$Tx = y\,, \tag{1}$$

where $T : \mathcal{X} \to \mathcal{Y}$ is a linear operator between Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$. By assuming that (1) is ill-posed, we mean that solutions do not depend continuously on the data. Thus, they have to be regularized in order to obtain reasonable approximations, in particular if only approximate data y^δ with some upper bound on the noise level $\|y^\delta - y\|_{\mathcal{Y}} \le \delta$ are available. For large scale problems, iterative regularization methods are an attractive alternative to standard approaches like Tikhonov regularization, cf. [4]. For illustration, we consider Landweber's method, where stable approximations are found by the iteration [3]

$$x^\delta_{k+1} = x^\delta_k + T^*(y^\delta - Tx^\delta_k), \qquad k \ge 1,\ x^\delta_0 = x_0.$$

Stability is introduced by stopping the iteration at the right time, e.g., as soon as $\|Tx^\delta_k - y^\delta\| \le \tau\delta$ for some $\tau > 1$.

One of the major drawbacks of iterative methods is that usually many iterations are required to obtain optimal convergence rates (in terms of the noise level of the available data). For acceleration, we propose preconditioning in Hilbert scales, cf. [1, 2]: Assume that L is a densly, self adjoint, defined strictly positive operator in $\mathcal{X}$. Then L induces a family of spaces $\{\mathcal{X}_r\}_{r\in\mathbb{R}}$ with norms $\|x\|_s := \|L^s x\|$. In our examples L will typically be some differential operator. The main assumption of our approach to preconditioning is that the smoothing properties of the operator T can be estimated in terms of the operator L generating the Hilbert scale, i.e., we assume

$$\|T^*y\|_a = \|L^a T^* y\| \le \overline{m}\|y\|, \qquad \forall y \in \mathcal{Y},$$

and consider, e.g., the preconditioned Landweber iteration

$$x^\delta_{k+1} = x^\delta_k + L^{-2s}T^*(y^\delta - Tx^\delta_k),$$

where $-a/2 \le s \le 0$. As an alternative approach, we investigate itzerative regularization methods, in which the position of T^* and L^{-2s} is interchanged; such methods are applicable to a different class of problems under similar smoothness conditions on T^* as above.

We will show that by preconditioning in Hilbert scales, the number of iterations needed to obtain optimal convergence rates with respect to the noise level δ can essentially be reduced to the square root. Furthermore, we will indicate under what kind of nonlinearity conditions our results can be applied to nonlinear inverse problems and illustrate the effect of preconditioning by numerical examples.

References

[1] H. Egger, *Semiiterative regularization in hilbert scales*, SIAM J. Numer. Anal., (2005). accepted.

[2] H. Egger and A. Neubauer, *Preconditioning Landweber iteration in Hilbert scales*, Numer. Math., 101 (2005), pp. 643–662.

[3] M. Hanke, A. Neubauer, and O. Scherzer, *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems*, Numer. Math., 72 (1995), pp. 21–37.

[4] A. N. Tikhonov, *Regularization of incorrectly posed problems*, Soviet Math. Dokl., 4 (1963), pp. 1624–1627.

Equivalence of Laguerre–based model order reduction and moment matching

Rudy Eid
Lehrstuhl für Regelungstechnik, Technische Universität München
Boltzmannstr. 15, D-85748 Garching bei München, Germany
eid@tum.de

Behnam Salimbahrami
Lehrstuhl für Regelungstechnik, Technische Universität München
Boltzmannstr. 15, D-85748 Garching bei München, Germany
salimbahrami@tum.de

Boris Lohmann
Lehrstuhl für Regelungstechnik, Technische Universität München
Boltzmannstr. 15, D-85748 Garching bei München, Germany
lohmann@tum.de

In order reduction of large scale systems, one of the most successful approaches is based on matching some of the coefficients of the Taylor series expansion of the transfer function around a suitable point. This can be achieved by reducing the original system through applying a projection where the so called Krylov subspaces and numerically robust algorithms are involved [2].
Lately, to improve the results of moment matching, to preserve stability, or to provide an error bound, some researchers have tried to couple the Krylov-based reduction approach to different orthonormal series (Laguerre, Legendre, ...) as an alternative to the Taylor series. Instead of matching the moments, these alternative approaches match some of the coefficients of the corresponding series of the reduced and original models' transfer functions.
In [3], the Laguerre series was first used for the purpose of order reduction of state-space models, based on the fact that the related series expansion can be transformed to a Taylor series by applying a certain bilinear transformation. This property simplifies the reduction procedure and makes it possible to use the numerically robust and efficient algorithms implemented for moment matching. Later, this reduction approach has been further developed and generalized by several other authors based on [3] or using an independent approach [1].
The Laguerre-based order reduction is receiving more and more attention and started to be considered as an independent reduction approach which is developing parallel to that of moment matching. Hence, it is of great importance to investigate the connection between these two approaches.
This talk proves the equivalence between the Laguerre-based reduction methods and the classical moment matching approach by two ways: First, we use the basic definition of the moments as the values of the derivatives of the transfer function at the expansion point and second, by proving that the Krylov subspaces involved in moment matching and Laguerre-based approaches are exactly the same. Furthermore, as a generalization of this equivalence, it is shown that by matching the moments around a certain frequency point, a family of coefficients called generalized Markov parameters match, where the Laguerre coefficients are just a special case.

References

[1] Y. Chen, V. Balakrishnan, C. K. Koh, and K. Roy, *Model reduction in time domain using Laguerre polynomial and Krylov methods*, Proc. Design, Automation and Test in Europe, (2002).

[2] Freund, R. W., *Model reduction methods based on Krylov subspaces*, Acta Numerica, 12 (2003), pp. 267–319.

[3] L. Knockaert and D. De Zutter, *Laguerre-SVD reduced-order modeling*, IEEE Transactions on Microwave Theory and Techniques, 48 (2000), pp. 1469–1475.

Convergence of Krylov approximations for matrix functions

Michael Eiermann
Institut für Numerische Mathematik und Optimierung, Technische Universität Bergakademie Freiberg
D-09596 Freiberg, Germany
eiermann@math.tu-freiberg.de

Oliver G. Ernst
Institut für Numerische Mathematik und Optimierung, Technische Universität Bergakademie Freiberg
D-09596 Freiberg, Germany
ernst@math.tu-freiberg.de

The evaluation of $f(A)\boldsymbol{b}$, where $A \in \mathbb{C}^{n\times n}$, $\boldsymbol{b} \in \mathbb{C}^n$ and $f : \mathbb{C} \supset D \to \mathbb{C}$ is a function for which $f(A)$ is defined, is a common computational task. Besides the solution of linear systems of equations, which involves the reciprocal function $f(\zeta) = 1/\zeta$, by far the most important application is the time evolution of a system under a linear operator, in which case $f(\zeta) = f_t(\zeta) = e^{t\zeta}$ and time acts as a parameter t.
In many of the applications mentioned above the matrix A is large and sparse or structured, typically resulting from discretization of an infinite-dimensional operator. In this case evaluating $f(A)\boldsymbol{b}$ by first computing $f(A)$ is usually unfeasible, so that most of the algorithms for the latter task cannot be used. The standard approach for approximating $f(A)\boldsymbol{b}$ directly is based on a Krylov subspace $\mathcal{K}_m(A, \boldsymbol{b})$ of A with initial vector $\boldsymbol{b}$.
We investigate the convergence properties of Krylov approximations of the form $p_{m-1}(A)\boldsymbol{b}$, where p_{m-1} is an interpolation polynomial for f whose nodes are either given a-priori or are the Ritz values of A with respect to $\mathcal{K}_m(A, \boldsymbol{b})$. We are particularly interested in problems where f is an entire function.

Max algebra and its applications

Ludwig Elsner
Fakultät für Mathematik, Universität Bielefeld
Postfach 10 01 31, D-33501 Bielefeld, Germany
elsner@math.uni-bielefeld.de

This lecture gives an overview about recent as well as unpublished results of the author (joint with P. van den Driessche) concerning aspects of the max algebra.
We discuss among others methods for calculating the max eigenvalue and max eigenvectors of a nonnegative matrix, applications to the analytic hierarchy process (AHP), and the treatment of the special case of Toeplitz matrices with an application to music theory.

Transient behavior of differential-algebraic equations via pseudospectra of matrix pencils

Mark Embree
Department of Computational and Applied Mathematics, Rice University
6100 Main Street – MS 134, Houston, Texas 77005–1892 USA
embree@rice.edu

Howard Elman
Department of Computer Science, University of Maryland
College Park, Maryland 20742 USA
elman@cs.umd.edu

Lloyd N. Trefethen
Oxford University Computing Laboratory
Wolfson Building, Parks Road, Oxford OX1 3QD, United Kingdom
lnt@comlab.ox.ac.uk

The asymptotic stability of a variety of mechanical systems governed by the linear differential equation $Bx'(t) = Ax(t)$ can be determined by finding the rightmost eigenvalues of the generalized eigenvalue problem $Ax = \lambda Bx$. In many cases this asymptotic analysis should be complemented by some characterization of the system's transient behavior, which can be derived from sources including eigenvectors, the numerical range, polynomial numerical hulls, and pseudospectra of related operators. We argue in [2] that to generalize the conventional ε-pseudospectrum to the matrix pencil setting, one should follow Ruhe's suggestion of considering the standard ε-pseudospectrum of $B^{-1}A$, measured in the norm most appropriate for the physical problem. Unlike alternative definitions of the pseudospectrum of a matrix pencil, this one remains invariant to premultiplication of the system by an invertible matrix, a transformation that also leaves unchanged the solutions of the underlying dynamical system.
Clearly this approach fails when B is singular, an essential case corresponding to systems that are modeled by differential-algebraic equations. This talk proposes a generalized definition for the pseudospectrum of a matrix pencil that is tailored to this setting, from which one can glean insight into the transient behavior of differential-algebraic systems. We shall compare this definition to alternatives proposed by Frayssé et al. [1] and van Dorsselaer [3] that are better suited for eigenvalue perturbation studies. We will also address computational challenges that arise when applying this analysis to large-scale systems, and illustrate these ideas with examples from fluid mechanics.

References

[1] V. Frayssé, M. Gueury, F. Nicoud, and V. Toumazou, *Spectral portraits for matrix pencils*, Tech. Rep. TR/PA/96/19, CERFACS, August 1996.

[2] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, 2005.

[3] J. L. M. van Dorsselaer, *Pseudospectra for matrix pencils and stability of equilibria*, BIT, 37 (1997), pp. 833–845.

Superlinear convergence of Krylov subspace methods: potential theory and invariant subspaces

Oliver G. Ernst
Institut für Numerische Mathematik und Optimierung, TU Bergakademie Freiberg
Akademiestr. 6, D-09596 Freiberg, Germany
ernst@math.tu-freiberg.de

Michael Eiermann
Institut für Numerische Mathematik und Optimierung, TU Bergakademie Freiberg
Akademiestr. 6, D-09596 Freiberg, Germany
eiermann@math.tu-freiberg.de

Despite over half a century of research, the convergence theory of Krylov subspace methods for solving linear systems of equations is still incomplete. Even for equations with normal matrices, rigorous convergence bounds typically predict asymptotic linear rates of convergence, although experience and many heuristic explanations indicate that often the convergence rate accelerates during the course of the iteration, so that convergence is in fact superlinear.
In this talk we give an overview of current approaches for convergence analysis and, based on recent work of Kuijlaars and Beckermann for the symmetric positive definite case, show how logarithmic potential theory may be employed to derive rigorous superlinear asymptotic convergence results for families of matrices whose spectral measures converge to a limit measure. We also discuss the role of approximate invariant subspaces contained in a Krylov space in producing superlinear convergence.

Structured linear algebra problems in image registration

Bernd Fischer
Institute of Mathematics, University of Lübeck
Wallstraße 40, D-23560 Lübeck, Germany
fischer@math.uni-luebeck.de

Jan Modersitzki
Institute of Mathematics, University of Lübeck
Wallstraße 40, D-23560 Lübeck, Germany
modersitzki@math.uni-luebeck.de

Image registration is one of today's challenging image processing problems. The objective is to find a geometrical transformation that aligns points in one view of an object with corresponding points in another view of the same object or a similar one. Particularly in medical imaging, there are many instances that demand for registration. It is inevitable whenever images acquired from different subjects, at different times, or from different scanners, need to be combined or compared for analysis or visualization. Typical examples include the treatment verification of pre- and post-intervention images, study of temporal series of images, the monitoring of time evolution of an agent injection subject to a patient-motion, and the fusion of computer tomography and magnetic resonance images.
Image registration, in particular medical image registration, has been subject to extensive studies in the past years. Its versatile and important applications have attracted researchers from various branches, including numerical analysts. Several techniques have been proposed to solve the registration problem. Most of them may be phrased in terms of the following model:
Given two images, a reference R and a template T, find a deformation $\mathbf{u}$, such that

$$\mathcal{D}[R, T; \mathbf{u}] + \alpha\mathcal{S}[\mathbf{u}] + \beta\mathcal{C}^{\text{soft}}[\mathbf{u}] = \min, \text{ subject to } \mathcal{C}^{\text{hard}}[\mathbf{u}] = 0.$$

Here, $\mathcal{D}$ denotes a distance measure which is designed to push the deformable template into the direction of the reference. It is well known that the intensity based registration is inherently an ill-posed problem. The regularizer $\mathcal{S}$ is introduced to overcome this problem and in addition may be used to incorporate features into the registration model, like the elastic behavior of the underlying tissue. The additional soft or hard constraints are designed to force the displacement to satisfy explicit criteria, like for example landmark or volume preserving imposed constraints.
There are various issues connected to this general formulation of the registration problem. The first one, concerns the proper choice of the defining building blocks. A problem which is clearly application dependent. The next question regards the existence and uniqueness of different realizations of the problem. Finally, it is by no means obvious on which numerical solution strategy one may employ. However, no matter on how one tries to minimize the joint functional, one ends up with the problem of solving a huge but highly structured linear system. For example, the registration of 512^3 images leads to a linear system of the order of 400.000.000 unknowns.
In this talk we will touch upon all three issues and we will point out some of the mathematical challenges associated with the image registration problem. The theoretical work is accompanied by various instructive (real life) examples.

Multigrid methods, structured matrices, and the Helmholtz equation

Rainer Fischer
Fakultät für Informatik, Technische Universität München
Boltzmannstr. 3, D-85748 Garching b. München, Germany
fischerr@in.tum.de

Thomas Huckle
Fakultät für Informatik, Technische Universität München
Boltzmannstr. 3, D-85748 Garching b. München, Germany
huckle@in.tum.de

Discretization of the Helmholtz equation with constant coefficients results in structured linear systems which are associated with generating functions. Depending on the type of boundary conditions one obtains matrices of a certain class such as two-level Toeplitz, tau, or DCT-III matrices. By solving these systems with normal equations, we have the advantage that the corresponding generating functions are nonnegative, although they have a whole curve of zeros.
The multigrid methods we develop are especially designed for structured matrix classes, making heavy use of the associated generating functions. They are based on the theory developed in [2, 1] for generating functions with isolated zeros. In [4] we have extended these multigrid methods from the isolated zero case to functions with zero curves. This is done by representing the whole zero curve on coarser grids, either by computing coarse grid matrices with the Galerkin approach or by rediscretization on coarser grids. Moreover, we have proved optimal convergence of the two-grid method and obtained fast convergence of the multilevel method in all numerical experiments as long as zero curves are not too large. Since zero curves become larger on coarser grids, the number of levels in our multigrid method is limited. To overcome this problem we propose a different approach which is based on splitting the original problem into a fixed number k of coarse grid problems which correspond to generating functions with isolated zeros. Locally representing one part of the zero curve, these coarse grid problems are straightforward to solve with the multigrid methods from [1] . We combine this splitting technique with the methods from [4] to construct a faster and more robust multigrid solver.
Our multigrid method can not only be used as a solver, but also as a preconditioner for the conjugate gradient method. When a zero curves becomes too large, it is approximated in the neighborhood of k points by anisotropic auxiliary problems. Since their generating functions have isolated zeros, they can be solved with the multigrid methods developed in [3] . The preconditioner is then constructed as the sum of the auxiliary problems.

References

[1] A. Arico, M. Donatelli, and S. Serra-Capizzano, *V-cycle optimal convergence for certain (multilevel) structured linear systems*, SIAM J. Mat. Anal. Appl., 26 (2004), pp. 186–214.

[2] G. Fiorentino and S. Serra, *Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions*, SIAM J. Sci. Comp., 17 (1996), pp. 1068–1081.

[3] R. Fischer and T. Huckle, *Multigrid methods for anisotropic BTTB systems*, Recommended for publication in Lin. Alg. Appl. (Special issue on the 80th birthday of F. L. Bauer).

[4] ———, *Multigrid methods for strongly ill-conditioned structured matrices*, submitted to ECOMAS Proceedings of the 8th European Multigrid Conference, (2005).

On the solution of symmetric indefinite systems that arise in optimization

Anders Forsgren
Optimization and Systems Theory, Department of Mathematics, Royal Institute of Technology (KTH)
SE-100 44 Stockholm, Sweden
andersf@kth.se

A major computational work in an interior method for optimization or a sequential-quadratic-programming method for optimization lies in the solution of a sequence of linear system of equations whose matrix is symmetric and indefinite. For interior methods, these matrices become increasingly ill-conditioned as the solution is approached. We discuss properties of the system of equations that arise, as well as propopse solution methods of factorization type and of iterative type. The solution methods are designed so as to detect local convexity of the model based on the inertia of the matrix.
The talk is based on joint work with Walter Murray, Philip E. Gill and Joshua D. Griffin.

Subgraph-based preconditioners for interior-point approaches to network flow problems

Antonio Frangioni
Dipartimento di Informatica – Università di Pisa
Largo B. Pontecorvo, 3 – I-56125 Pisa, Italy
frangio@di.unipi.it

Claudio Gentile
Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti" – Consiglio Nazionale delle Ricerche
Viale Manzoni, 30 – I-00185 Roma, Italy
gentile@iasi.cnr.it

Interior Point algorithms for Linear and Quadratic Programs provide a challenging environment for numerical analysis: a *sequence* of large-scale linear systems have to be solved efficiently with increasing accuracy, despite the fact that they become increasingly ill-conditioned. For general-purpose LP/QP codes, direct solvers based on Cholesky factorization are commonly used; however, these methods become prohibitively expensive for very large, very sparse problems such as network flow ones, in which the Cholesky factor may suffer a catastrophic fill-in [3]. After introducing the basic elements of the Interior Point approach from a numerical analysis perspective, we review experiences about using Preconditioned Conjugate Gradient (PCG) methods employing *subgraph-based preconditioners*. These preconditioners are based on computing "large weight" tree-like subgraphs of the graph associated with the system matrix; originally proposed by Vaidya, they have been computationally tested in [4] and extended in [2, 1].
In the context of Interior Point approaches, the preferred strategy has been to select a *triangulated* subgraph as the support of the preconditioner, so as to entirely avoid fill-in. From the theoretical analysis in [8] and [7], even the tree preconditioner can be expected to possess good spectral properties in the final iterations of an IP algorithm, even in the degenerate case; however, the picture is far less clear in the initial and middle iterations.
We report about previous results, focussing on the choice of the specific family of subgraphs between "simple" trees [9] and the "more advanced" Brother-Connected Trees of depth two [5], as well as about more recent results which show that even the specific algorithm used for computing the (approximately) maximum-weight subgraphs has an impact on the overall efficiency of the approach [6]. Our results show that a delicate balance has to be attained between the cost of forming and factoring a more complex preconditioner and the corresponding decrease in PCG iterations, for which developing exact guidelines is not straightforward. However, weight-based automatic switching rules for deciding when the BCT preconditioner can be used which seem to provide good average performances on a large set of instances.

References

[1] M. Bern, J. Gilbert, B. Hendrickson, N. Nguyen, and S. Toledo, *Support-graph preconditioners*, tech. rep., School of Computer Science, Tel-Aviv University, Tel-Aviv, 2001.

[2] E. Boman, D. Chen, B. Hendrickson, and S. Toledo, *Maximum-Weight-Basis-Preconditioners*, Numerical Linear Algebra and Applications, 11 (2004), pp. 695–721.

[3] J. Castro, *A specialized interior-point algorithm for multicommodity network flows*, SIAM Journal on Optimization, 10 (2000), pp. 852–877.

[4] D. Chen and S. Toledo, *Vaidya's preconditioners: implementation and experimental study*, Electronic Transactions on Numerical Analysis, 16 (2003), pp. 30–49.

[5] A. Frangioni and C. Gentile, *New Preconditioners for KKT Systems of Network Flow Problems*, SIAM Journal on Optimization, 14(3) (2004), pp. 894–913.

[6] ———, *Prim-based BCT preconditioners for Min Cost Flow Problems*, tech. rep., Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti", Consiglio Nazionale delle Ricerche, 2005.

[7] A. Frangioni and S. Serra Capizzano, *Spectral Analysis of (Sequences of) Graph Matrices*, SIAM J. on Matrix An. and Appl., 23(2) (2001), pp. 339–348.

[8] J. Jùdice, J. Patricio, L. Portugal, M. Resende, and G. Veiga, *A study of preconditioners for network interior point methods*, Computational Optimization and Applications, 24 (2003), pp. 5–35.

[9] M. Resende and G. Veiga, *An Implementation of the dual affine scaling algorithm for minumum cost flow on bipartite uncapacited networks*, SIAM Journal on Optimization, 3(3) (1993), pp. 516–537.

Extensions of certain graph-based algorithms for preconditioning

David Fritzsche
Department of Mathematics, Temple University
Philadelphia, PA 19122-6094
david.fritzsche@temple.edu

Andreas Frommer
Fachbereich C – Mathematik und Naturwissenschaften, Bergische Universität Wuppertal
Gauß-Str. 20, D-42097 Wuppertal, Germany
frommer@math.uni-wuppertal.de

Daniel B. Szyld
Department of Mathematics, Temple University
Philadelphia, PA 19122-6094
szyld@temple.edu

The original TPABLO algorithms are a collection of algorithms which compute a symmetric permutation of a linear system such that the permuted system has a relatively full block diagonal with relatively large nonzero entries. This block diagonal can then be used as a preconditioner. We propose and analyze three extensions of this approach: we incorporate an nonsymmetric permutation to obtain a large diagonal, we use a more general parametrization for TPABLO, and we use a block Gauss-Seidel preconditioner which can be implemented to have the same execution time as the corresponding block Jacobi preconditioner. Since our approach allows for efficient use of level 3 BLAS operations, it outperforms direct solvers and rivals standard ILU preconditioners on many test problems on a single processor system, while having good potential for efficient parallelization.

Computing $\mathrm{sign}(A)v$ with applications in QCD

Andreas Frommer
Fachbereich C – Mathematik und Naturwissenschaften, Bergische Universität Wuppertal
Gauß-Str. 20, D-42097 Wuppertal, Germany
frommer@math.uni-wuppertal.de

The computation of $\mathrm{sign}(Q)v$, with Q being the hermitian Wilson fermion matrix is the core operation in modern lattice QCD computations based on the so-called overlap fermion model. Here, Q is very large and sparse, so that computing $\mathrm{sign}(Q)$ as a matrix is definitely prohibitive, although $\mathrm{sign}(Q)v$ will have to be computed for several vectors v.
In this talk we present various projection approaches to compute $\mathrm{sign}(Q)v$ using Krylov subspaces and deflation techniques. We will give theoretical results on the quality of the approximations obtained and we show how efficient computational methods may be devised. In certain situations it is possible to obtain *a posteriori* error bounds.
This talk is based on joint work with Japser van den Eshof, Stefan Krieg, Thomas Lippert, Katrin Schäfer and Henk A. van der Vorst.

Perturbation of unitoid matrices

Susana Furtado
Faculdade de Economia da Universidade do porto
Rua Dr. Roberto Frias, 4200 Porto, Portugal
sbf@fep.up.pt

Charles R. Johnson
Mathematics Department, College of William and Mary
Williamsburg, VA 23187-8795, USA
crjohnso@math.wm.edu

A matrix $A \in M_n$ is called unitoid if it is congruent to a diagonal matrix. Necessary and sufficient conditions are given on the "canonical angles" of a unitoid matrix so that sufficiently small perturbations remain unitoid. This, in particular, resolves the question of when simultaneous diagonalizability of two Hermitian matrices is retained under perturbation.

PageRank computations for compressed link-structures [1]

Efstratios Gallopoulos
Computer Engineering and Informatics Department, University of Patras
CEID, University of Patras, Patras 26500, Greece
stratis@ceid.upatras.gr

Giorgos Kollias
Computer Engineering and Informatics Department, University of Patras
CEID, University of Patras, Patras 26500, Greece
gdk@hpclab.ceid.upatras.gr

PageRank is a simple iterative algorithm for computing a content-neutral ordering of Web pages. We investigate the effect of compression of the link-structure on this ordering. Compression patterns are either dynamically configured by successive iterates or statically selected by the specifics of a particular execution environment.
In the first context, we successively compress into supernodes those parts of the link-structure consisting of web pages tagged by low PageRank values. These dynamic perturbations of the link matrix generally yield a different ordering that appears to favor pages with maximum importance per inlink ratio. We present the dynamics of the orderings during iterations and compare to the non-perturbed case.
We also use compression forcibly adapted to concurrent execution environments, computed during a preprocessing stage. Here we favor uncoupled node aggregates to minimize communication, but doing so introduces deviations from the original PageRank solution. We comment on these perturbed solutions and experiment with possible mechanisms for keeping deviations to a minimum.

[1] This work was partially supported by PYTHAGORAS I, project B365016, under the EPEAK II program of the Greek Ministry of Educational and Religious Affairs

Structured initialization for head start in nonnegative matrix factorization

Efstratios Gallopoulos
Computer Engineering & Informatics Department, University of Patras
26500 Patras, Greece
stratis@ceid.upatras.gr

Christos Boutsidis
Computer Engineering & Informatics Department, University of Patras
26500 Patras, Greece
mpoutsid@ceid.upatras.gr

Approximate nonnegative matrix factorization, that is the approximation of a nonnegative matrix $A \in \mathbb{R}^{m\times n}$ by a product of nonnegative factors $W \in \mathbb{R}^{m\times k}$ and $H \in \mathbb{R}^{k\times n}$ of reduced rank $k < \min\{m, n\}$, has been attracting much attention because of its intrinsic interest and its important applications in areas such as Information Retrieval In particular, the nonnegative factors W and H can frequently provide useful information regarding the data in A; see e.g. [1, 3].
In the literature, there exist many iterative algorithms that attempt to minimize some distance function defined on $A - WH$. A common issue in these algorithms is their proper initialization [2, 4]; currently, one of the most frequent approaches is to initialize by means of random nonnegative matrices. In this presentation we describe some issues related to the initialization process and our recent results in this area. In particular, we show interesting structured alternatives to random initialization, based on the singular value decomposition, that appear to be very effective.
Work conducted in the context of and supported in part by a University of Patras KARATHEODORI grant.

References

[1] M. Berry, M. Browne, A. Langville, P. Pauca, and R. Plemmons, *Algorithms and applications for approximate nonnegative matrix factorization*. Submitted. Available from `www.wfu.edu/ plemmons/papers.htm`, Jan. 2006.

[2] C. Boutsidis and E. Gallopoulos, *On SVD-based initialization for nonnegative matrix factorization*, Tech. Rep. HPCLAB-SCG-6/08-05, Computer Eng. & Inform. Dept., University of Patras, Patras, Greece, Oct. 2005.

[3] D. D. Lee and H. S. Seung, *Learning the Parts of Objects by Non-Negative Matrix Factorization*, Nature, 401 (1999), pp. 788–791.

[4] S. Wild, J. Curry, and A. Dougherty, *Improving non-negative matrix factorizations through structured intitialization*, Pattern Recognition, 37 (2004), pp. 2217–2232.

Computing the hierarchy closure of the space of singular systems under derivative feedback

Ma. Isabel Garcia-Planas
Universitat Politècnica de Catalunya
C. Minerí 1, Esc C. 1-3, 08038 Barcelona Spain
maria.isabel.garcia@upc.edu

It is well known that computing the fine canonical structure elements of triples of matrices $(E, A, B) \in M_n(\mathbb{C}) \times M_n(\times M_{n\times m}(\mathbb{C})$ under feedback and derivative feedback corresponding to the singular systems $E\dot{x} = Ax+Bu$ are ill-posed problem because of arbitrary small perturbations in the entries may drastically change the canonical structure.
Besides knowing the canonical structure of a triple of matrices (E, A, B) is important to know what are the canonical structures what are nearby because they explain the behavior of the system under small perturbations.
One of the underlying theories to study tools for computing the qualitative information for triples of matrices is stratification of the space of triples of matrices under discrete invariants, because it reveals the complete hierarchy of nearby structures.

Perturbation of eigenvalues of matrix pencils and min-plus algebra

Stéphane Gaubert
INRIA
Domaine de Voluceau, 78153, Le Chesnay Cédex, France
Stephane.Gaubert@inria.fr

Marianne Akian
INRIA
Domaine de Voluceau, 78153, Le Chesnay Cédex, France
Marianne.Akian@inria.fr

Ravindra Bapat
Indian Statistical Institute
New Delhi, 110016, India
rbb@isid.ac.in

A result of Višik, Ljusternik and Lidskiĭ gives simple formulæ for the first order asymptotics $\mathcal{L}_\epsilon \sim \lambda\epsilon^\Lambda$ of the eigenvalues $\mathcal{L}_\epsilon$ of a perturbed matrix $\mathcal{A}_\epsilon = c + \epsilon d$, when ϵ tends to 0 (c and d are given complex matrices). When the matrix c is nilpotent, the leading exponents Λ are the inverses of the sizes of the Jordan blocks of c and the leading coefficients λ are obtained from the eigenvalues of certain Schur complements constructed from c and d, except when the matrices involved in this construction are singular.
In this talk, we present new perturbation formulæ obtained in [1, 2] using min-plus algebra techniques, which allow us to solve many singular cases. We also sketch the application of the methods of [1, 2] to the numerical computation of matrix eigenvalues.
To describe these results more precisely, let us consider a matrix pencil whose coefficients depend on a positive parameter ϵ, and have asymptotic equivalents of the form $a\epsilon^A$ when ϵ goes to zero (the leading coefficient a is complex, and the leading exponent A is real). We show that, generically, the leading exponents of the eigenvalues of the matrix pencil are the "eigenvalues" of a min-plus matrix pencil, and that their leading coefficients are the eigenvalues of auxiliary matrix pencils constructed from certain optimal assignment problems [2].
When considering the special case of the eigenvalues of a matrix $\mathcal{A}_\epsilon$ whose coefficients have equivalents of the form $a_{ij}\epsilon^{A_{ij}}$, when ϵ goes to zero, we give other formulæ for the leading exponents and leading coefficients of the eigenvalues of $\mathcal{A}_\epsilon$. These formulæ involve min-plus Schur complements of the matrix $A = (A_{ij})$ and usual Schur complements of the matrix $a = (a_{ij})$, provided the matrix A satisfies a structural condition expressed in terms of the existence of perfect matchings in certain digraphs [1].
The proof of the above results involves diagonal scalings. We shall discuss here how such diagonal scalings may be used in the numerical computation of eigenvalues and eigenvectors of matrices with complex entries, in particular when these entries have different orders of magnitude.

References

[1] M. Akian, R. Bapat, and S. Gaubert, *Min-plus methods in eigenvalue perturbation theory and generalised Lidskiĭ-Višik-Ljusternik theorem*. arXiv:math.SP/0402090, 2004.

[2] ———, *Perturbation of eigenvalues of matrix pencils and the optimal assignment problem*, C. R. Acad. Sci. Paris, Série I, 339 (2004), pp. 103–108. See also arXiv:math.SP/0402438.

On the development of numerical algorithms: a personal voyage

Gene H. Golub
Scientific Computing and Computational Mathematics Program
Stanford University
California, 94305-9025, USA
golub@sccm.stanford.edu

For over 40 years, I have been actively involved in developing numerical algorithms. The methods developed came about through the need of solving problems which often came from applications. Interestingly enough, each method often led to the need of others. In this talk, I will describe how these algorithms were developed sequentially.
Here are some algorithms/methods I shall describe.

- Cyclic Reduction
- Fast Poisson Solvers
- Domain Decomposition
- Generalized CG
- Bidiagonalization of Matrices
- Singular Value Decomposition

There are others and I will welcome questions.

Matrices and moments: perturbation for least squares

Gene H. Golub
Scientific Computing and Computational Mathematics Program
Stanford University
California, 94305-9025, USA
golub@sccm.stanford.edu

Zheng Su
Scientific Computing and Computational Mathematics Program
Stanford University
California, 94305-9025, USA
su@sccm.stanford.edu

Given a matrix A, $(m \times n)$, a vector $\vec{b}$, and an approximate solution vector, we are interested in determining approximate error bounds induced by the approximate solution. We are able to obtain bounds for the perturbation using the Theory of Moments.
For an $n \times n$ symmetric, positive definite matrix A and a real vector $\vec{u}$, we study a method to estimate and bound the quadratic form

$$\frac{\vec{u}' F(A) \vec{u}}{\vec{u}' \vec{u}}$$

where F is a differentiable function. This problem arises in many applications in least squares theory e.g. computing a parameter in a least squares problem with a quadratic constraint, regularization and estimating backward perturbations of linear least squares problems. We describe a method based on the theory of moments and numerical quadrature for estimating the quadratic form. A basic tool is the Lanczos algorithm which can be used for computing the recursive relationship for orthogonal polynomials. We will present some numerical results showing the efficacy of our methods and will discuss various extensions of the method.

Direct solution of saddle point problems of dimension 10^9 arising in optimization with interior point methods

Jacek Gondzio
School of Mathematics, University of Edinburgh
King's Buildings, Edinburgh, EH9 3JZ, UK
J.Gondzio@ed.ac.uk

Andreas Grothey
School of Mathematics, University of Edinburgh
King's Buildings, Edinburgh, EH9 3JZ, UK
A.Grothey@ed.ac.uk

We consider very large block-structured indefinite linear systems arising in optimization with interior point methods. In the area of optimization they are known as reduced Karush-Kuhn-Tucker systems and in the PDE community they are known as saddle point problems. We apply a direct method which exploits the block-structure of the problem in the linear algebra operations. More specifically, we use the implicit inverse representations, which allows us for a thorough exploitation of block-sparsity. We employ these techniques in the modern implementation of interior point algorithm. The program called OOPS (Object-Oriented Parallel Solver: http://www.maths.ed.ac.uk/~gondzio/parallel/solver.html) can efficiently handle very large problems and achieves scalability on a number of different computing platforms. We illustrate its performance on a collection of problems with sizes reaching 10^9 variables arising from asset liability management and portfolio

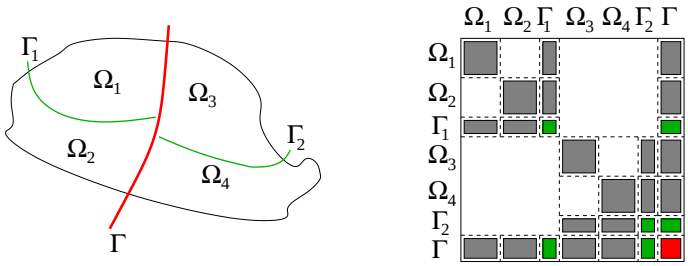
Domain decomposition based $\mathcal{H}$-matrices for FEM

Lars Grasedyck
Scientific Computing – MPI for Mathematics in the Sciences Leipzig
Inselstr. 22, D-04103 Leipzig, Germany
lgr@mis.mpg.de

Ronald Kriemann
Scientific Computing – MPI for Mathematics in the Sciences Leipzig
Inselstr. 22, D-04103 Leipzig, Germany
rok@mis.mpg.de

Sabine LeBorne
Department of Mathematics – Tennessee Technological University
Box 5054, Cookeville, TN 38505
sleborne@tntech.edu

Most direct methods for sparse linear systems perform an LU factorisation of the original matrix after some reordering of the indices in order to reduce fill-ins. One such popular reordering method is the so-called *nested dissection* which exploits the concept of separation. The main idea is to separate a (matrix) graph into three parts, two of which have no coupling between each other. The third one, referred to as an interior boundary or separator, contains couplings with (possibly both of) the other two parts. The nodes of the separator are numbered last. This process is then repeated recursively in each subgraph. An illustration of the resulting sparsity pattern is shown in the following figure for the first two decomposition steps.



A favourable property of such an ordering is that a subsequent LU factorisation maintains a major part of this sparsity structure, i.e., there occurs no fill-in in the large, off-diagonal zero matrix blocks. For a regular two-dimensional grid the computational complexity amounts to $\mathcal{O}(N^{1.5})$ for a matrix $A \in \mathbb{R}^{N \times N}$. In order to obtain a (nearly) optimal complexity, we propose to approximate all nonzero, off-diagonal blocks in $\mathcal{H}$-matrix representation and compute them using $\mathcal{H}$-matrix arithmetic. The (small) blocks along the diagonal and the corresponding LU factorisations will be stored as full matrices. We apply the new domain decomposition based $\mathcal{H}$-LU factorisation [2, 1] to a volume-conduction model of the head [4], where the system has to be solved for many right-hand sides, and compare the new method to the optimised algebraic multigrid method PEBBLES. Both approaches use the leadfield-bases acceleration from [3].

References

[1] L. Grasedyck and S. LeBorne, *$\mathcal{H}$-matrix preconditioners in convection-dominated problems*, *SIAM J. Mat. Anal.* accepted.

[2] M. Lintner, *The eigenvalue problem for the 2d Laplacian in $\mathcal{H}$-matrix arithmetic and application to the heat and wave equation*, *Computing*, (2004), pp. 293–323.

[3] C. Wolters, L. Grasedyck, and W. Hackbusch, *Efficient computation of lead field bases and influence matrix for the FEM-based EEG and MEG inverse problem. part I: Complexity considerations*, *Inverse Probems*, (2004), pp. 1099–1116.

[4] C. Wolters, M. Kuhn, A. Anwander, and S. Reitzinger, *A parallel algebraic multigrid solver for finite element method based source localization in the human brain*, *Comp. Vis. Sci.*, (2002), pp. 165–177.

Preconditioners for the discretized time-harmonic Maxwell equations in mixed form

Chen Greif
The University of British Columbia
Vancouver, B.C., Canada
greif@cs.ubc.ca

Dominik Schötzau
The University of British Columbia
Vancouver, B.C., Canada
schoetzau@math.ubc.ca

We introduce a new preconditioning technique for iteratively solving linear systems arising from finite element discretizations of the mixed formulation of the time-harmonic Maxwell equations. The preconditioners are block diagonal with positive definite blocks and are based on discrete augmentation using the scalar Laplacian. They are motivated by spectral equivalence properties of the discrete operators. Specifically, we show that augmenting the curl-curl operator by a discrete grad-div operator, weighed by the scalar Laplacian, yields almost immediate convergence when preconditioned MINRES is used. We also show that if the augmented term is replaced by the vector mass matrix we still obtain fast convergence. Similar (operator-independent) algebraic principles can be applied in general settings and give rise to preconditioners that work effectively for saddle-point linear systems whose (1,1) block is highly singular. Analytical observations and numerical results demonstrate the scalability and the convergence properties of the proposed approach.

Preconditioning of linear systems arising in the optimal control of MHD

Roland Griesse
Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences
Altenbergerstraße 69, A–4040 Linz, Austria
roland.griesse@oeaw.ac.at

Magnetohydrodynamics, or MHD, deals with the mutual interaction of electrically conducting fluids and magnetic fields. We recall the velocity–current formulation of the incompressible stationary MHD system and present an optimal control approach motivated from practical applications.
Some emphasis in this presentation will be given to the linear systems arising in the simulation and optimal control of these equations. For instance, the linearized forward problem requires the repeated solution of linear systems of the following form,

$$\begin{pmatrix} M & & & -F & \\ G[\mathbf{J}]^\top & A & B^\top & C[\mathbf{A}]^\top & \\ & B & & & \\ H[\mathbf{u}] & -C[\mathbf{A}] & & D & E^\top \\ & & & E & \end{pmatrix} \begin{pmatrix} \mathbf{A} \\ \mathbf{u} \\ p \\ \mathbf{J} \\ \phi \end{pmatrix} = r. \tag{1}$$

The variables $\mathbf{A}$, $(\mathbf{u}, p)$, and $(\mathbf{J}, \phi)$ represent the magnetic vector potential, fluid velocity and pressure, and electric current and potential, respectively. The saddle point sub-problems involving the matrices

$$\begin{pmatrix} A & B^\top \\ B & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} D & E^\top \\ E & 0 \end{pmatrix}$$

originate in the linearization of the Navier-Stokes and the current-potential systems, respectively.
We shall address several preconditioners for (1), in particular in the context of optimal control, where (1) arises as a sub-problem in the solution of Karush-Kuhn-Tucker (KKT) systems.

Characterizing the fill of sparse Gaussian elimination on bipartite graphs

Laura Grigori
INRIA
Campus de Beaulieu, F-35042 Rennes, France
Laura.Grigori@inria.fr

In this talk we discuss the structure prediction problem of the Gaussian elimination when used to solve the linear system $Ax = b$, where A is a $n \times n$ nonsingular and unsymmetric matrix and b is a n-vector. When A is sparse, during the Gaussian elimination, or equivalently the $A = LU$ factorization, some of the elements that are zero in the original matrix A can become nonzero in the factors L and U.
A graph-theoretic elimination model was proposed by Rose and Tarjan [2] to describe the creation of nonzero elements, called fill-ins, using the directed graph of matrix A. One of the main results introduced in [2] is the Fill Path Lemma. It describes the fill-in elements of the factors L and U in terms of paths in the directed graph of A, represented in the so called filled graph of A. We discuss this lemma further in the paper.
The Fill Path Lemma shows in particular that fill-ins can be computed without referring to the numerical values of the matrix. Hence the Gaussian elimination can be performed by first determining the nonzero structure of the factors and then carrying out the numerical computation of the factors itself. Knowledge of this structure is used not only to avoid operations on zeros, but also to allocate memory or set up data structures. All the solvers implementing the Gaussian elimination or the incomplete LU factorization compute fill-ins using algorithms derived from the Fill Path Lemma and based on directed graphs.
It is well known that the filled graph of A represents a symbolic bound of the nonzero structure of the factors L and U. That is, it assumes that during Gaussian elimination the result of adding or subtracting two nonzeros is never zero and ignores possible numeric cancellations that occur during the factorization. The goal of this work is to model the creation of fill-in elements using bipartite graphs and develop related algorithmic aspects of the elimination process.
New results describe necessary and sufficient conditions, in terms of paths in the bipartite graph of A, for a fill-in element of L and U to appear during the factorization. These results allow to identify submatrices of A that are structurally singular, that is, singular due to the arrangements of their nonzeros, and independently of their numerical values. These submatrices are related to numeric cancellations occurring in the LU factorization, that are missed by the model based on directed graphs.
The graph theoretic model based on bipartite graphs allows an exact estimation of the nonzero structure of the factors L and U. It assumes that the nonzero values in A are algebraically independent from each other; in other words, it assumes that any computed zero is due to combinatorial properties of the nonzero structure rather than to coincidence in choice of values.
The theoretical results are used to derive an algorithm for computing the nonzero structure of L and U using bipartite graphs. Its complexity is bounded by $O(nm)$ where n is the order and m is the number of nonzeros of matrix A. This is equivalent to the complexity of FILL2 [2], one of the first algorithms proposed for computing fill-in and based on the Fill Path Lemma.

References

[1] L. Grigori, *Characterizing the fill of sparse gaussian elimination on bipartite graphs*, Submitted to SIAM J. Matrix Anal. Appl., (2005).

[2] D. J. Rose and R. E. Tarjan, *Algorithmic aspects of vertex elimination on directed graphs*, SIAM J. Appl. Math., 34 (1978), pp. 176–197.

Extracting spectral information from large "singularly perturbed" matrices

Luka Grubišić
Institut für Reine und Angewandte Mathematik, Rheinisch-Westfälische Technische Hochschule in Aachen
Templergraben 55, D-52062 Aachen, Germany
luka.grubisic@iram.rwth-aachen.de

Consider a class of matrices which are produced as hierarchical discretisations of a singularly perturbed elliptic (self adjoint) operator.
The question we pose is: How to measure the quality of an approximate eigenvector ψ, $\|\psi\| = 1$?
For the sake of being definite assume that we have constructed a vector ψ which approximates the eigenvector v, $\|v\| = 1$, $Hv = \lambda v$. Then, typically, one is satisfied with ψ such that $\sin\angle(\psi, v)$ be "tiny". One expects that in such a situation ψ can be used to extract spectral information on the matrix H.
Unfortunately, small $\sin\angle(\psi, v)$ does not always imply that quality information on λ can be extracted from ψ. Surely an approximate eigenvector which yields insufficient information on the accompanying eigenvalue cannot be considered as a satisfactory solution to a spectral approximation problem.
The so-called "standard" or "relative" $\sin\Theta$ theory (Davis–Kahan, Ipsen, Li, Mathias–Veselić...) gives only a partial answer to this problem. Indeed, in a "singular" situation the relative residual is typically large (as it should be for a poor approximation ψ) but the estimated quantity $\sin\Theta(\psi, v)$ can be (and in the presented example it will be) very tiny. So we are in a situation in which we have to rely on a possibly very un-sharp inequality to obtain a correct hint on the "accuracy" of ψ.
We reconsider the approximation problem in the geometry of the "energy" norm $\|\cdot\|_H$. The obtained estimates are an energy norm variant of the Mathias–Veselić eigenvector inequalities. The almost sharpness of the new estimates can be traced back to the scaled version of the celebrated Wilkinson's Schur Complement Trick. Issues of cluster stability/robustnes as well as issues of invariant subspace approximations will also be tackled.
The class of matrices we consider inherits a lot of structure from the differential operators which lay in the background. We discuss how to design a mesh refinement procedure which produces a hierarchical sequence of matrices whose eigenvectors will be progressively better resolved by a triangulation. As a measure of the agreement between a triangulation and an approximate eigenvector we use the osc-function of Morin, Nochetto and Siebert. We argue that for our class of "discretisation" matrices the (realistic) residuum should reflect both the algebraic and the analytic properties of the approximate eigenvector ψ (equally for a more general invariant subspace setting).
Numerical examples which illustrate the theory will also be presented. Typical matrices of interest have their origin (or better to say their singularities have) in problems of the Large Coupling Limit (in both Quantum and Classical Mechanics) or can be seen as Laplace matrices of weighted graphs where the weights of some edges tend to explode.
Acknowledgement: Some of the results are a product of a joint work with Prof. Krešimir Veselić, Hagen and Prof. Zlatko Drmač, Zagreb.

An iterative SVD-Krylov based method for model reduction

Serkan Gugercin
Virginia Polytechnic Institute and State University
Department of Mathematics, Virginia Tech., 460 McBryde Hall, Blacksburg, VA, 24061-0123
gugercin@math.vt.edu

We propose a model reduction algorithm which combines the two model reduction frameworks, namely the SVD and Krylov-based frameworks. The proposed technique is a two-sided projection method where one side carries the SVD (Gramian) information and the other side the Krylov information. While the SVD-side depends on the observability gramian, the Krylov-side is obtained via iterative rational Krylov steps. The reduced model is asymptotically stable and matches the moments of the original system at the mirror images of the reduced system poles; hence it is the best $\mathcal{H}_2$ approximation among all reduced models having the same reduced system poles. We present both continuous and discrete-time formulations and discuss large-scale implementation issues in detail. Several numerical results and comparison to existing methods will be presented to illustrate the effectiveness of the proposed method.

Iterative solution of a nonsymmetric algebraic Riccati equation

Chun-Hua Guo
Department of Mathematics and Statistics, University of Regina
Regina, SK S4S 0A2, Canada
chguo@math.uregina.ca

Nicholas J. Higham
School of Mathematics, The University of Manchester
Sackville Street, Manchester, M60 1QD, UK
higham@ma.man.ac.uk

We study the nonsymmetric algebraic Riccati equation whose four coefficient matrices are the blocks of a nonsingular M-matrix or an irreducible singular M-matrix M. The solution of practical interest is the minimal nonnegative solution. We show that Newton's method with zero initial guess can be used to find this solution without any further assumptions. We also present a qualitative perturbation analysis for the minimal solution, which is instructive in designing algorithms for finding more accurate approximations. For the most practically important case, in which M is an irreducible singular M-matrix with zero row sums, the minimal solution is either stochastic or substochastic and the Riccati equation can be tranformed into a unilateral matrix equation by a procedure of Ramaswami. The minimal solution of the Riccati equation can then be found by computing the minimal nonnegative solution of the unilateral equation using the Latouche–Ramaswami algorithm. We show that the Latouche–Ramawami algorithm, combined with a shift technique suggested by He, Mini, and Rhee, is breakdown-free in all cases and is able to find the minimal solution more efficiently and more accurately than the algorithm without a shift. Our approach is to find a proper stochastic solution using the shift technique even if it is not the minimal solution. We show how we can easily recover the minimal solution when it is not the computed stochastic solution.

From qd to LR and QR

Martin H. Gutknecht
Seminar for Applied Mathematics, ETH Zurich
CH-8092 Zurich, Switzerland
mhg@math.ethz.ch

Beresford N. Parlett
Department of Mathematics, University of California
Berkeley, CA 94720-3840, USA
parlett@Math.Berkeley.edu

It is well known that Francis' QR algorithm for computing eigenvalues emerged from Rutishauser's LR algorithm, but it is less obvious how Rutishauser discovered the matrix interpretation of his qd algorithm, which is just the LR algorithm for tridiagonal matrices, and why he seems to have been much less excited on LR than on qd. Moreover, it is far from obvious how he discovered qd in the first place. We try to shed some light on these questions.

The technique of hierarchical matrices

Wolfgang Hackbusch
Max-Planck-Institut *Mathematik in den Naturwissenschaften*
Inselstr. 22, D-04103 Leipzig, Germany
wh@mis.mpg.de

The discretisation of partial differential equations (PDEs) leads to large systems of equations. In particular, the boundary element method (BEM) produces fully populated matrices. Several methods try to reduce the costs for storage and matrix-vector multiplication from $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log^q n)$. The technique of hierarchical matrices presented here supports *all* matrix operations, i.e., addition, multiplication, inversion of matrices and LU decompositions up to certain approximation errors. This is of interest also for the sparse matrices from FEM, since the dense inverse or certain Schur complements can be approximated. Concerning the applications, we mention five topics.

FEM preconditioning: Linear equations with sparse FEM matrices A are usually solved iteratively, provided a good preconditioner is available. The technique of hierarchical matrices allows to approximate the LU-factors (if they exist) which lead to a perfect black-box preconditioner. The same technique applies to BEM problems.

Domain decomposition: Ideas from DDM can be combined with the hierarchical matrices. Eliminating the interior unknowns in a domain decomposition method with non-overlapping subdomains, one obtains a dense matrix (Schur complement) for the nodal points on the skeleton. The hierarchical matrix technique allows the elimination as well as the treatment of the skeleton matrix.

Matrix equations: The Lyapunov and Riccati equation arise in control theory and define a system of n^2 equations for the n^2 unknown entries of X. Therefore the best possible solve seems to need a work of $\mathcal{O}(n^2)$. If the coefficient matrix A arises from an elliptic operator (as in control problems with a state governed by an elliptic boundary value problem), it turns out that the solution X can be well approximated by a hierarchical matrix. The costs add up to $\mathcal{O}(nk^2 \log^3 n)$ even in the case of the nonlinear Riccati equation.

Matrix functions: The matrix exponential function $\exp(-tA)$ is of general interest. We are able to compute $\exp(-tA)$ with accuracy ε with a cost of order $\mathcal{O}(n \log^p \frac{1}{\varepsilon} \log^q n)$. Similarly, other matrix functions can be computed, in particular, the sign-function $\mathrm{sign}(A)$ is a very interesting function.

Problems in high spatial dimensions: Related techniques can be applied to problems in high spatial dimensions when Kronecker products of matrices can be used. Examples are mentioned where matrices of size $N = 1024^{2048} \approx 1.2 \times 10^{6165}$ are treated.

Concerning publications about hierarchical matrices compare http://www.mis.mpg.de/scicomp/ or http://www.hmatrix.org/index.html.

Stability and self excited vibrations of axially moving rods with frictional contact

Peter Hagedorn
Fachgebiet Dynamik – Fachbereich Maschinenbau, Technische Universität Darmstadt
Hochschulstraße 1, D-64289 Darmstadt, Germany
hagedorn@dyn.tu-darmstadt.de

Gottfried Spelsberg-Korspeter, Daniel Hochlenert
Fachgebiet Dynamik – Fachbereich Maschinenbau, Technische Universität Darmstadt
Hochschulstraße 1, D-64289 Darmstadt, Germany
{speko,hochlenert}@dyn.tu-darmstadt.de

Axially moving and vibrating elastic rods with bending stiffness (beams) occur frequently in technical applications. Examples are conveyor belts, band saws etc. In most cases the structural vibrations are unwanted and therefore possible excitation mechanisms have to be identified and suppressed by appropriate design.
This paper is devoted to the formulation and identification of a possible excitation mechanism for self excited vibrations in moving continua under frictional contact. The underlying model is a travelling beam sliding through two idealized brake pads. Using a Ritz discretization approach it is shown that self excited vibrations due to Coulomb friction occur resulting from an instability of the trivial solution of the linearized discretized equations of motion. Furthermore it is shown that the consistent formulation of the contact between the surface of the beam and the brake pads is essential for the stability behavior. The insights gained from the travelling beam also serve as an explanation for brake squeal and the methods used can also be applied to a rotating Kirchhoff plate.

Eigenproblems in chemistry: from multiresolution and (near) linear scaling algorithms for SCF to multi billion CI vectors

Robert J. Harrison
Computational Chemical Sciences
Oak Ridge National Laboratory
P.O. Box 2008 MS6378, Oak Ridge, TN 37831-6367, USA
harrisonrj@ornl.gov

Comparison of conjugate gradient method for nonsymmetric matrices

Hidehiko Hasegawa
Graduate School of Library, Information and Media studies, University of Tsukuba
hasegawa@slis.tsukuba.ac.jp

Tomohiro Sogabe
Graduate School of Engineering, Nagoya University
sogabe@na.cse.nagoya-u.ac.jp

Takeshi Ogita, Tamito Kajiyama
Japan Science and Technology Agency
ogita@waseda.jp, kajiyama@is.s.u-tokyo.ac.jp

In this talk, we show some comparisons of Conjugate Gradient (CG) methods for nonsymmetric matrices. Even though the CG method is for symmetric matrices, it can be applied to nonsymmetric matrices with some modifications. Since there are few reports on the comparison of their convergence behavior on the modern computers, we tested the following two methods:

- CGNR: Conjugate Gradient method for the normal equation $A'Ax = A'b$[2]
- Conjugate Gradient method to solve $Ax = b$ and $A'y = c$ simultaneously[1]

The condition number of $A'A$ is a square of that of original matrix A, which may lead to poor convergence for the CGNR method. However, there are some test problems where it shows smoother and even faster convergence behavior than the BiCG method to $Ax = b$. Second, to solve $Ax = b$ and $A'y = c$ simultaneously, the dimension of a coefficient matrix becomes twice of that of original matrix A and A'. In this case, there is a difficulty to give a right-hand side of additional equation $A'y = c$, but the convergence is not as bad as expected.
We will also show results of preconditioned methods and ones with the use of quadruple precision operations for accurate computation. To perform these numerical experiments effectively, we used a tool called SILC[3].

References

[1] E. H. Ayachour, *Expanded systems and the ILU preconditioner for solving non-Hermitian linear systems*, Linear Algebra and its Applications, 293 (1999), pp. 243 – 256.

[2] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 409 – 436.

[3] T. Kajiyama and *et. al*, *SILC: a flexible and environment independent interface to matrix computation libraries*, In Proceedings of the Sixth International Conference on Parallel Processing and Applied Mathematics (PPAM2005), to appear.

GMRES methods for least squares problems

Ken Hayami
National Institute of Informatics
2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan
hayami@nii.ac.jp

Tokushi Ito
Business Design Laboratory Co., Ltd.
1-7-1, Yuraku-cho, Chiyoda-ku, Chiyoda-ku, Tokyo 100-0006, Japan
ito@kta.biglobe.ne.jp

Consider the least squares problem : $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|_2$, where $A \in \mathbf{R}^{m\times n}$, and $m \geq n$ or $m < n$, and A may be rank-deficient.
Previous iterative methods for solving this problem were mainly based on solving the normal equation: $A^{\mathrm{T}}A\boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{b}$ using the preconditioned conjugate gradient (CG) method[1]: the CGLS approach. However, this approach may be slow to converge for highly ill-conditioned problems.
Zhang and Oyanagi [6] proposed the CR-LS(k) method, which uses an $n \times m$ matrix B and applies the Orthomin(k) method[5] directly to $\min_{\boldsymbol{z}\in\mathbf{R}^m} \|\boldsymbol{b} - AB\boldsymbol{z}\|_2$.
In this talk, we follow this idea and introduce an appropriate $B \in \mathbf{R}^{n\times m}$ and apply the more robust GMRES (Generalized Minimal Residual) method [4] to $\min_{\boldsymbol{z}\in\mathbf{R}^m} \|\boldsymbol{b} - AB\boldsymbol{z}\|_2$ as well as $\min_{\boldsymbol{x}\in\mathbf{R}^n} \|B\boldsymbol{b} - BA\boldsymbol{x}\|_2$ [3, 2].
We give a sufficient condition concerning B for the proposed methods to give a least squares solution of the original problem without breakdown for arbitrary $\boldsymbol{b}$, for over-determined, under-determined and possibly rank-deficient problems. We also present some convergence analysis.
Finally, numerical experiments for full rank problems with $m \geq n$ show that the BA approach with $B = (\mathrm{diag}(A^\top A))^{-1}A^\top$, converges faster than previous methods based on CGLS for highly ill-conditioned problems. For the case $m < n$, the AB approach with $B = A^\top$ gave similar results.

References

[1] Björck, A., *Numerical Methods for Least Squares Problems*, SIAM, 1996.

[2] Hayami, K. and Ito, T., The solution of least squares problems using GMRES methods, *Proceedings of the Institute of Statistical Mathematics*, Vol. 53, No. 2, 2005, (to appear, in Japanese).

[3] Ito, T. and Hayami, K., Preconditioned GMRES methods for least squares problems, *NII Technical Reports*, National Institute of Informatics, Tokyo, NII-2004-006E, pp. 1–29, May, 2004.

[4] Saad, Y. and Schultz, M. H., GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM J. Sci. Stat. Comput.*, Vol. 7, No. 3 (1986), pp. 856–869.

[5] Vinsome, P.K.W., ORTHOMIN – an iterative method for solving sparse sets of simultaneous linear equations, *Proc. 4th Sym. on Reservoir Simulations*, Soc. of Petroleum Engineers of AIME, pp. 149–159, 1976.

[6] Zhang, S.-L. and Oyanagi, Y., Orthomin(k) method for linear least squares problem, *Journal of Information Processing*, Vol. 14, No. 2, pp. 121–125, 1991.

PDE based techniques for image registration

Stefan Henn
Mathematisches Institut
Heinrich-Heine Universität Düsseldorf, Universitätsstr.1,
40225 Düsseldorf, Germany
henn@am.uni-duesseldorf.de

Kristian Witsch
Mathematisches Institut
Heinrich-Heine Universität Düsseldorf, Universitätsstr.1,
40225 Düsseldorf, Germany
witsch@am.uni-duesseldorf.de

In recent years there has been an increasing number of papers devoted to the application of partial differential equations (PDEs) in image processing, such as image restoration, segmentation, optical flow computation and image registration.
We consider the image registration problem, i.e. to find a reasonable displacement field, such that a transformed template image becomes similar to a so–called reference image; see [7]. The behavior of image registration problems is often governed by an energy functional. Typically, these energy functionals consist of two terms: a data term, which measures the disparity between the template and the reference image (e.g. least squares, mutual information, etc.) and a regularizing term (e.g. classical Tikhonov, linear elastic, Poisson, Biharmonic, total variation, etc.) that encourages global or local smoothness of the displacement field; e.g. see [1, 2, 4, 3, 4, 5, 6].
In this talk we present different PDE based techniques for the image registration problem, such as the classical multigrid method, multilevel methods, scale space approaches for the underlying images as well as scale space approaches for the displacement field.

References

[1] C. Frohn-Schauf, S. Henn, L. Hömke, and K. Witsch, *Total variation based image registration*, PDE-Based Image Processing and Related Inverse Problems, Lecture Notes in Computer Science, Springer-Verlag, (2006).

[2] V. Grimm, S. Henn, and K. Witsch, *A unified higher order pde based image registration approach*, Journal of Numerical linear algebra with applications (NLAA), to appear.

[3] S. Henn, *A Levenberg-Marquardt scheme for nonlinear image registration*, BIT Numerical Mathematics, 43(4) (2003), pp. 743–759.

[4] ———, *A multigrid method for a fourth-order diffusion equation with application to image processing*, SIAM J. Sci. Comput. (SISC), 27(3) (2006), pp. 831–849.

[5] S. Henn and K. Witsch, *Iterative multigrid regularization techniques for image matching*, SIAM J. Sci. Comput. (SISC), 23(4) (2001), pp. 1077–1093.

[6] ———, *Image registration based on multiscale energy information*, SIAM Journal on Multiscale Modeling and Simulation (MMS), 4(2) (2005), pp. 584–609.

[7] J. Modersitzki, *Numerical Methods for Image Registration*, Oxford Science Publications, 2003.

Iterations for matrix roots: convergence and stability

Nicholas J. Higham
The University of Manchester
School of Mathematics, The University of Manchester, Sackville Street, Manchester, M60 1QD, UK
higham@ma.man.ac.uk

A large literature exists on iterations for computing matrix functions such as the sign function, matrix roots, and the unitary polar factor. Two important issues are convergence and numerical stability. For the standard matrix square root iterations, most of which are related to Newton's method, convergence is relatively easy to analyze and stability is well understood. For matrix pth roots with $p > 2$ Newton methods were until recently little used, for two reasons: their convergence in the presence of complex eigenvalues was not well understood and the iterations have been found to have poor numerical stability. The subtlety of the question of convergence is clear from the scalar case, since the starting values for which Newton's method for $z^p - 1 = 0$ converges to some pth root of unity form fractal Julia sets in the complex plane for $p > 2$. We show that if the eigenvalues of $A \in \mathbb{C}^{n\times n}$ lie in a certain convex set in the complex plane then the inverse Newton iteration converges to $A^{-1/p}$, and we develop a hybrid Schur–Newton algorithm that can compute pth roots of general matrices. We give a general definition of stability in terms of a Frechét derivative and show that it can distinguish between several variants of the Newton iteration for the matrix pth root. Finally, an application to roots of transition matrices from Markov models is described.

Computing $f(A)b$ for matrix functions f

Nicholas J. Higham
The University of Manchester
School of Mathematics, The University of Manchester, Sackville Street, Manchester, M60 1QD, UK
higham@ma.man.ac.uk

For matrix functions f we investigate how to compute a matrix-vector product $f(A)b$ without explicitly computing $f(A)$. A general method is described that applies quadrature to the matrix version of the Cauchy integral theorem. Methods specific to the logarithm, based on quadrature, and fractional matrix powers, based on solution of an ordinary differential equation initial value problem, are also presented.

Golub–Kahan bidiagonalization and stopping criteria in solving ill-posed problems

Iveta Hnětynková
Faculty of Mathematics and Physics, Department of Numerical Mathematics, Charles University
Sokolovská 83, 186 75 Prague 8, Czech Republic
hnetynkova@cs.cas.cz

Martin Plešinger
Faculty of Mechatronics, Department of Modelling of Processes, Technical University of Liberec
Hálkova 6, 461 17 Liberec 1, Czech Republic
martin.plesinger@tul.cz

Zdeněk Strakoš
Institute of Computer Science, Academy of Sciences of Czech Republic
Pod Vodárenskou věží 2, 182 07 Prague 8, Czech Republic
strakos@cs.cas.cz

Golub-Kahan bidiagonalization has been used for solving of large ill-posed problems for years. The outer bidiagonalization process is combined with an inner regularization applied to the problem with a lower bidiagonal $(k+1) \times k$ matrix. Stopping criteria are typically based on the estimation of the L-curve [5] using the L-ribbon [3], the discrepancy principle and the generalized cross validation [1, 2], see also [4, Chap. 6].
Recently it has been proved that the Golub-Kahan bidiagonalization leads to a fundamental decomposition of data, which reveals the so called core problem [6]. Application to ill-posed problems has been studied by D. Sima and S. Van Huffel [7]. In this contibution we investigate a possibility of using an information from the bidiagonalization process for constructing an effective stopping criteria.

This work has been supported by the National Program of Research "Information Society" under project 1ET400300415. The research of Iveta Hnětynková has been partly supported by the Grant GAUK 6/2005/R.

References

[1] Å. Björck, *A bidiagonalization algorithm for solving large and sparse ill-posed systems of linear equations*, BIT, 28 (1988), pp. 659–670.

[2] Å. Björck, E. Grimme, and P. Van Dooren, *An implicit shift bidiagonalization algorithm for ill-posed systems*, BIT, 34 (1994), pp. 510–534.

[3] D. Calvetti, G. H. Golub, and L. Reichel, *Estimation of the L-curve via Lanczos bidiagonalization*, BIT, 39 (1999), pp. 603–619.

[4] P. C. Hansen, *Rank-deficient and discrete ill-posed problems*, SIAM Monographs on Mathematical Modeling and Computation, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998. Numerical aspects of linear inversion.

[5] P. C. Hansen and D. P. O'Leary, *The use of the L-curve in the regularization of discrete ill-posed problems*, SIAM J. Sci. Comput., 14 (1993), pp. 1487–1503.

[6] C. C. Paige and Z. Strakoš, *Core problems in linear algebraic systems*, SIAM J. Matrix Anal. Appl., (to appear in 2005).

[7] D. M. Sima and S. Van Huffel, *Using core formulations for ill-posed linear systems*, PAMM, 5 (2005), pp. 795–796.

Homogeneous and SPAM Jacobi–Davidson

Michiel E. Hochstenbach
Department of Mathematics, Case Western Reserve University
10900 Euclid Avenue, Cleveland OH 44106-7058, USA
hochsten@case.edu

Gerard L. G. Sleijpen
Mathematical Institute, Utrecht University
Budapestlaan 6, NL-3508 TA Utrecht, The Netherlands
sleijpen@math.uu.nl

We address two challenges for the Jacobi–Davidson method:

- the method generally requires many matrix-vector products in the inner iteration; this is unfavorable for "less sparse" matrices;
- for generalized eigenvalue problems, we would like to cope with finite and infinite eigenvalues in one procedure.

Inspired by this, we give two novel Jacobi–Davidson variants:

- Subspace-projected approximate matrix (SPAM) Jacobi–Davidson, which may be useful when we have an approximation to the matrix of which the action is (much) less expensive;
- Homogeneous Jacobi–Davidson, which may be useful to compute finite and infinite eigenvalues in one consistent procedure.

Subspace extraction for matrix functions

Michiel E. Hochstenbach
Department of Mathematics, Case Western Reserve University
10900 Euclid Avenue, Cleveland OH 44106-7058, USA
hochsten@case.edu

Marlis Hochbruck
Mathematisches Institut, Lehrstuhl für Angewandte Mathematik, Heinrich-Heine-Universität Düsseldorf
Universitätsstraße 1, D-40225 Düsseldorf, Germany
marlis@am.uni-duesseldorf.de

We discuss various approaches to approximate $x = f(A)\,b$, for a large (sparse) matrix A, by using a search space method. The search space may be a Krylov space, but this is not necessary. We examine one-sided and two-sided harmonic extraction techniques from a search space, where we use ideas coming from the numerical solution of eigenvalue problems. The methods are derived in different ways. The harmonic techniques may result in a smoother convergence for functions with a (near) discontinuity or pole.
If time allows, we will also discuss a preconditioning technique for the matrix exponential. This may be useful when we have an approximation to the matrix $M \approx A$ for which the action with $\exp(M)$ is (much) less expensive to compute than that with $\exp(A)$.
This talk is partly based on [1].

References

[1] M. Hochbruck and M. E. Hochstenbach, *Subspace extraction for matrix functions*, Preprint, Dept. Math., Case Western Reserve University, September 2005. Submitted.

Iterative methods for linear inverse problems with random noise

Thorsten Hohage
Institut für Numerische und Angewandte Mathematik, Universität Göttingen
Lotzestr. 16-18, D-37083 Göttingen, Germany
hohage@math.uni-goettingen.de

In this talk [2] we study the solution of linear ill-posed operator equations with random noise:

$$Y = Kf + \text{“random noise''}$$

Here f is an unknown vector in a Hilbert space, and K is bounded (but not necessarily compact) linear operator with an unbounded Moore-Penrose inverse. We will treat several commonly used noise models including the white noise model, inverse regression with a finite number of measurement points, and errors-in-variables in a unified framework.
During the past the convergence analysis of linear statistical inverse problems has mainly focused on spectral cut-off (or truncated singular value decomposition) and Tihonov regularization (see [3, 5, 6, 7] among others). It has been shown in many cases that spectral cut-off estimators achieve the optimal rates (minimax rates) (see [4, 6, 8]). However, numerical efficiency of estimators has not been a major issue in the statistical literature so far. Therefore, we will focus on the analysis of matrix-free methods such as Landweber iteration or ν-methods which are natural and necessary for the solution of large-scale inverse problems as they occur e.g. in parameter identification problems for partial differential equations. Whereas the estimation of the bias (or approximation error) of iterative methods is known from deterministic inverse problems theory, the variance term (or propagated data noise error) must be treated differently to obtain optimal rates. Unlike for deterministic inverse problems, the bound on the variance term depends also on spectral properties of the operator. We suggest a condition on the operator which allows an optimal estimation of the variance term for all commonly used linear regularization methods and verify this assumption for a number of interesting inverse problems. Our examples include the backwards heat equation, satellite gradiometry, L^2-boosting, errors-in-variables with not necessarily independent random variables, and operators in Hilbert scales. Finally, we touch upon the topic of a completely data-driven choice of the stopping index (see [1, 3, 6]).

References

[1] F. Bauer and T. Hohage, *A Lepskij-type stopping rule for regularized Newton methods*, Inverse Problems, 21 (2005), pp. 1975–1991.

[2] N. Bissantz, T. Hohage, A. Munk, and F. Ruymgaart, *Convergence rates of general regularization methods for statistical inverse problems and applications*. submitted.

[3] L. Cavalier, G. K. Golubev, D. Picard, and A. B. Tsybakov, *Oracle inequalities for inverse problems*, Ann. Stat., 30 (2002), pp. 843–874.

[4] I. M. Johnstone and B. W. Silverman, *Speed of estimation in positron emission tomography and related inverse problems*, Ann. Stat., 18 (1990), pp. 251–280.

[5] B. A. Mair and F. Ruymgaart, *Statistical inverse estimation in Hilbert scales*, SIAM J. Appl. Math., 56 (1996), pp. 1424–1444.

[6] P. Mathé and S. Pereverzev, *Regularization of some linear ill-posed problems with discretized random noisy data*, Math. Comp., (to appear).

[7] D. W. Nychka and D. Cox, *Convergence rates for regularized solutions of integral equations from discrete noisy data*, Ann. Stat., 17 (1989), pp. 556–572.

[8] M. S. Pinsker, *Optimal filtration of square-integrable signals in gaussian noise*, Probl. Inf. Transm., 16 (1980), pp. 52–68.

Regularization of inverse problems by exponential integrators

Michael Hönig
Mathematisches Institut, Lehrstuhl für Angewandte Mathematik, Heinrich-Heine-Universität Düsseldorf
Universitätsstraße 1, D-40225 Düsseldorf, Germany
hoenig@am.uni-duesseldorf.de

Marlis Hochbruck
Mathematisches Institut, Lehrstuhl für Angewandte Mathematik, Heinrich-Heine-Universität Düsseldorf
Universitätsstraße 1, D-40225 Düsseldorf, Germany
marlis@am.uni-duesseldorf.de

In this talk we discuss the regularization of inverse problems as an application of the evaluation of matrix functions. We study the asymptotic regularization (Showalter's method) of a linear inverse problem $A f = g$. The key idea is that the solution $u(t)$ of the evolution equation $u'(t) = A^*(g - Au(t))$, $u(0) = 0$ yields a stable approximation to the solution of the inverse problem for *large* t. We propose to apply exponential integrators for solving the evolution equation and prove that this leads to regularization schemes of optimal order. The main computational work for these integrators is spent on the approximation of a product of a matrix function and vector using Krylov subspace methods. We will discuss some practical aspects such as suitable stopping criteria. Furthermore, we present numerical experiments which show the competitiveness of these regularization schemes.

References

[1] M. Hochbruck and C. Lubich, *On Krylov subspace approximations to the matrix exponential operator*, SIAM Journal on Numerical Analysis, 34 (1997), pp. 1911–1925.

[2] A. Rieder, *Runge-Kutta integrators yield optimal regularization schemes*, Inverse Problems, 21 (2005), pp. 453–471.

Fourier analysis for twogrid methods and optimal smoothing

Thomas Huckle
Fachbereich Informatik, Technische Universität München
Boltzmann-Str. 3, D-85748 Garching, Germany
huckle@in.tum.de

The convergence of twogrid methods for the Laplace operator can be described by the equation $e_h^{(i+1)} = (I_h - M_2 L_h)^{\nu_2}(I_h - P_H^h A_H^{-1} R_h^H L_h)(I_h - M_1 L_h)^{\nu_1} e_h^{(i)} = S_h^{\nu_2} K_h^H S_h^{\nu_1} e_h^{(i)} = M_h^H e_h^{(i)}$. Here, M_j denote the post-, resp. presmoother, applied ν_j-times, P_H^h and R_h^H denote the prolongation and restriction operator, and A_h, resp. A_H, the given linear system on different grids. To find optimal smoothers, M_j have to be chosen such that the iteration matrix M_h^H has minimum norm. In standard Multigrid analysis the Fourier analysis often is applied on the matrices K_h^H and S_h separately to derive convergence criteria. Here, we try to analyse the convergence based on the actual iteration matrix itself.
It is well known that for the 1D Laplacian the Red Black Gauss Seidel smoother gives $M_h^H = 0$ and therefore the Multigrid solver can be seen as a direct solver in this case [1]. This is caused by the special property that the coarse grid correction matrix K_h^H and the matrix S_h remove the error in two orthogonal subspaces such that the application of both matrices leads to the exact solution. Unfortunately this nice property is lost in higher dimensions.
For the Laplacian with constant coefficients on a rectangular domain the resulting matrices are multilevel Toeplitz matrices and their behaviour can be described by generating functions or stencils [2]. The prolongation operator leads to a distinction between odd and even rows/columns and gives multilevel Block Toeplitz matrices in K_h^H. Therefore, we have to write the generating functions in block form, too. In the following we consider only smoothers that can be written in terms of multilevel Block Toeplitz matrices and take into account the odd/even blocking introduced by the prolongation/restriction. In the 1D case the generating functions are 2×2 matrix functions, in the 2D case they are 4×4 matrix functions.
By this approach we can analyse the iteration matrix M_h^H in terms of the generating matrix functions. We will define new versions of Red Black smoothers based on different choices for the generating block functions, and compare the norm of the related iteration functions.

References

[1] W. Briggs, V. E. Henson, and S. McCormick, *A Multigrid Tutorial*, SIAM, Philadelphia, 2000.

[2] R. Wienands and W. Joppich, *Practical Fourier Analysis for Multigrid Methods*, Chapman & Hall, Boca Raton, Florida, 2005.

Numerical schemes for three dimensional irregular shape quantum dots over curvilinear coordinate systems

Tsung-Min Hwang
Department of Mathematics, National Taiwan Normal University
88, Sec. 4, Ting Chou Road, Taipei, 116, Taiwan
min@math.ntnu.edu.tw

Wei-Cheng Wang
Department of Mathematics, National Tsing Hua University
101, Section 2 Kuang Fu Road, Hsinchu, 300, Taiwan
wangwc@math.nthu.edu.tw

Weichung Wang
Department of Applied Mathematics, National University of Kaohsiung
No.700, Kaohsiung University Rd., Nan Tzu Dist., Kaohsiung, 811, Taiwan
wwang@nuk.edu.tw

We present efficient and stable numerical discretized schemes to simulate three-dimensional quantum dot heterostructure with irregular shape. The main goal is to compute all the bound state energies (eigenvalues) and associated wave functions (eigenvectors). A curvilinear coordinate system that fits the target quantum dot shape is first determined. Three finite difference discretizations of the Schrödinger equation with uniform meshes are then developed on the original and the skewed curvilinear coordinate system. The resulting large-scale generalized eigenvalue systems are solved by a modified Jacobi–Davidson method. Intensive numerical experiments show that the scheme using both grid points on the original and skewed curvilinear coordinate system can converge to the eigenpairs quickly and stably with second order accuracy.

Roots of quaternions

Drahoslava Janovská
Institute of Chemical Technology, Prague
Technická 5, 166 28 Prague, Czech Republic
janovskd@vscht.cz

Gerhard Opfer
University of Hamburg
Bundesstrasse 55, D-20146 Hamburg, Germany
opfer@math.uni-hamburg.de

Similarly as for real or complex numbers, we define roots of a quaternion $a \in \mathbb{H}$ as the solutions of the equation

$$g(x) := x^n - a = 0, \quad a \in \mathbb{H}\backslash\mathbb{R}, \quad n \in \mathbb{N}, \quad n \geq 2. \tag{1}$$

At first, we summarize some basic facts about quaternions. There are more details in our previous paper [1]. Results concerning matrices with quaternion elements are surveyed by Zhang in [3]. Householder's transformation matrices for quaternion-valued matrices are defined and their main properties are listed in [2].
We start with some information on explicit formulas for roots of quaternions. Then we adjust the common Newton formula for the $n-$th root of a real (positive) or complex number to the case of quaternions. Because of the non commutativity of the multiplication we obtain two slightly different formulas. Under a simple condition, both formulas produce the same sequence. We also apply the Gâteaux derivative and the Jacobi matrix of the partial derivatives to formula (1) and show that under the same condition the same formulas come out which proves that the convergence is locally quadratic. The Gâteaux derivative gives also rise to the damped Newton form which turns up to be very successful and superior to the ordinary Newton technique.

References

[1] D. Janovská and G. Opfer, *Givens' transformation applied to quaternion valued vectors*, BIT Numerical Mathematics, 43 (2003), pp. 991–1002.

[2] ———, *Givens' and householder' transformations applied to quaternion-valued matrices*, Proceedings of SANM'03, Hejnice,University of West Bohemia in Pilsen, (2004), pp. 17–24.

[3] F. Zhang, *Quaternions and matrices of quaternions*, Linear Algebra Appl., 251 (1997), pp. 21–57.

The analytic SVD: on the non-generic points on the path

Vladimír Janovský
Faculty of Mathematics and Physics, Charles University, Prague
Sokolovská 83, 186 75 Prague, Czech Republic
janovsky@karlin.mff.cuni.cz

Drahoslava Janovská
Institute of Chemical Technology, Prague
Technická 5, 166 28 Prague, Czech Republic
janovskd@vscht.cz

In [3], a pathfollowing algorithm has been suggested for a numerical continuation of a few singular values and the relevant left/right singular vectors of a matrix A that depends smoothly on a parameter t. This objective could be linked to Analytical Singular Value Decomposition, see [1]. Note that the proposed technique is suited for large sparse matrices.
The algorithm in [3] may get stuck at isolated points. They are related to *non-generic points*, see [1]. The formulation in [3] makes it possible to investigate these singularities as classical singularities of a mapping and apply techniques of *dimensional reduction*, see e.g. [2]. Using the reduction, it is possible to perform an *asymptotic analysis* of the singular point. This analysis can be exploited to improve the performance of the algorithm in a neighborhood of singular points.
Finally, we will discuss a genericity of the above scenario: Let $A(t) + \varepsilon B(t)$, $a \leq t \leq b$. Does the path persist an arbitrary sufficiently small perturbation?

References

[1] A. Bunse-Gerstner, R. Byers, Mark, V. Mehrmann, and K. Nichols, Nancy, *Numerical computation of an analytic singular value decomposition of a matrix valued function*, Numer. Math., 60 (1991), pp. 1–39.

[2] W. Govaerts, *Numerical Methods for Bifurcations of Dynamical Equilibria*, SIAM, Philadelphia, 2000.

[3] D. Janovská, V. Janovský, and K. Tanabe, *Computing the analytic singular value via a pathfollowing.* submitted to Proceedings of ENUMATH, 2005.

Learning to live with linear dependencies in quantum chemistry: From Hartree-Fock theory to local correlation methods

Curtis Janssen
Scalable Computing Research and Development Department
Sandia National Laboratories
P.O. Box 969, Livermore, CA 94551-0969, USA
cljanss@sandia.gov

Ida Nielsen, Joe Kenny
Scalable Computing Research and Development Department
Sandia National Laboratories
P.O. Box 969, Livermore, CA 94551-0969, USA
{ibniels,jpkenny}@sandia.gov

Advancements in computing hardware, theoretical methods, and parallel algorithms are greatly extending the range of molecular systems that can be treated by *ab initio* quantum mechanical methods. Most of these methods employ the linear combination of atomic orbital approach, where a nonorthogonal basis of atom-centered Gaussian-type functions is used to represent electronic orbitals. As either the size of the molecular system increases (the number of atoms, each carrying a fixed number of basis functions, grows) or the accuracy of the quantum mechanical method increases (the number of basis functions per atoms grows), the degree of approximate linear dependency of the basis set increases (that is, the condition number of the overlap matrix, $S_{ij} = \int \Phi_i(\bar{r})\Phi_j(\bar{r})d\bar{r}$, for the atomic orbitals, $\Phi_i(\bar{r})$, grows). This results in numerical difficulties causing slow or non-convergence, unless the basis set size is reduced. However, automated reduction of the basis set size engenders altered answers and a non-continuous potential energy surface. This talk will illustrate how these issues affect practical computations, and discuss approaches that are currently used to deal with the linear dependencies. The goal of the talk is to highlight these issues, as they affect quantum chemistry, for the mathematical community and to stimulate discussion around best practices to deal with them.

Computing critical delays for neutral time delay systems using a quadratic eigenproblem

Elias Jarlebring
Institut *Computational Mathematics*, TU Braunschweig
Pockelsstr 14, 38106 Braunschweig, Germany
e.jarlebring@tu-bs.de

We present a way to compute the delays $h_1, \ldots, h_m$ of a *neutral time delay system (TDS)*

$$B_0\dot{x}(t) + \sum_{k=1}^{m} B_k\dot{x}(t-h_k) = A_0x(t) + \sum_{k=1}^{m} A_kx(t-h_k),$$

such that the system has an *imaginary eigenvalue*. These delays, called *critical delays*, can be computed by solving a *quadratic eigenproblem* (QEP) resulting from the vectorization of a *quadratic Lyapunov-type condition*. The QEP is the vectorization of a matrix equation and is hence, even for moderately sized TDS, of very large dimension. We show how we can adapt an eigenvalue solver to treat this problem by exploiting the Lyapunov structure of the QEP. The results are generalizations of the previously presented methods for retarded TDS.

StratiGraph for interactive investigation of canonical structures of matrices and matrix pencils

Pedher Johansson
Department of Computing Science, Umeå University
SE-901 87 Umeå, Sweden
pedher@cs.umu.se

Erik Elmroth, Stefan Johansson, Bo Kågström
Department of Computing Science, Umeå University
SE-901 87 Umeå, Sweden
{elmroth, stefanj, bokg}@cs.umu.se

Computing the fine-canonical-structure elements of matrices and matrix pencils are ill-posed problems [10]. Therefore, besides knowing the canonical structure (Jordan, Kronecker, etc) of a matrix or a matrix pencil, it is equally important to know what are the nearby canonical structures that explain the behavior under small perturbations. We present the StratiGraph tool and a prototype toolbox for interactive analysis of these nearness problems [4, 7, 8].
The StratiGraph tool facilitates the application of the complex theory of stratification for matrices and matrix pencils [1, 2, 6]. The toolbox computes canonical structures and determines upper and lower bounds on the shortest distance to a matrix or matrix pencil with a specified structure [5]. The two tools are tightly integrated, making e.g., distance bounds available in StratiGraph and stratification information available in Matlab. The presentation includes an overview of the underlying theory and demonstrations of how the tools can be used to investigate nearness problems for the standard and generalized eigenvalue problems as well as the matrix pencil special cases of controllability and observability matrix pairs [3, 9].

References

[1] A. Edelman, E. Elmroth, and B. Kågström, *A geometric approach to perturbation theory of matrices and matrix pencils. Part I: Versal deformations*, SIAM Journal on Matrix Analysis and its Applications, 18 (1997), pp. 653–692.

[2] A. Edelman, E. Elmroth, and B. Kågström, *A geometric approach to perturbation theory of matrices and matrix pencils. Part II: A stratification-enhanced staircase algorithm*, SIAM Journal on Matrix Analysis and its Applications, 20 (1999), pp. 667–699.

[3] E. Elmroth, P. Johansson, S. Johansson, and B. Kågström, *Orbit and bundle stratification for controllability and observability matrix pairs in StratiGraph*, In proceedings of MTNS 2004, (2004).

[4] E. Elmroth, P. Johansson, and B. Kågström, *Computation and presentation of graphs displaying closure hierarchies of Jordan and Kronecker structures*, Numerical Linear Algebra with Applications, 8 (2001), pp. 381–399.

[5] E. Elmroth, P. Johansson, and B. Kågström, *Bounds for the distance between nearby Jordan and Kronecker structures in a closure hierarchy*, Journal of Mathematical Sciences, 114 (2003), pp. 1765–1779.

[6] E. Elmroth and B. Kågström, *The set of 2-by-3 matrix pencils – Kronecker structures and their transitions under perturbations*, SIAM Journal on Matrix Analysis and its Applications, 17 (1996), pp. 1–34.

[7] P. Johansson, *StratiGraph User's Guide*, UMINF 03.21, ISSN-0348-0542, Dept. of Computing Science, Umeå University, Sweden, 2003.

[8] P. Johansson, *Software Tools for Matrix Canonical Computations*, PhD Thesis, Department of Computing Science, Umeå University, S-901 87 Umeå, Sweden, 2006 (*in preparation*).

[9] S. Johansson, *Stratification of Matrix Pencils in Systems and Control: Theory and Algorithms*, Licentiate Thesis, UMINF-05.17, Department of Computing Science, Umeå University, S-901 87 Umeå, Sweden, May 2005.

[10] B. Kågström, *Singular matrix pencils*, in Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide, Z. Bai et al., eds., SIAM Publications, 2000, pp. 260–277.

Linear equation solvers via robust and optimal control

Jens Jordan
Mathematisches Institut, Universität Würzburg
Am Hubland 97074 Würzburg, Germany
jordan@mathematik.uni-wuerzburg.de

Uwe Helmke
Mathematisches Institut, Universität Würzburg
Am Hubland 97074 Würzburg, Germany
helmke@mathematik.uni-wuerzburg.de

Alexander Lanzon
Research School of Information Sciences and Engineering, Australian National University
Canberra ACT 0200, Australia
alexander.lanzon@anu.edu.au

Solving linear systems of equations is the core problem of numerical linear algebra. In the last 50 years, there has been a huge development of very different powerful iterative algorithms and software packages. Nevertheless, at present, algorithms that combine global stability with low computational complexity are available only for restricted classes of matrices.
We present a new iterative approach which is inspired by feedback stabilization schemes from robust control and yields a control system whose parameters can be tuned to achieve prescribed convergence properties. In particular, we analyze the control system

$$\begin{bmatrix} x_{t+1} \\ u_{t+1} \end{bmatrix} = \begin{bmatrix} I & \Phi \\ -A & -\Xi \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} b$$

with output function $z_t = Ax_t - b$. We apply standard techniques from robust and optimal control to stabilize the dynamics of the system. The resulting sequence of state-vectors $(x_t)_{t\in\mathbb{N}}$ converges to the solution of $Ax = b$. We show that certain choices of the parameters Φ and Ξ yield known algorithms, such as GMRES(m), LQRES or splitting methods. On the other hand, other choices for Φ and Ξ yield new families of algorithms which are globally stable. We optimize a subfamily of those algorithms and compare their convergence properties with the classical algorithms.

Multishift variants of the QZ algorithm with aggressive early deflation

Bo Kågström
Department of Computing Science and HPC2N
Umeå University
S-901 87 Umeå, Sweden
bokg@cs.umu.se

Daniel Kressner
Department of Computing Science and HPC2N
Umeå University
S-901 87 Umeå, Sweden
kressner@cs.umu.se

In this talk, we present novel variants of the QZ algorithm [7] for solving the generalized eigenvalue problem. An extension of the small-bulge multishift QR algorithm [2, 6] is developed, which chases chains of many small bulges instead of only one bulge in each QZ iteration. This allows the effective use of level 3 BLAS operations, which in turn can provide efficient utilization of high performance computing systems with deep memory hierarchies.
Moreover, an extension of the aggressive early deflation strategy [3] is proposed, which can identify and deflate converged eigenvalues long before classic deflation strategies would. Consequently, the number of overall QZ iterations needed until convergence is considerably reduced.
As a third ingredient, we reconsider the deflation of infinite eigenvalues and present a new deflation algorithm, which is particularly effective in the presence of a large number of infinite eigenvalues. Combining all these developments, our implementation significantly improves existing implementations of the QZ algorithm [1, 4]. This is demonstrated by numerical experiments with random matrix pairs as well as with matrix pairs arising from various applications.
We refer to our paper [5] for more details and a more complete listing of relevant references. In this talk, we will also discuss extensions of this work to distributed parallel algorithms (joint work with Björn Adlerborn).

References

[1] E. Anderson, Z. Bai, C. H. Bischof, S. Blackford, J. W. Demmel, J. J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. C. Sorensen, *LAPACK Users' Guide*, SIAM, Philadelphia, PA, third ed., 1999.

[2] K. Braman, R. Byers, and R. Mathias, *The multishift QR algorithm. I. Maintaining well-focused shifts and level 3 performance*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 929–947.

[3] ———, *The multishift QR algorithm. II. Aggressive early deflation*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 948–973.

[4] K. Dackland and B. Kågström, *Blocked algorithms and software for reduction of a regular matrix pair to generalized Schur form*, ACM Trans. Math. Software, 25 (1999), pp. 425–454.

[5] B. Kågström and D. Kressner, *Multishift variants of the QZ algorithm with aggressive early deflation*, Report UMINF-05.11, Department of Computing Science, Umeå University, Umeå, Sweden, Nov. 2005. Submitted to *SIAM J. Matrix Anal. Appl.*

[6] B. Lang, *Effiziente Orthogonaltransformationen bei der Eigen- und Singulärwertzerlegung.* Habilitationsschrift, 1997.

[7] C. B. Moler and G. W. Stewart, *An algorithm for generalized matrix eigenvalue problems*, SIAM J. Numer. Anal., 10 (1973), pp. 241–256.

[1] Also kressner@math.hr, Department of Mathematics, Bijenička 30, 10000 Zagreb, Croatia.

Factorized sparse approximate inverses for statics' simulations

Alexander Kallischko
Lehrstuhl für Informatik V, Technische Universität München
Boltzmannstr. 3, D-85748 Garching, Germany
kallisch@in.tum.de

Thomas Huckle
Lehrstuhl für Informatik V, Technische Universität München
Boltzmannstr. 3, D-85748 Garching, Germany
huckle@in.tum.de

Ralf Mundani
Lehrstuhl für Informatik V, Technische Universität München
Boltzmannstr. 3, D-85748 Garching, Germany
mundani@in.tum.de

For systems $Ax = b$ of linear equations with symmetric positive definite $A \in \mathbb{R}^{n \times n}$, preconditioning can be done with the FSPAI (Factorized SPAI, see [2, 3]) algorithm, which yields the sparse approximate Cholesky factor L_M of a symmetric positive definite preconditioner $M = L_M L_M^T \approx A^{-1}$. The sparsity pattern of L_M can be set a priori, as well as being captured or updated automatically. Promising choices for start patterns are the pattern of A, of A^2 or simply a diagonal pattern for the dynamic case.
In order to reduce the computation time for FSPAI, the effects of some important matrix reordering algorithms are investigated. According to recent results for the classic SPAI algorithm for general matrices [1] there will also be a focus on block-based approaches for the factorized version. Additionally, such methods benefit from FSPAI's high intrinsic parallelity.
One sample application for FSPAI comes from the field of civil engineering. Within statics' simulations, when ordering the finite elements of a finite element discretisation (of p-version type) in a hierarchical way, efficient solvers based on the classical nested dissection algorithm can be applied to compute the respective stresses and displacements [4]. The only drawback of this approach is an ill-conditioned and quite full matrix K at the hierarchy's top level. Hence, K entails a huge amount of iterations when solved with the conjugent gradient method. Here, the FSPAI approach seems perfectly suited for preconditioning the system of linear equations $K \cdot u = d$ on the top level and, thus, to drastically reduce the necessary amount of iterations.
For the construction of a sparse preconditioner we apply a Jacobi prescaling $\tilde{A} = DAD, D = diag(\frac{1}{\sqrt{a_{jj}}})$, $j = 1, \ldots, n$ and a threshold on $\tilde{A}$ which reduces the matrix to its entries greater than a certain ε. Now, FSPAI is computed for the sparse $\tilde{A}$ and can be used as a preconditioner for the preconditioned conjugate gradient method in order to solve the original system $Ax = b$.

References

[1] S. Barnard and M. Grote, *A block version of the spai preconditioner*, 1999.

[2] T. Huckle, *Factorized sparse approximate inverses for preconditioning*, The Journal of Supercomputing, 25 (2003), pp. 109–117.

[3] I. Kaporin, *New convergence results and preconditioning strategies for the conjugate gradient method.*, Numer. Linear Algebra Appl., 1 (1994), pp. 179–210.

[4] R.-P. Mundani, *Hierarchische Geometriemodelle zur Einbettung verteilter Simulationsaufgaben*, dissertation, Fakultät Informatik, Elektrotechnik und Informationstechnik, Universität Stuttgart, 2005.

Computing vibrational eigenstates using Lanczos algorithms and zeroth order eigenstates

Hans O. Karlsson
Department of Quantum Chemistry, Uppsala University
Box 518, SE-75120 Uppsala, Sweden
Hans.Karlsson@kvac.uu.se

Highly excited states play an important role in photo-induced chemistry and physics. To understand and interpret the result of experiments it is important to theoretically be able to describe and compute these vibrational states. From a computational point of view this means that we need to compute a few interior eigenstates of a very large Hermitian matrix.
This is an area of high activity both among applied mathematicians and theoretical chemists, and there are today several methods available to adress this problem.
In this talk we will discuss the feasability of using Lanczos algorithm with chemically motivated start vectors to compute highly excited states, without factorization of the underlying matrix.

A regularized Gauss–Newton method for nonlinear imaging problems in diffuse optical tomography

Misha Kilmer
Mathematics Department, Tufts University
503 Boston Ave., Tufts University, Medford, MA, USA
misha.kilmer@tufts.edu

Eric de Sturler
Department of Mathematics, Virginia Polytechnic Institute
460 McBryde, Virginia Tech, Blacksburg, VA 24061-0123, USA
sturler@vt.edu

The goal in diffuse optical tomography for medical imaging is the joint reconstruction of parameterized images of absorption and scattering of light in the body. The reconstruction requires the approximate solution of a nonlinear least squares problem for the image parameters. While traditional approaches such as damped Gauss-Newton (GN) and Levenberg-Marquardt (LM) have been shown to be effective at solving the imaging problem in its various forms, a considerable number of additional function and Jacobian evaluations are involved in determining the correct step length and/or damping parameter. In 3D imaging problems, depending on the particular image model, the cost of one function evaluation is, at a minimum, the cost of several dense matrix-vector products and in the worst, but more realistic, case requires the solution of several large-scale PDEs. Therefore, it is crucial to keep the number of function and Jacobian evaluations to a minimum.
The ill-conditioning of the Jacobian, together with the presence of noise in the data, motivates us to devise a regularized, trust-region-based Gauss-Newton approach for determining search directions. Although LM can be thought of as a regularized analogue to determining the GN direction, LM has the property of damping possibly important contributions to the search direction in spectral components corresponding to small singular values. On the other hand, the Gauss-Newton direction is too influenced by components due to small singular values early on, causing the line search to work hard to refine the step length. We propose a method that systematically evaluates the potential contribution of each of the spectral components corresponding to the GN-direction and constructs the new direction relative to this contribution within the confines of trust-region. Examples show the success of our method in minimizing function evaluations with respect to other well-known methods.

Kronecker products, tensor decompositions and 3D image processing applications

Misha Kilmer
Department of Mathematics, Tufts University
503 Boston Avenue, Medford, MA, 02052, USA
misha.kilmer@tufts.edu

James Nagy
Department of Mathematics and Computer Science, Emory University
Mathematics and Science Center, Suite W401,400 Dowman Drive, Atlanta, GA, 30322, USA
nagy@mathcs.emory.edu

Lisa Perrone
Department of Mathematics, Tufts University
503 Boston Avenue, Medford, MA, 02052, USA
lisa.perrone@tufts.edu

In 3D image processing applications, such as image restoration, it is not uncommon for the matrices involved to have a significant amount of structure. In this work, we are concerned with matrices having a triply Toeplitz or triply Toeplitz+Hankel structure, depending on the boundary conditions imposed on the desired image. In our applications, these matrices are severely ill-conditioned, with no gap in the singular value spectrum. Due to the presence of noise in the data, an exact or least-squares solution will therefore be hopelessly contaminated by noise.
It is well-known that matrices with such structure admit fast matrix-vector products, and therefore iterative regularization methods from the Krylov or steepest descent families are viable options for producing approximate solutions. These methods can be slow to converge, necessitating the use of preconditioners. To construct a useful preconditioner, however, requires a rank-revealing factorization of the matrix which is typically not computationally feasible.
Therefore, we propose a structured Kronecker product approximation to the matrix operator, and use this approximation to build the preconditioner. We show that our matrix approximation problem is equivalent to a problem of approximating a third-order tensor (i.e. three-way array), representative of a weighted discrete PSF, by a low-rank PARAFAC tensor decomposition. Through this equivalence, we show that in a special case, our matrix Kronecker approximations are optimal in the Frobenius norm. We use our Kronecker product approximation to build a suitable preconditioner for the imaging problem. Through examples in microscopy and medical imaging, we show that the Kronecker approximation preconditioners provide a powerful tool that can be used to improve efficiency of iterative image restoration algorithms.

Unfolding of eigenvalue surfaces near a diabolic point due to a complex perturbation

Oleg Kirillov
Institute of Mechanics, Moscow State Lomonosov University
Michurinskii pr. 1, 119192 Moscow, Russia
kirillov@imec.msu.ru

Alexei Mailybaev, Alexander Seyranian
Institute of Mechanics, Moscow State Lomonosov University
Michurinskii pr. 1, 119192 Moscow, Russia
mailybaev@imec.msu.ru and seyran@imec.msu.ru

- Since the papers by Von Neumann and Wigner (1929) and Teller (1937), it is known that the energy surfaces in quantum physics may cross forming two sheets of a double cone: a diabolo. The apex of the cone is called a diabolic point (Berry and Wilkinson (1984)). This kind of crossing is typical for systems described by real symmetric Hamiltonians with at least two parameters and Hermitian Hamiltonians depending on three or more parameters.

- From a mathematical point of view, the energy surfaces are described by eigenvalues of real symmetric or Hermitian operators dependent on parameters, and the diabolic point is a point of a double eigenvalue with two linearly independent eigenvectors. In modern problems of quantum physics, crystal optics, physical chemistry, acoustics and mechanics, it is important to know how the diabolic point bifurcates under arbitrary complex perturbations forming topological singularities of eigenvalue surfaces like a double coffee filter with two exceptional points or a diabolic circle of exceptional points.

- In the present work we study effects of complex perturbations in multiparameter families of real symmetric and Hermitian matrices. In the case of real symmetric matrices we study the unfolding of eigenvalue surfaces near a diabolic point under real and complex perturbations. The origination of singularities such as a "double coffee filter" and a "diabolic circle" is analytically described. Unfolding of a diabolic point of a Hermitian matrix under an arbitrary complex perturbation is analytically treated. We emphasize that the unfolding of eigenvalue surfaces is described qualitatively as well as quantitatively by using only the information at the diabolic point, including eigenvalues, eigenvectors and derivatives of the system matrix taken at the diabolic point.

- As a physical application, singularities of the surfaces of refractive indices in crystal optics (Berry and Dennis (2003)) are studied. Asymptotic formulae for the metamorphoses of these surfaces depending on properties of a crystal are established and discussed in detail. Singular axes for general crystals with weak absorption and chirality are found. A new explicit condition distinguishing the absorption-dominated and chirality-dominated crystals is established in terms of components of the inverse dielectric tensor. Numerical examples are given to illustrate the general theory.

References

[1] M. Berry and M. Dennis, *The optical singularities of birefringent dichroic chiral crystals*, Proc. R. Soc. Lond. A, 459 (2003), pp. 1261–92.

[2] M. Berry and M. Wilkinson, *Diabolical points in the spectra of triangles*, Proc. R. Soc. Lond. A, 39 (1984), pp. 215–43.

[3] E. Teller, *The crossing of potential surfaces*, J. Phys. Chem., 41 (1937), pp. 109–16.

[4] J. von Neumann and E. P. Wigner, *Über das Verhalten von Eigenwerten bei adiabatischen Prozessen*, Zeitschrift für Physik, 30 (1929), pp. 467–70.

Combinatorial considerations for a bound on subdominant eigenvalues of stochastic matrices

Steve Kirkland
Department of Mathematics and Statistics, University of Regina
Regina, Saskatchewan, Canada, S4S 0A2
kirkland@math.uregina.ca

Suppose that A is a stochastic matrix of order n, and let $\lambda_2(A)$ denote the second largest of the moduli of the eigenvalues of A. A standard upper bound on $\lambda_2(A)$ is given by

$$\lambda_2(A) \leq \tau(A) \equiv \frac{1}{2} \max\{\sum_{k=1}^{n} |a_{i,k} - a_{j,k}| | i, j = 1, \ldots, n\}.$$

It is not difficult to see that $\tau(A) \leq 1$, with equality holding if and only if there is a pair of vertices in the directed graph of A having no common out-neighbours.
In this talk, we investigate the influence of the directed graph of A on the performance of the function τ as a bound on λ_2. Specifically, given a directed graph D on n vertices, let S_D be the collection of $n \times n$ stochastic matrices A such that the directed graph of A is a subgraph of D. We address the following questions.
i) Which directed graphs D have the property that for each $A \in S_D, \lambda_2(A) = \tau(A)$?
ii) Which directed graphs D have the property that for each $c \in (0, 1)$, there is a matrix $A \in S_D$ such that $\lambda_2(A) = c = \tau(A)$?
We also investigate, for various classes of directed graphs D, the absolute and relative errors $\tau(A) - \lambda_2(A)$ and $\frac{\tau(A)-\lambda_2(A)}{\lambda_2(A)}$ as A ranges over S_D.

Fourth-order time stepping for stiff PDEs in the low dispersion limit

Christian Klein
MPI for Mathematics in the Sciences
Inselstr. 22, D-04103 Leipzig, Germany
klein@mis.mpg.de

Solutions to dispersive PDEs without dissipation develop a zone of rapid modulated oscillations near shocks of the dispersionless system. An asymptotic description of the oscillatory zone is available for certain integrable equation as Korteweg-de Vries. To study the situation numerically, we apply the exponential time-differencing fourth-order Runge-Kutta method for solving stiff PDEs by Cox and Matthews [1] as modified by Kassam and Trefethen [2] and extend it to 2+1 dimensional systems. We compare the efficiency of the approach with integrating factor and time-splitting methods for exact solutions and for the low dispersion limit. As an example we study the Kadomtsev-Petviashvili and the Davey-Stewartson equations.

References

[1] S. Cox and P. Matthews, *Exponential time differencing for stiff systems*, J. Comp. Phys., **176**, 430-455, (2001).

[2] A.-K. Kassam and L. Trefethen, *Fourth-order time-stepping for stiff pdes*, SIAM J. Sci. Comp., **26**, 1214, (2005).

The quality of the Gauss–Arnoldi quadrature for $\langle (zI - A)^{-1}\varphi, \varphi\rangle$ and application to Padé-like approximation

Leonid Knizhnerman
Mathematical Modelling Department, Central Geophysical Expedition
Russia, 123298, Moscow, Narodnogo Opolcheniya St., house 38, building 3
mmd@cge.ru

The quality of the Gauss–Arnoldi quadrature for computation of the quantity $\langle (zI - A)^{-1}\varphi, \varphi\rangle$ has been studied, where A is a bounded operator in Hilbert space and φ is a nonzero vector. For normal operators, we have established a necessary condition and a sufficient condition for the quadrature to be effective. We reckon that the quadrature is non-trivial, if its error decreases exponentially faster than the error of Arnoldi's approximant to $(zI - A)^{-1}\varphi$ does, as the number of an Arnoldi step tends to infinity. For non-normal operators, we have presented a pessimistic example.
It has been shown that in certain cases the Gauss–Arnoldi quadrature is related to Padé-like approximation (with poles at Ritz values) of Markov type functions (with generating measures supported in the complex plane) and, in particular, can serve as a means to localise the poles of "a rational perturbation". The corresponding error estimates, formulated in terms of the spectrum and Green function, are non-trivial even in cases when the classical Padé approximation malfunctions or its ability is not theoretically validated.
The results of numerical experiments illustrate the theoretical assertions.

Accurate eigenvalues of elliptic PDEs using finite elements

Plamen Koev
Department of Mathematics, Massachusetts Institute of Technology
77 Massachusetts Ave. 2-376, Cambridge, MA 02138, U.S.A.
plamen@math.mit.edu

James Demmel
Department of Mathematics and Computer Science Division, University of California, Berkeley
Berkeley, CA 94720, U.S.A.
demmel@cs.berkeley.edu

Stephen Vavasis
Department of Computer Science, Cornell University
4130 Upson Hall, Cornell University, Ithaca, NY 14853
vavasis@cs.cornell.edu

We will present our recent work on computing eigenvalues of elliptic PDEs using tetrahedral finite elements. The eigenvalues are computed to high relative accuracy in $O(n^3)$ time using only working precision. The complexity of our algorithm is independent of the material properties. The requirements on the discretization are modest and easy to attain using existing meshing algorithms.

References

[1] E. G. Boman, B. Hendrickson, and S. Vavasis, *Solving elliptic finite element systems in near-linear time with support preconditioners*. preprint, 2005.

[2] J. Demmel, P. Koev, and S. Vavasis, *Accurate eigenvalues for elliptic problems using finite elements*. in preparation.

Accurate eigenvalues of sign-regular matrices

Plamen Koev
Department of Mathematics, Massachusetts Institute of Technology
77 Massachusetts Ave. 2-376, Cambridge, MA 02138, U.S.A.
plamen@math.mit.edu

We will present new algorithms for computing all eigenvalues of certain sign regular matrices to high relative accuracy. The sign regular matrices are characterized by the property that all minors of order k have the same sign $\varepsilon_k \in \{1, -1\}$ for all $k = 1, 2, \ldots, n$, where n is the size of the matrix. The ones under consideration have $\varepsilon_k = (-1)^{k(k-1)/2}$.
All eigenvalues of a sign regular matrix are real, but need not be positive. We exploit the connection between sign regular and totally positive matrices, and in particular, the characterization of the latter as products of nonnegative bidiagonals.
All eigenvalues are computed to high relative accuracy using only working precision in $O(n^3)$ time.
Neither the lack of symmetry in the matrix nor the value of the condition number $\kappa(A) \equiv \|A\| \cdot \|A^{-1}\|$ have any adverse effect on the accuracy of our algorithm.

References

[1] F. Gantmacher and M. Krein, *Oscillation Matrices and Kernels and Small Vibrations of Mechanical Systems*, AMS Chelsea, Providence, RI, revised ed., 2002.

[2] P. Koev, *Accurate eigenvalues and SVDs of totally nonnegative matrices*, SIAM J. Matrix Anal. Appl., 27 (2005), pp. 1–23.

Reduction of the problem size through linear algebra for the problem of reactive transport in porous media

Serge Kräutle
Institut für Angewandte Mathematik, Friedrich-Alexander-Universität Erlangen-Nürnberg
Martensstr.. 3, D-91058 Erlangen, Germany
kraeutle@am.uni-erlangen.de

Peter Knabner
Institut für Angewandte Mathematik, Friedrich-Alexander-Universität Erlangen-Nürnberg
Martensstr. 3, D-91058 Erlangen, Germany
knabner@am.uni-erlangen.de

We are considering a system consisting of partial differential equations (PDEs), ordinary differential equations (ODEs) and algebraic equations (AEs), as it frequently occurs for the computation of reactive flow in forous media:

$$\begin{array}{lll} \frac{d}{dt}c_i + Lc_i & = \sum_{j=1}^{J} s_{ij} R_j(\mathbf{c}, \bar{\mathbf{c}}), & i = 1, ..., I \\ \frac{d}{dt}\bar{c}_i & = \sum_{j=1}^{J} \bar{s}_{ij} R_j(\mathbf{c}, \bar{\mathbf{c}}), & i = 1, ..., \bar{I} \\ Q_j(\mathbf{c}, \bar{\mathbf{c}}) & = 0, & j = 1, ..., J_{eq} \end{array}$$

Here, L denotes a linear transport (i.e., advection-dispersion/diffusion) operator, the c_i denote the concentrations of I mobile chemical species (in the groundwater), the $\bar{c}_i$ denote the concentrations of the $\bar{I}$ immobile chemical species (adsorbed to the soil), the R_j denote the (nonlinear) rates of J chemical reactions, depending on the concentrations, the Q_j denote some (nonlinear) equilibirum conditions, if some of the reactions are assumed to be so fast that they can be considered to be in equilibrium. It is common use to assume the chemical law of mass action for equilibrium reactions, i.e.,

$$Q_j(c, \bar{c}) = \sum_{i=1}^{I} s_{ji} \ln c_i + \sum_{i=1}^{\bar{I}} s_{ji} \ln \bar{c}_i - \ln k_j, \quad j = 1, ..., J_{eq}, \quad (k_j = \text{equilibrium constant}).$$

Our intention is to propose a method how to reformulate such a system such that in the resulting formulation, many of the differential equations decouple from the remaining, now smaller, nonlinear system. This is an urgent task, since the given nonlinear systems, typically consisting of $I + \bar{I} = 10...20$ coupled nonlinear (partial) differential equations, require, in two and even more in three space dimensions, very large computer resources. The efficient numerical solution of the equations under consideration is needed, for example, to predict the evolution of contamination sites in the groundwater/in the Earth, and the effect of possible remediation strategies of the contamination site.
Standard methods [3], but also some newer developments [2] for the equations under consideration usually make use of an operator splitting between the transport and the chemistry, which often leads to additional splitting errors or other disadvantages.
In contrast to this, the reformulation we propose [1] does not use any artificial splitting. It consists of (1) taking linear combinations of the given ODEs/PDEs, (2) introducing new variables (change of basis in the solution space $\mathbb{R}^{I+\bar{I}}$), and (3) solving the nonlinear AEs for certain variables and substitution of these variables in the remaining differential equations. Steps (1)-(2) lead to a decoupling of certain, then linear, PDEs, and step (3) leads to a further size reduction of the system.
Some example problems to which we apply the reduction method are considered.

References

[1] S. Kräutle and P. Knabner, *A new numerical reduction scheme for fully coupled multicomponent transport-reaction problems in porous media*, Water Resour. Res., 41 (2005). doi:10.1029/2004WR003624.

[2] S. Molins, J. Carrera, C. Ayora, and M. Saaltink, *A formulation for decoupling components in reactive transport problems*, Water Resour. Res., 40 (2004). doi:10.1029/2003WR002970.

[3] C. Steefel and K. MacQuarrie, *Approaches to modeling of reactive transport in porous media*, in: Reactive transport in porous media, Reviews in Mineralogy, Vol. 34, Lichtner, Steefel, and Oelkers (editors), (1996).

Structure-preserving iterative refinment of invariant subspaces

Daniel Kressner
Department of Computing Science, Umeå University
SE-90187 Umeå, Sweden
kressner@cs.umu.se

Ralph Byers
Department of Mathematics, University of Kansas
Lawrence, KS 66045-7523
byers@math.ku.edu

Starting from recent results on structured condition numbers for invariant subspaces and associated matrix equations [1], a novel framework for structure-preserving iterations for refining an invariant subspace of a matrix or a deflating subspace of a matrix pencil is presented. We discuss several variants, such as Newton methods and Rayleigh quotient iterations on Grassmann manifolds. The *potential* merits of the new methods applied to structured matrices compared to existing general methods are better local and global convergence properties, higher efficiency and higher attainable final accuracy. The framework is general enough to cover many of the structures that admit structured Schur factorizations, including Hamiltonian matrices, symplectic matrices, and palindromic matrix pencils.

References

[1] R. Byers and D. Kressner, *Structured condition numbers for invariant subspaces*, 2005. To appear in SIAM J. Matrix Anal. Appl.

Parallel hierarchical matrices

Ronald Kriemann
Max-Planck-Institute for Mathematics in the Sciences
Inselstr. 22-26, D-04103 Leipzig, Germany
rok@mis.mpg.de

Many algorithms in the context of $\mathcal{H}$-matrices contain a large degree of parallelism. Consider for example the construction of a matrix, where each matrix block can be built independent of the others. Similar properties apply to matrix-vector multiplication, matrix-multiplication and, to some degree, to matrix inversion (see [1, 2]). Another advantage of parallelising $\mathcal{H}$-matrices is the distribution of the consumed memory to several machines, allowing the computation of larger problems.
In this talk parallel versions of the basic $\mathcal{H}$-matrix arithmetics for solving FEM-/BEM-problems will be presented. To cover most of the today available computer architectures, the algorithms will be discussed for shared and distributed memory systems. Numerical experiments are finally used to demonstrate the parallel efficiency of the discussed methods.

References

[1] M. Bebendorf and R. Kriemann, *Fast Parallel Solution of Boundary Integral Equations and Related Problems*, Computing and Visualization in Science, 8 (2005), pp. 121–135.

[2] R. Kriemann, *Parallel H-Matrix Arithmetics on Shared Memory Systems,*, Computing, 74 (2005), pp. 273–297.

Special product triangular skew-Hermitian splitting iteration methods for strongly non-Hermitian positive definite linear equation systems

Lev Krukier
Rostov State University
Computing Center, Stachki Avenue 200/1, Building 2, Rostov-on-Don 344090, Russia
krukier@rsu.ru

Tatiana Martynova
Rostov State University
Computing Center, Stachki Avenue 200/1, Building 2, Rostov-on-Don 344090, Russia
martynova@rsu.ru

Zhong-Zhi Bai
Institute of Computational Mathematics and Scientific/Engineering Computing
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
P.O. Box 2719, Beijing 100080, P.R. China
bzz@lsec.cc.ac.cn

By further generalizing the modified skew-Hermitian triangular splitting iteration methods studied by Wang and Bai (BIT Numerical Mathematics 44(2004)363-386), in this paper, we present a new iteration scheme, called the product-type skew-Hermitian triangular splitting iteration methods, for solving the strongly non-Hermitian systems of linear equations with positive definite coefficient matrices. We discuss the convergence property and the optimal parameters of this method. Some numerical experiments have been executed.

Multiscale methods for PDE-constrained control problems: optimal preconditioners and fast iterative solvers

Angela Kunoth
Institut für Angewandte Mathematik & Institut für Numerische Simulation
Universität Bonn
Wegelerstr.6, D-53115 Bonn, Germany
kunoth@iam.uni-bonn.de

The fast numerical solution of PDE-constrained control problems provides a formidable challenge, in particular, for problems constrained by time-dependent PDEs. For a single elliptic PDE which may be considered as the core system to be solved in any case, there exist by now three classes of preconditioners, all of which are of multiscale structure (multigrid, BPX and wavelet preconditioners), which assure uniformly bounded spectral condition numbers of the system matrix independent of the discretization. In my talk I wish to address fast multiscale solvers for control problems constrained by elliptic PDEs with distributed as well as Dirichlet boundary control. I will present the main ingredients for obtaining theoretical estimates which guarantee optimality of the multilevel preconditioners. Together with employing iterative solvers and a nested iteration scheme, I will show that they therefore provide the solution ingredients of the control problem (state, costate and control) up to discretization error accuracy in optimal linear complexity. Corresponding numerical results confirming these estimates will be shown. Finally, I would like to address optimal preconditioners for control problems constrained by parabolic PDEs.

Derivation of high-order spectral methods for time-dependent PDE using modified moments

James V. Lambers
Department of Petroleum Engineering, Stanford University
367 Panama St Rm 065, Stanford, CA 94305-2220, USA
lambers@stanford.edu

This talk summarizes recent and ongoing work on an alternative approach to the solution of diffusion problems and wave propagation problems in the variable-coefficient case that leads to a new numerical method, called a Krylov subspace spectral method [3], for which the Fourier components of the computed solution can easily be represented as a function of time, as in the constant-coefficient case.
The basic idea behind the method, applied to a PDE of the form $du/dt + L(x, D)u = 0$, is to use Gaussian quadrature in the spectral domain to compute Fourier components of the solution by treating them as off-diagonal elements of $\exp[-L\Delta t]$ for a matrix L discretizing $L(x, D)$ and time step Δt, using algorithms developed in [1], as opposed to applying Gaussian quadrature in the spatial domain as in traditional spectral methods.
For each component, a different approximation of the solution operator by a restriction to a low-dimensional Krylov subspace is employed, and each approximation is optimal in some sense for computing the corresponding component. This strategy allows accurate resolution of all desired components, for both high and low frequencies, without having to resort to smoothing techniques to ensure stability. In fact, by viewing the Fourier components in terms of directional derivatives of moments, we can demonstrate unconditional stability given sufficient smoothness of the coefficients of $L(x, D)$.
We will also see how this approach can be fashioned into practical numerical solvers of PDE. By exploiting the simple structure of differential operators, we can perform the Lanczos algorithm on an operator with several starting vectors simultaneously, generating Lanczos vectors and Jacobi matrices that are functions of the wave number. In addition, we present an algorithm for rapidly updating the output of the Lanczos algorithm when the starting vector is changed, which also improves the performance significantly.
Next, we take advantage of the fact that like the Fourier method applied to constant-coefficient problems, Krylov subspace spectral methods compute solutions to variable-coefficient problems whose Fourier components have a simple-closed form representation as a function t that can easily be evaluated and analytically differentiated. This representation yields several benefits, including (1) a new approach to deferred correction that uses a highly accurate residual, (2) an algorithm that computes Fourier components independently using different time scales, as in the constant-coefficient case, and (3) a high-order method for second-order wave equations (see [2]).
Numerical results will be presented for both parabolic and hyperbolic problems in one, two and three space dimensions.

References

[1] G. Golub and G. Meurant, *Matrices, moments and quadrature*, in Proceedings of the 15th Dundee Conference, D. F. Griffiths and G. A. Watson, eds., Longman Scientific & Technical, 1994.

[2] P. Guidotti, J. V. Lambers, and K. Sølna, *Analysis of wave propagation in 1d inhomogeneous media*, Numerical Functional Analysis and Optimization, 27 (2006).

[3] J. V. Lambers, *Krylov subspace spectral methods for variable-coefficient initial-boundary value problems*, Electronic Transactions on Numerical Analysis, 20 (2005), pp. 212–234.

Spectral conditions for stability of vibrating systems

Peter Lancaster
Department of Mathematics and Statistics, University of Calgary
Calgary AB T2N 1N4, Canada
lancaste@ucalgary.ca

Let $L(\lambda) = \lambda^2 M + \lambda D + K$ correspond to a vibrating system with real symmetric coefficients and M nonsingular. The well-known sufficiant condition for stability $M > 0,\ D > 0,\ K > 0$ is far from being necessary. Using recent work on inverse problems for matrix polynomials (in collaboration with U. Prells), I will discuss conditions on the spectral data for a vibrating system which will ensure the above conditions on $M,\ D$ and K, and hence stability. Problems of elliptic, hyperbolic and mixed types will be mentioned.

GCR vs Flexible GMRES

Julien Langou
Innovative Computing Lab., Computer Science Department, University of Tennessee
1122 Volunteer Blvd., Claxton Bldg. , TN 37996 Knoxville, USA
langou@cs.utk.edu

When dealing with a nonsymmetric linear system, the most used solver is certainly the GMRES method [2]. Another equivalent method in exact arithmetic is the GCR method [4]. While GMRES needs only to store one basis (the Arnoldi basis), GCR needs to store two basis (the residual basis and the search direction basis). Therefore GMRES is superior to GCR in term of memory storage but also in term of FLOPS (few AXPY more). This analysis holds in exact arithmetic, with no restart, one right-hand side and with a fixed preconditioner.
While GCR handles naturally variable preconditioners, a modification of GMRES is necessary so that GMRES handles variable preconditioners. The resulting algorithm is named Flexible GMRES [1]. Flexible GMRES and GCR have the same memory requirement. Therefore the question arises which one to choose between Flexible GMRES and GCR.
In this talk I will present five experiments that has arisen in various context of my research and that leads me to the conclusion that GCR was better than Flexible GMRES.

1. GCR allows for truncation as opposed to restart. Truncation is clearly superior to restart in particular if the matrix is close to a symmetric matrix.
2. Block GCR is superior to Block (Flexible) GMRES because deflation is possible in the block of residuals while it is not straightforward in the Arnoldi sequence.
3. When comparing the inner/outer scheme Flexible GMRES/GMRES vs GCR/GMRES in the context of electromagnetism application, it turns out that GCR/GMRES (also known as GMRESR [3]) is always better than Flexible GMRES/GMRES.
4. One way to use the iterative refinement procedure is to perform a dense LU factorization done in single precision and iterate in double precision arithmetic to obtain normwise backward error of the order of the double precision machine epsilon. This procedure is interesting when single precission is significantly faster than double precision (e.g. SSE2, Cell processor). The classic scheme for the iterative refinement is Richardson. If we remplace it with GCR, convergence is significantly improves for all matrices. When approaching dealing with matrices with condition number as high as 10^8, we have observed that GCR is always able to converge in less than 20 steps while FGMRES (and GMRES, and Richardson) have some failures mode.
5. In the context of fault tolerant algorithms, we consider the case where the results of a matrix-vector product or a preconditioning step is sometimes completely false. We interpret a false matrix-vector product as a flexible preconditioner step. In this context the use of GCR or Flexible GMRES enables us to obtain convergence of the algorithm without detecting nor recovering from the errors. Once more in this context, GCR is clearly the algorithm of choice. Convergence happens in the number of iterations without fault plus the number of faults while the convergence is seriously damaged in Flexible GMRES.

Finally we will conclude our discussion by considering the limiting accuracy of these two solvers.

References

[1] Y. Saad, *A flexible inner-outer preconditioned gmres algorithm*, SIAM J. Sci. Statist. Comput., 14 (1993), pp. 461–469.

[2] Y. Saad and M. H. Schultz, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.

[3] H. A. Van der Vorst and V. C., *GMRESR: a family of nested gmres methods*, Numer. Lin. Alg. with Appl., 1 (1994), pp. 369–386.

[4] D. M. Young and K. C. Jea, *Generalized conjugate-gradient acceleration of nonsymmetrizable iterative methods*, Linear Algebra Appl., 34 (1980), pp. 159–195.

Discrete dynamical systems and preconditioned eigensolvers

Richard Lehoucq
Sandia National Labs
Albuquerque, NM (USA)
rblehou@sandia.gov

Mark Embree
Rice University
Houston, TX (USA)
embree@rice.edu

We present an isomorphism between discrete dynamical systems and a class of preconditioned eigen-iterations for the eigenvalue problem. This isomorphism provides a framework to investigate the convergence and stability of existing eigen-iterations and propose new techniques. The isomorphism with discrete dynamical systems allows us to consider geometric properties. In particular, we demonstrate that there are first integrals, or invariants, that must be satisfied by the eigen-iteration. Satisfaction of these invariants is demonstrated to be crucial to the numerical stability of the eigen-iteration. Moreover, these invariants provide a mechanism for monitoring the stability of the eigen-iteration.

A primal based penalty preconditioner for elliptic saddle point systems

Richard Lehoucq
Sandia National Laboratories
P.O. Box 5800, MS 1110, Albuquerque, NM 87185-1110, USA
rblehou@sandia.gov

Clark Dohrmann
Sandia National Laboratories
P.O. Box 5800, MS 0847, Albuquerque, NM 87185-0847
crdohrm@sandia.gov

A primal based penalty preconditioner is presented for a linear set of equations arising from elliptic saddle point problems. We show that the eigenvalues of the preconditioned matrix are positive real and demonstrate that a variant of the preconditioner can be combined with the conjugate gradient algorithm. Our approach is motivated by two basic observations. First, the solution of a problem with constraints is often similar to the solution of a problem where the constraints are penalized. Second, certain methods of solution not available for a constrained problem are possible for its penalized counterpart so motivating a primal based Schur complement approach. Numerical examples for elliptic 2D and 3D problems are presented that confirm theoretical results and demonstrate the effectiveness of the preconditioner.

Challenges for the analysis and understanding of Krylov subspace methods

Jörg Liesen
Institute of Mathematics, TU Berlin
Straße des 17. Juni 136, 10623 Berlin, Germany
liesen@math.tu-berlin.de

Krylov subspace methods such as CG and GMRES have become household names in the area of scientific computing, and they are widely used in applications throughout science and engineering. The basic algorithms can be written down in a few lines, and the whole CG method can be programmed (say, in MATLAB) in a few minutes. Nevertheless, it seems that the *understanding* of Krylov subspace methods does not progress with the same speed as the algorithmic developments in this area. In this talk we will discuss some difficult questions which still largely remain open.
We will first consider the question of *convergence* of Krylov subspace methods. It is essential to realize that they behave nonlinearly, ideally (in exact arithmetic) terminate in finitely many steps, and usually can be stopped early, namely when the user-specified accuracy of the approximation is reached. How should the term "convergence" be used in this context, and how can the convergence behavior be properly described and analyzed? In particular, the behavior of Krylov subspace methods depends not only on the matrix, but also on the specific right hand side (e.g. the boundary conditions and/or source terms in case of discretized PDEs). Is this dependence significant in practical computations, and if yes, then how does this affect the analysis of convergence?
Next, we will study the question of the *computational cost* of Krylov subspace methods. This naturally combines the convergence and numerical stability in the description and analysis of the *numerical convergence behavior*. How should this behavior be evaluated, and what is the relation of computational cost to the widely used concept of complexity of numerical algorithms?

Stopping criteria for mixed finite element methods

Daniel Loghin
University of Birmingham
Watson Building, Edgbaston Birmingham UK
loghind@maths.bham.ac.uk

Mario Arioli
Rutherford Appleton Laboratory
Chilton, Didcot, Oxfordshire, UK
M.Arioli@rl.ac.uk

Black-box stopping criteria employed by Krylov-type solvers can be shown to be often wasteful and irrelevant to finite element practice. The concept of energy-norm as suitable stopping criterion – now classical in the context of CG convergence – is also applicable to non-symmetric iterations. In particular, the norm induced by the finite element application can be used to indicate convergence of Krylov-type solvers. The main issues here are the cheap and accurate estimation of such criteria.
The case of mixed finite element methods is of particular interest due to the large class of problems formulated in this fashion. The main difficulty here is the indefiniteness of the resulting system matrix, often bearing no relation to the discrete finite element norms underlying the formulation. Nevertheless, consideration of such norms leads to good estimators of convergence for non-symmetric solvers. We demonstrate this for the case of GMRES used in the solution of some standard linear saddle-point problems. We also investigate the more interesting (and realistic) case of non-linear saddle-point problems and give a number of guidelines aimed at useful estimation of the proposed criteria. We illustrate the gains on Newton-GMRES algorithms for problems arising from CFD and domain decomposition.

Definite pencils and hyperbolic polynomials

D. Steven Mackey
School of Mathematics, University of Manchester
Manchester, M60 1QD, UK
smackey@ma.man.ac.uk

Nicholas J. Higham
School of Mathematics, University of Manchester
Manchester, M60 1QD, UK
higham@ma.man.ac.uk

Françoise Tisseur
School of Mathematics, University of Manchester
Manchester, M60 1QD, UK
ftisseur@maths.manchester.ac.uk

Among matrix polynomial eigenproblems, those for definite pencils and for hyperbolic quadratics can be regarded as the closest analogs of the standard eigenproblem for Hermitian matrices. In this talk we generalize the usual definition of hyperbolic quadratic, and see how this generalization both simplifies and unifies the notions of hyperbolic quadratic polynomial and definite pencil. This connection is then further strengthened by showing that any (generalized) hyperbolic quadratic can always be linearized by a definite pencil.

Nonlinear eigenvalue problems in energy band calculation of semiconductor quantum dots

Marta Markiewicz
Institute of Numerical Simulation, Hamburg University of Technology
Schwarzenbergstr. 95, D-21073 Hamburg, Germany
m.markiewicz@tuhh.de

Heinrich Voss
Institute of Numerical Simulation, Hamburg University of Technology
Schwarzenbergstr. 95, D-21073 Hamburg, Germany
voss@tuhh.de

Quantum dots are semiconductor nanostructures in which the free carriers are confined by the potential barriers to a small region in all three directions of space. Moreover, if the size of the region is smaller than an electron wavelength the energy states become quantized at discrete energy levels as it happens in an atom. This property makes the quantum dots very attaractive for applications in micro and optoelectronic devices and allows to use them for investigating systems at the atomic level.
The energy states of a quantum dot embedded into a matrix are often computed as the eigensolutions of the Schroedinger equation

$$H_0\phi(r) := -\nabla\frac{\hbar^2}{2m(r,\lambda)}\nabla\phi(r) + V(r)\phi(r) = \lambda\phi(r), \quad r \in \Omega_d \cup \Omega_m,$$

where the nonparabolic effective mass approximation $m(r,\lambda)$ depends nonlinearly on the eigenparameter. This description does not take into account the spin effects. The electron spin plays an important role in the quantum dot design. The interaction of the electron spin with its angular momentum (spin–orbit interaction) can significantly alter the electron energy spectrum. To include the spin–orbit interaction the Hamiltonian is augmented by the following term

$$V_{so} = i\nabla\beta(r,\lambda)\cdot[\hat{\sigma}\times\nabla],$$

where $\beta(r,\lambda)$ is a spin–orbit coupling parameter, which is a nonlinear function of the eigenparameter as well.
The resulting Hamiltonian $H = H_0 + V_{so}$ defines a rational eigenvalue problem, since $m(r,\lambda)^{-1}$ and $\beta(r,\lambda)$ both depend rationally on the energy.
The spin–orbit interaction constributes to the Hamiltonian only on the dot–matrix interface. We discretize the Hamiltonian with the finite element method and include this extra boundary term directly into the discretized problem. Thus, we obtain a matrix rational eigenvalue problem, with relatively small rank spin–orbit contribution, since V_{so} is nonzero only on the interface, which we solve by the nonlinear Arnoldi method.

Determining hyperbolicity and ellipticity of quadratic eigenvalue problems

Roy Mathias
School of Mathematics, University of Birmingham
Birmingham B15 2TT, United Kingdom
mathias@maths.bham.ac.uk

I will present new methods to determine whether a quadratic eigenvalue problem is hyperbolic or elliptic. These methods are computationally more efficient and conceptually simpler than existing methods, and in the case of hyperbolicty, more robust.
This is joint work with Nick Higham and Francoise Tisseur–both at the University of Manchester.

A new multilevel ILU-preconditioner with pivoting by columns and row interchanges to improve sparsity

Jan Mayer
Fakultät für Mathematik, Universität Karlsruhe
Kaiserstraße 12, D-76128 Karlsruhe, Germany
jan.mayer@math.uni-karlsruhe.de

In this talk, we will present a new multilevel preconditioner intended to be used with iterative methods for solving a sparse linear system. It is based on Crout's formulation of Gaussian elimination and implements pivoting by columns. This preconditioner uses additional row interchanges which are aimed at producing a sparser preconditioner. Several different stratigies for permuting rows, for dropping elements to promote sparsity and for terminating one level in order to begin a new one will be discussed. We will conclude with numerical results.

Triangular representations of second order systems

Christian Mehl
Institut für Mathematik, Technische Universität Berlin
Sekretariat MA 4-5, 10623 Berlin
mehl@math.tu-berlin.de

The problem of solving second order systems leads to the quadratic eigenvalue problem $P(\lambda) = 0$, where $P(\lambda) = \lambda^2 M + \lambda C + k$ is a quadratic matrix polynomial. Usually, this problem is then solved by linearization. However, when the leading coefficient matrix M is singular, then Gohberg, Kaashoek, and Lancaster observed that the so-called concept of *strong linearization* has to be used in order to properly reflect the structure at the eigenvalue infinity. A corresponding equivalence relation can be imposed on the set of all matrix polynomials.
In this talk, we consider the question whether any quadratic matrix polynomial allows an equivalent representation, i.e., a representation with the same spectral information, where all coefficient matrices are upper triangular. Moreover, we discuss how the index of a second order system is reflected in the spectral information of the underlying matrix polynomial.

The mathematical diaries of Helmut Wielandt: a gold mine for problems and solutions in applied and numerical linear algebra

Volker Mehrmann
Institut für Mathematik, TU Berlin
Str. des. 17. Juni 136, D-10623 Berlin, Germany
mehrmann@math.tu-berlin.de

When Helmut Wielandt died in 2001, his family donated 19 mathematical diaries (from the early fourties to the eighties) and a lot of mathematical correspondence to the public for scientific editing and archivation. Since 2002 the Deutsche Forschungsgemeinschaft supports a research project by O. Holtz, I. Ipsen, W. Knapp, V. Mehrmann, P. Schmidt and H. Schneider to edit this work, and soon the complete diaries will be available to the public as facsimiles, latex and pdf files (unfortunately in german).
Some of the diaries are already at `http://www.math.tu-berlin.de/numerik/mt/Wielandt/` where also Helmut Wielandt's list of publications, part of his mathematical correspondence and related publications can be found.
Besides reviews of work by other people, the diaries contain a tremendous amount of unpublished work by Helmut Wieland as well as open problems and short and elegant proofs of known results.
Most of the work in the diaries is concerned with applied and numerical linear algebra and the theory of finite groups but also results in function theory, operator theory and differential equations can be found.
As examples, we will present results concerning concepts of normality and quasi-commutative matrices.
We will also discuss some historical facts from the life of Helmut Wielandt which shed some light on the life of a mathematician in Nazi Germany.

A singular value characterization of the distance to uncontrollability for higher order systems and a practical algorithm exploiting the singular value characterization

Emre Mengi
Courant Institute of Mathematical Sciences
Computer Science Department, New York University
251 Mercer Street, New York, NY 10012 USA
mengi@cs.nyu.edu

A fundamental question concerning the kth order continuous time-invariant dynamical system

$$K_k x^{(k)}(t) + \ldots + K_1 x'(t) + K_0 x(t) = Bu(t), \quad x(0) = c_1, x'(0) = c_2, \ldots x^{(k-1)} = c_{k-1} \tag{1}$$

is the dimension of the subspace of reachable configurations at a given time t' where $B \in \mathbb{C}^{n\times m}$, $K_0, K_1, \ldots K_k \in \mathbb{C}^{n\times n}$, $x(t) \in \mathbb{C}^n$, $u(t) \in \mathbb{C}^m$ and $c_1, c_2, \ldots, c_{k-1} \in \mathbb{C}^n$ are given initial conditions. For simplicity we assume that the leading coefficient is non-singular and additionaly when perturbations to the leading coefficient are allowed, the leading coefficient remains non-singular under perturbations. We define the space of reachable configurations at time t' as

$$\mathcal{R}_{t'} = \{(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{k-1}) \ : \ \exists u(t) \ \text{ the system (1) satisfies}, \varepsilon_0 = x(t'), \ \varepsilon_1 = x'(t'), \ \ldots, \ \varepsilon_{k-1} = x^{(k-1)}(t')\}.$$

The system (1) (alternatively the tuple $(K_k, \ldots, K_1, K_0, B)$) is controllable if and only if

$$\dim(\mathcal{R}_{t'}) = nk,$$

that is all of the configurations are reachable.
Controllability of the system (1) is equivalent to the rank problem

$$\forall \lambda \ \ \text{rank}([P(\lambda) \ \ B]) = n$$

with

$$P(\lambda) = \sum_{j=0}^{m} \lambda^j K_j.$$

Therefore it can be checked efficiently but not reliably under rounding errors. Furthermore a controllable system may still have nearby uncontrollable systems which potentially is an indicator of a problem with the model. To overcome these problems we define the distance to the closest uncontrollable system as

$$\tau(P, B, \alpha) = \inf\{\|\Delta K_k \ \ldots \ \Delta K_0 \ \Delta B\| : \text{the tuple } (K_k + \alpha_k \Delta K_k, \ldots, K_0 + \alpha_0 \Delta K_0, B + \Delta B) \text{ is uncontrollable}\}$$

where $\alpha = (\alpha_k, \ldots, \alpha_1, \alpha_0)$ consists of non-negative real numbers not all of them zero and $\|\cdot\|$ denotes either the spectral or the Frobenius norm. With $k = 1$, $K_1 = I$, $K_0 = -A$ and $\alpha = (0, 1)$ we recover the standard definition of the distance to uncontrollability for the first order system $x'(t) = Ax + Bu$.
In this work we provide a singular value minimization characterization for the higher order distance to uncontrollability, namely

$$\tau(P, B, \alpha) = \inf_{\lambda\in\mathbb{C}} \sigma_{\min} \left[\begin{array}{cc} \dfrac{P(\lambda)}{\sqrt{s_\alpha(|\lambda|)}} & B \end{array} \right], \ \text{ where } \ s_\alpha(|\lambda|) = \sum_{j=0}^{k} \alpha_j^2 |\lambda|^{2j}.$$

We describe a practical trisection algorithm locating the global minimum of the equivalent optimization problem. The trisection algorithm depends on the extraction of the imaginary eigenvalues of even-odd matrix polynomials (a matrix polynomial with even Hermitian coefficients and odd skew-Hermitian coefficients) of size $2n$ and degree $2k$ and has a cost of $O\left(\frac{1}{\arccos(1-(\frac{tol}{k})^2)} n^3 k^4\right)$ where tol denotes the required accuracy.

On algebraic multilevel methods for non-symmetric system

Christian Mense
Technische Universität Berlin, Institut für Mathematik, MA 3-3
Strasse des 17. Juni 136, D-10623 Berlin, Germany
mense@math.tu-berlin.de

Reinhard Nabben
Technische Universität Berlin, Institut für Mathematik, MA 3-3
Strasse des 17. Juni 136, D-10623 Berlin, Germany
nabben@math.tu-berlin.de

The 'Algebraic Multilevel Iteration' (AMLI) developed by Axelsson and Vassilevski [2, 3] is an algebraic multigrid method, which uses an approximation of the Schur complement as a coarse grid operator. Convergence results for the AMLI method are given in [2, 3, 4, 1]. However, the results can be applied only to so called Stieltjes matrices, i.e. symmetric positive definite M-Matrices. In [5] some results are given for the multilevel approximate block factorization or AMLI method applied to some special non-symmetric M-matrices which arise from a specific discretization of a certain PDE. But so far a general convergence analysis for the AMLI method for a wide class of non-symmetric matrices is still missing. In this talk we give convergence results for the AMLI method applied to arbitrary non-symmetric M-matrices.
The AMLI method can be expressed as an inexact additive Schwarz method. We use this connection to Schwarz methods to introduce the second method which we analyze in this talk. It is the multiplicative counterpart of the AMLI method which we call the MAMLI method (multiplicative algebraic multilevel method).
We establish that the rate of convergence of the MAMLI method is better than the rate of convergence of the AMLI method. Furthermore we present a bound for the asymptotic convergence rate for both methods for the symmetric and for the non-symmetric case. Also different kinds of approximations to speed up the convergence rate of these methods are considered.

References

[1] O. Axelsson and M. Neytcheva, *Algebraic multilevel iteration method for Stieltjes matrices*, Numer. Linear Algebra Appl., 1 (1994), pp. 213–236.

[2] O. Axelsson and P. Vassilevski, *Algebraic multilevel preconditioning methods, I*, Numerische Mathematik, 56 (1989), pp. 157–177.

[3] ———, *Algebraic multilevel preconditioning methods, II*, SIAM Journal Numerical Analysis, 27 (1990), pp. 1569–1590.

[4] M. Neytcheva, O. Axelsson, and K. Georgiev, *An application of the AMLI method for solving convection-diffusion problems with potential velocity field*, in Algebraic multilevel iteration methods with applications, Vol. I, II (Nijmegen, 1996), Katholieke Univ. Nijmegen, Dept. Math., Nijmegen, 1996, pp. 197–210.

[5] Y. Notay, *A robust algebraic multilevel preconditioner for non-symmetric M-matrices*, Numer. Linear Algebra Appl., 7 (2000), pp. 243–267.

Estimates of the error in the preconditioned conjugate gradient algorithm

Gérard Meurant
CEA/DIF
BP12, 91680 Bruyeres le Chatel, France
gerard.meurant@cea.fr

During the last ten years there has been a renewed interest in computing bounds or estimates of norms of the errors in the Conjugate Gradient algorithm for solving linear systems with symmetric positive definite matrices, see [2], [5], [7].
The techniques used in [3], [5] are based on writing the norms of the error (A or l_2 norms) as Riemann–Stieltjes integrals and to obtain bounds or estimates of these integrals with Gauss quadrature rules. If we have estimates of the smallest and largest eigenvalues of A, we can use the Gauss–Radau and Gauss–Lobatto quadrature rules which can provide upper bounds. Such estimates can be used to devise reliable stopping criteria for stopping the iterations, particularly in problems arising from the finite element method, see [1]. It has been shown in [4], [8] that these techniques can be efficiently used in finite precision computations despite the rounding errors that delay CG convergence.
In this lecture we shall summarize these techniques and show how they can be extended to the Preconditioned Conjugate Gradient, see [6], [9]. We shall provide several numerical examples of the effectiveness of the estimates of the error norms.

References

[1] M. Arioli, *Stopping criterion for the conjugate gradient algorithm in a finite element method framework*, Numer. Math., 97 (2004), pp. 1–24.

[2] G. Golub and G. Meurant, *Matrices, moments and quadrature*, in Numerical Analysis 1993, D.F. Griffiths & G.A. Watson Eds., Pitman Research Notes in Mathematics, 303 (1994), pp. 105–156.

[3] ———, *Matrices, moments and quadrature 2 or how to compute the norm of the error in iterative methods*, BIT, 37 (1997), pp. 687–705.

[4] G. Golub and Z. Strakoš, *Estimates in quadratic formulas*, Numerical Algorithms, 8 (1994).

[5] G. Meurant, *The computation of bounds for the norm of the error in the conjugate gradient algorithm*, Numerical Algorithms, 16 (1997), pp. 77–87.

[6] ———, *Numerical experiments in computing bounds for the norm of the error in the preconditioned conjugate gradient algorithm*, Numerical Algorithms, 22 (1999), pp. 353–365.

[7] ———, *The Lanczos and Conjugate Gradient, algorithms, from theory to finite precision computations*, SIAM, 2006.

[8] Z. Strakoš and P. Tichý, *On error estimation in the conjugate gradient method and why it works in finite precision computation*, Electr. Trans. Numer. Anal., 13 (2002), pp. 56–80.

[9] ———, *Error estimation in preconditioned conjugate gradient*, accepted in BIT Numerical Mathematics, (2005).

KKT systems in constrained image registration

Jan Modersitzki
Institute of Mathematics, University of Lübeck
Wallstraße 40, Lübeck, Germany
modersitzki@math.uni-luebeck.de

Eldad Haber
Dept. of Mathematics and Computer Science, Emory University, Atlanta
400 Dowman Drive, Atlanta, G, 30322, USA
haber@mathcs.emory.edu

In this talk, we discuss KKT system arising from constrained registration problems. We focus on two challenges: non linear differential constraints and large scaled but structured problems demanding for special treatment, including discretization and multigrid techniques.
Registration is one of the fundamental image processing problems. The goal of registration (also: alignment, warping, or fusion) is to find a proper transformation of a given template image T, such that the reasonable transformed template image T is similar to a given reference image R. Particularly in medical imaging, registration is central to a comparison or integration of data obtained from different times or devices.
Registration is most commonly phrased as an optimization problem, where the goal is to minimize an application conform distance measure with respect to a continuous vector field $u = u(x)$. Since the problem is severely ill-posed, an additional regularization becomes inevitable. In contrast to many other ill-posed problems, the particular regularization is crucial to the solution. It is therefore desirable to provide as much a priori information as possible. Here, we discuss different approaches, where this additional information is supplied in terms of constraints. Particularly, we focus on volume preserving constraints. Our variational approach reads

$$D(T(u), R) + \alpha S(u) \stackrel{u}{=} \min \quad \text{s.t} \quad C(u)(x) = 0 \quad \text{for all} \quad x, \tag{1}$$

where $T(u)$ is the deformed template image, $S(u)$ is a regularizer, $\alpha > 0$ a regularization parameter and $C(u)$ are the constraints. The basis for a *discretize than optimize* approach is a proper discretization of this functional. Particularly for the pointwise and differential constraints, such a discretization is non-trivial. Our optimization scheme is based on a sequential quadratic programming, that is a quadratic approximation to the objective function and a linearization of the constraints. Using a Gauß-Newton type scheme, the computational expensive part of our algorithm is the solution of a KKT system for the updates of the displacement u and the Lagrange multiplier p,

$$\begin{pmatrix} H & C_u \\ C_u^\top & 0 \end{pmatrix} \begin{pmatrix} \delta_u \\ \delta_p \end{pmatrix} = \text{rhs}.$$

Here, H is a Gauß-Newton approximation to the Hessian of the objective function, C_u is a linearization of the constraints. Commonalities and differences with Stokes problem are discussed. The numerical solution is computationally expensive since the number of unknowns is $3N$ and the number of constraints is N, where N ranges from 30.000 to 100.000.000. Thus direct solution schemes, as they are standardly used in optimization, can not be applied. We thus comment on multigrid used as a linear solver within an overall Gauß-Newton scheme.
Finally, we comment on first results obtained for the registration of magnetic resonance scans of a female breast and positron emission tomography images from heart movement.

A rank-one variable metric approach for solving nonsymmetric linear systems of equations

Adel Mohsen
Eng. Math. Dept., Eng. Faculty, Cairo University
Giza 12211, Egypt
amhsn22@yahoo.com

Josef Stoer
Inst. Angew. Math., Univ. Würzburg
D-97074 Würzburg, Germany

A solution method for the linear system $Ax = b$ with an $m \times n$-matrix of rank r is presented. The method generates a sequence $\{H\}$ of $n \times m$ matrices so that AH are spd and $\{H\}$ leads to an approximate pseudo-inverse of A in no more than r steps.
The properties of the algorithm are established and examples comparing the method with CGN, CGS and GMRES are given.
Another algorithm using the spd matrices $V = AH$ is also proposed. This has the advantage that it can easily monitor the accuracy of the solution.

Multiple LU factorizations of a singular matrix

Juan M. Molera
Departamento de Matemáticas, Universidad Carlos III de Madrid
Avenida de la Universidad 30, 28911 Leganés, Spain
molera@math.uc3m.es

Froilán M. Dopico
Departamento de Matemáticas, Universidad Carlos III de Madrid
Avenida de la Universidad 30, 28911 Leganés, Spain
dopico@math.uc3m.es

Charles R. Johnson
Department of Mathematics, The College of William and Mary
Williamsburg, VA 23187-8795, USA
crjohnso@math.wm.edu

A singular matrix A may have more than one LU factorizations. In this work the set of all LU factorizations of A is explicitly described when the lower triangular matrix L is nonsingular. To this purpose, a canonical form of A under left multiplication by unit lower triangular matrices is introduced. This canonical form allows us to characterize the matrices that have an LU factorization and to parametrize all possible LU factorizations. Formulae in terms of quotient of minors of A are presented for the entries of this canonical form.

Upper bounds for infinity norm of the inverse of SDD and $\mathcal{S}$-SDD matrices

Nenad Moraca
Faculty of Science, Department of Mathematics and Informatics, University of Novi Sad
Trg D. Obradovica 4, 21000 Novi Sad, Serbia and Montenegro
nemoraca@gmail.com

The class of $\mathcal{S}$-SDD matrices is the class of nonsingular matrices introduced by Gao and Wang in 1992 and further investigated by Cvetkovic, Kostic and Varga in 2004.

Definition 1 *Given any matrix $A = [a_{ij}] \in \mathbf{C}^{n,n}, n \geq 2$ and given any nonempty proper subset S of N, then A is an S-strictly diagonally dominant (S-SDD) iff*

$$\begin{cases} i)\ |a_{ii}| > r_i^S(A) \text{ for all } i \in S, \\ ii)\ (|a_{ii}| - r_i^S(A))(|a_{jj}| - r_j^{\overline{S}}(A)) > r_i^{\overline{S}}(A) r_j^S(A) \text{ for all } i \in S, j \in \overline{S}. \end{cases}$$

Let $T(A)$ be the set of non SDD rows of a matrix A

$$T(A) = \{i \in N \mid |a_{ii}| \leq r_i(A)\}.$$

Proposition 1 *Let $A \in \mathbf{C}^{n,n}, n \geq 2$ and $S \in \mathcal{P}(N) \setminus \{\emptyset, N\}$. Then A is S-SDD if and only if A is $\overline{S}$-SDD.*

Proposition 2 *Let a non SDD matrix $A \in \mathbf{C}^{n,n}, n \geq 2$ be S-SDD for some nonempty proper subset S of N. Then it is either $T(A) \subseteq S$ or $T(A) \subseteq \overline{S}$.*

The main result is:

Theorem 1 *Let $A = [a_{ij}] \in \mathbf{C}^{n,n}, n \geq 2$ be S-SDD matrix for some nonempty proper subset S of N and let $\emptyset \subset T(A) \subseteq \overline{S}$. Then*

$$\|A^{-1}\|_\infty \leq \max_{i \in S, j \in \overline{S}} \frac{|a_{ii}| - r_i^S(A) + r_j^S(A)}{(|a_{ii}| - r_i^S(A))(|a_{jj}| - r_j^{\overline{S}}(A)) - r_i^{\overline{S}}(A) r_j^S(A)}.$$

Although SDD matrices are S-SDD for every nonempty proper subset S of N, the previous theorem cannot be applied for every S because $T(A) = \emptyset$.
Let $A = [a_{ij}] \in \mathbf{C}^{n,n}, n \geq 2$ be SDD matrix.

$$J(A) = \{j \in N \mid |a_{jj}| - r_j(A) = \min_{i \in N}(|a_{ii}| - r_i(A))\}.$$

We have the similar theorem for SDD matrices, just the set $T(A)$ is replaced with the set $J(A)$.
The new bound for infinity norm of the inverse of an SDD matrix is always less than Varah bound, except in the case when A is uniformly SDD ($J(A) = N$), when it is equal to Varah bound.
Varga's bound for H matrices ($\|A^{-1}\|_\infty \leq \|\mathcal{M}^{-1}(A)\|_\infty$, $\mathcal{M}(A)$ is comparison matrix), applied on $\mathcal{S}$-SDD matrices, is theoretically the best bound for matrices equimodular to A, however, our bound is more suitable for use in practical situations.
We shall also introduce iterative procedure for obtaining tighter bounds accompanied with some open questions.
Our bounds for $\|A^{-1}\|_\infty$ can be used for obtaining lower bounds for the smallest singular value, in the same way Varah used his bound.

Restarted Lanczos for nonsymmetric eigenvalue problems and linear equations

Ron Morgan
Baylor University
Dept. Mathematics, Waco, TX 76798
Ronald_Morgan@baylor.edu

First, a restarted nonsymmetric Lanczos method will be discussed. It can be used to compute both left and right eigenvectors even when storage is limited. Approximate eigenvectors are retained at the restart as with implicitly restarted Arnoldi. It uses a three-term recurrence, but some reorthogonalization is needed.
The next topic is a related method called BiCG with deflated restarting (QMR with deflated restarting may also be discussed). It simultaneously solves linear equations and computes left and right eigenvectors. For the case of multiple right-hand sides, the eigenvector information from solving the first right-hand side can help efficiently solve subsequent right-hand sides. A deflated BiCGStab can be used for this. Deflated BiCGStab has a projection over the eigenvectors followed by regular BiCGStab.

A new algorithm for frequency isolation in undamped spring-mass systems

Julio Moro
Departamento de Matemáticas, Universidad Carlos III de Madrid
Avda. Universidad 30, 28911-Leganés, Madrid (Spain)
jmoro@math.uc3m.es

Juan Egaña
Departamento de Matemáticas, Universidad Católica del Norte
Avda. Angamos 0610, Antofagasta (Chile)
jegana@ucn.cl

Given a vibrational system with eigenvalues close to resonance, the frequency isolation problem deals with redesigning the system in such a way that: i) all eigenvalues of the new system are outside the resonance region; and ii) the impact of redesign is minimized in some sense. We consider undamped spring-mass systems of the form

$$Mx'' + Kx = F(t),$$

where the mass matrix M is diagonal with positive diagonal entries, and the stiffness matrix K is symmetric tridiagonal with positive diagonal and negative off-diagonal entries. We present a new algorithm for frequency isolation, based on using directional derivatives of eigenvalues to identify directions in matrix space (M, K) in which the appropriate spectral variation is maximal. Unlike previous algorithms in the literature, which solve the problem as an inverse eigenvalue problem, followed by an optimization procedure, we take a direct approach, modifying the matrices M and K along directions of maximal variation, and employing a bisection procedure to detect the moment when all eigenvalues leave the resonance region. Using bisection keeps the computational cost low, and following the directions of maximal spectral variation leads to solutions which, in most cases, are closer to the initial values for M and K than those obtained by previous algorithms. Therefore, the cost of redesign is usually much lower, as will be shown in a number of experiments.

An eigenvalue and eigenvector problem in linear vibration systems

Peter C. Müller
Safety Control Engineering, University of Wuppertal
Gauß-Str. 20, D-42097 Wuppertal, Germany
mueller@srm.uni-wuppertal.de

Ming Hou
Department of Engineering, University of Hull
Hull, HU6 7RX, United Kingdom
m.hou@hull.ac.uk

A conservative chain-structured mass-spring vibration system with N degrees of freedom may be characterized by a diagonal mass matrix and a tri-diagonal symmetric stiffness matrix. For homogeneous chains with equal masses and equal springs the natural frequencies and the related eigen-modes have been determined explicitly, cf. e.g. [1]. Mikota [2] considered an in-homogeneous chain where the masses $m_k = m/k$ and the spring stiffnesses $c_k = (N + 1 - k)c$ for $k = 1, 2, \cdots, N$ have been applied. It was conjectured that the natural frequencies are given by $\Omega_k = k\Omega$ for $k = 1, 2, \cdots, N$, where Ω is the first natural frequency according to $\Omega^2 = c/m$. Although various aspects of this problem have been discussed recently in [3], the conjecture is still not proved.
In this contribution, the related eigenvalue and eigenvector problem is discussed in detail and a complete solution is given. Mikota's conjecture is verified. The eigenvectors are also determined analytically using binomial coefficients. Because the system matrix can be represented by the square of an anti-bidiagonal matrix, the eigenvalue problem is also solved.

References

[1] K. Klotter: Technische Schwingungslehre (in German), Vol. 2, 2nd edition. Springer, Berlin-Heidelberg 1981, Section 3.2.

[2] J. Mikota: Frequency tuning of chain structure multi-body oscillators to place the natural frequencies at Ω_1 and $N - 1$ integer multiples $\Omega_2, \ldots, \Omega_N$, Z. Angew. Math. Mech. 81 (2001), S 2, pp. S201-S202.

[3] P. C. Müller and M. Gürgöze: Natural frequencies of a multi-degree-of-freedom vibration system. Lecture, Annual Meeting of the "Gesellschaft für Angewandte Mathematik und Mechanik (GAMM)", Berlin, 27–31 March, 2006.

Lanczos methods for ill-posed problems in image processing

James Nagy
Mathematics and Computer Science Department, Emory University
Atlanta, GA, USA
nagy@mathcs.emory.edu

Julianne Chung
Mathematics and Computer Science Department, Emory University
Atlanta, GA, USA
jmchung@emory.edu

Ill-posed problems arise in many image processing applications, including microscopy, medicine and astronomy. Iterative methods are typically recommended for these large scale problems, but they can be difficult to use in practice. For example, it may be difficult to determine an appropriate stopping criteria for fast algorithms, such as the conjugate gradient method; noise contaminates the iterates very quickly, so an imprecise stopping criteria can lead to poor reconstructions. Lanczos based hybrid methods have been proposed to slow the introduction of noise in the iterates. These methods require choosing a regularization parameter for a small subproblem at each iteration. It has been shown that if these parameters are chosen optimally, then the Lanczos based methods can be very effective. In this talk we illustrate difficulties that can arise in practice when attempting to choose the regularization parameters automatically, and consider a modification of the generalized cross validation method for this purpose. Image processing examples are used to illustrate concepts and to test and compare algorithms.

New shifted algorithm with ensured convergence to singular values of bidiagonal matrices

Yoshimasa Nakamura
Graduate School of Informatics, Kyoto University, PRESTO JST
Kyoto, 606-8501, JAPAN
ynaka@amp.i.kyoto-u.ac.jp

Masashi Iwasaki
PRESTO, JST, Graduate School of Informatics, Kyoto University
Kyoto, 606-8501, JAPAN
imasa@amp.i.kyoto-u.ac.jp

A new algorithm with a shift of origin (mdLVs) for computing singular values σ_k of bidiagonal matrices is presented. A shift ${\theta^{(n)}}^2$ is introduced into the recurrence relation defined by a discrete-time integrable dynamical system as follows.

$$\bar{w}_{2k-1}^{(n+1)} = v_{2k-2}^{(n)} + v_{2k-1}^{(n)} - \bar{w}_{2k-2}^{(n+1)} - {\theta^{(n+1)}}^2, \quad \bar{w}_{2k}^{(n+1)} = \frac{v_{2k-1}^{(n)} v_{2k}^{(n)}}{\bar{w}_{2k-1}^{(n+1)}}$$

A shift strategy is given so that the singular value computation becomes numerically stable and has a cubic convergence rate and a higher relative accuracy. Namely,

$${\theta^{(n)}}^2 = \max\{0, {\vartheta_1^{(n)}}^2 - \varepsilon\}, \quad \vartheta_1^{(n)} := \min_k \left\{ \sqrt{w_{2k-1}^{(n)}} - \frac{1}{2}\left(\sqrt{w_{2k-2}^{(n)}} + \sqrt{w_{2k}^{(n)}} \right) \right\}$$

for any small positive ε. Therefore the algorithm is more accurate and faster than a credible LAPACK routine for singular values.
Secondly, a new double Cholesky factorization of symmetric tridiagonal matrices is also presented by using discrete-time integrable systems, which gives rise to a fast algorithm for the associate singular vectors.

$$B^\top B - \left(\frac{1}{\delta^{(0)}} - \frac{1}{\delta^{(\pm 1)}} \right) I = (B^{\pm})^\top B^{\pm}$$

A new bidiagonal SVD algorithm (I-SVD) is then designed. It has good properties with respect to both the computational time and the numerical accuracy. Performance of the new SVD algorithm is then discussed.
Combining the algorithm with the block Householder transform and the Murata-Horikoshi-Lang algorithm for bidiagonalization a new fast SVD algorithm for full matrices will be completed.

References

1) M. Iwasaki and Y. Nakamura, On the convergence of a solution of the discrete Lotka-Volterra system, Inverse Problems, **18**(2002), 1569–1578.

2) M. Iwasaki and Y. Nakamura, An application of the discrete Lotka-Volterra system with variable step-size to singular value computation, Inverse Problems, **20**(2004), 553–563.

3) M. Takata, M. Iwasaki, K. Kimura and Y. Nakamura, An evaluation of singular value computation by the discrete Lotka-Volterra system, Proceedings of The 2005 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA2005), Vol. IIC2005, pp. 410–416.

4) M. Takata, K. Kimura, M. Iwasaki and Y. Nakamura, Performance of a new scheme for bidiagonal singular value decomposition of large scale, Proceedings of IASTED International Conference on Parallel and Distributed Computing and Networks (PDCN 2006).

5) M. Iwasaki and Y. Nakamura, Accurate computation of singular values in terms of shifted integrable schemes, preprint, 2005.

Trading fill-out for fill-in in Jacobian accumulation

Uwe Naumann
Fachgruppe Informatik, RWTH Aachen
52056 Aachen, Germany
naumann@stce.rwth-aachen.de

Andrew Lyons
Department of Mathematics, Vanderbilt University
VU Station B Box 2895, Nashville TN 37235, USA
lyonsam@gmail.com

Ebadollah Varnik
Fachgruppe Informatik, RWTH Aachen
52056 Aachen, Germany
varnik@stce.rwth-aachen.de

The accumulation of the Jacobian matrix F' of a vector function F can be interpreted as an elimination procedure on the extended Jacobian E' as defined in [1]. The lower triangular matrix E contains the partial derivatives of all arithmetic operators and intrinsic functions used to implement F as a computer program. It is potentially very large as it may represent the entire computation performed by a numerical simulation program. Moreover, E is very sparse – a fact that we aim to exploit by building a sparse data structure that accounts for generated fill-in while making use of *fill-out*. As opposed to classical Gaussian elimination techniques nonzero elements in E will become zero during the elimination procedure. Fill-in can potentially be *absorbed* by fill-out allowing for a more compact representation, e.g., in compressed row storage.
Consider, for example, the computation of the gradient $(\frac{\partial d}{\partial a}, \frac{\partial d}{\partial b})^T$ for $c = a * b$; $d = \sin(c)$. The extended Jacobian before and after the elimination of the third row has the following formats:

$$\begin{pmatrix} -1 & & & \\ 0 & -1 & & \\ b & a & -1 & \\ 0 & 0 & \cos(c) & -1 \end{pmatrix} \quad \longrightarrow \quad \begin{pmatrix} -1 & & & \\ 0 & -1 & & \\ 0 & 0 & -1 & \\ b*\cos(c) & a*\cos(c) & 0 & -1 \end{pmatrix}$$

A single elimination step suffices to accumulate the desired gradient in the lower left corner generating the corresponding fill-in in two positions. Fill-out is generated in three positions allowing for the generated fill-in to be fully absorbed. In compressed storage only three entries need to be accounted for when exploiting fill-out. A vector of length five is required otherwise.
The minimization of the number of arithmetic operations required to accumulate F' has shown to be NP-complete only recently [4]. Another NP-complete problem arises when trying to rearrange the rows in E such that maximal use can be made of fill-out. The development and implementation of algorithms for minimizing the compressed storage while ensuring efficient data access is work in progress. We report on first theoretical and practical results showing the potential of the proposed approach. Prior related work contains [2] and [3].

References

[1] A. Griewank, *Evaluating Derivatives. Principles and Techniques of Algorithmic Differentiation*, SIAM, Apr. 2000.

[2] A. Griewank and U. Naumann, *Accumulating Jacobians as chained sparse matrix products*, Math. Prog., 3 (2003), pp. 555–571.

[3] U. Naumann, *Optimal accumulation of Jacobian matrices by elimination methods on the dual computational graph*, Math. Prog., 3 (2004), pp. 399–421.

[4] U. Naumann, *The complexity of derivative computation*, Tech. Rep. AIB-2005-15, RWTH Aachen, Aug. 2005. Submitted to Math. Prog.

Preconditioning of nonsymmetric saddle point systems as arising in modelling of visco-elastic problems

Maya Neytcheva
Department of Information Technology, Uppsala university
Box 337, 751 05 Uppsala, Sweden
Maya.Neytcheva@it.uu.se

Erik Bängtsson
Department of Information Technology, Uppsala university
Box 337, 751 05 Uppsala, Sweden
Erik.Bangtsson@it.uu.se

This presentation is concerned with numerical simulations of the so-called glacial rebound phenomenon. The problem originates from modelling the response of the solid Earth to large scale glacial advance and recession which may have provoked very large earthquakes in Northern Scandinavia. The need for such numerical simulations is due to ongoing investigations on safety assessment of radioactive waste repositories.
Within this study we use the isostatic model, based on the concept that the elevation of the Earth's surface seeks a balance between the weight of lithospheric rocks and the buoyancy of asthenospheric "fluid" (nearly-molten rock). The model describes the geophysical problem in terms of a system of partial differential equations which describe the equilibrium state of a pre-stressed visco-elastic material body, subject to surface and body forces. It includes a first-order term representing the so-called advection of pre-stress, the incorporation of which has proven to be crucial for the successful modelling of the underlying processes.
The continuous setting of the problem is to solve an integro-differential equation in a large time-space domain. This problem is then discretized using a finite element method in space and a suitable discretization in time, and gives rice to the solution of a large number of linear systems with nonsymmetric matrices of saddle point form. In the purely elastic case the resulting linear systems resemble the linearized Navier-Stokes equations but is of somewhat more general form.
We will outline the so-arising linear systems of equations, discuss possible preconditioning strategies, and will present some numerical experiments.

Approximation of matrix operators applied to a set of vectors

Jörg Niehoff
Mathematisches Institut, Lehrstuhl für Angewandte Mathematik, Heinrich-Heine-Universität Düsseldorf
Universitätsstraße 1, D-40225 Düsseldorf, Germany
niehoff@am.uni-duesseldorf.de

Marlis Hochbruck
Mathematisches Institut, Lehrstuhl für Angewandte Mathematik, Heinrich-Heine-Universität Düsseldorf
Universitätsstraße 1, D-40225 Düsseldorf, Germany
marlis@am.uni-duesseldorf.de

In this talk we consider the approximation of the product of a matrix function with a set of vectors, $f(A)b_i$, $i = 1, \ldots, s$, using Krylov subspace methods. We assume that A is a constant matrix and that the vectors b_i correspond to the evaluation of a smooth function at certain quadrature nodes. This problem arises in the implementation of exponential integrators, in which case f is related to the exponential function. We propose a projection technique and analyse its properties. Our analysis gives rise to a new practical stopping criterion for the iteration. Moreover, we present numerical examples which show a significant gain in computational time compared to a naive implementation, where a new Krylov basis is computed for each vector b_i.

References

[1] P. F. Fischer, *Projection techniques for iterative solution of $Ax = b$ with successive right-hand sides*, Comput. Methods Appl. Mech. Engtg., 163 (1998), pp. 193, 204.

[2] A.-K. Kassam and L. N. Trefethen, *Fourth-order time-stepping for stiff PDEs*, SIAM J. Sci. Comput., 26 (2005), pp. 1214–1233 (electronic).

[3] P. Lötstedt and M. Nilsson, *A minimal residual interpolation method for linear equations with multiple right-hand sides*, SIAM J. Sci. Comput., 25 (2004), pp. 2126, 2144.

Dynamical low-rank approximation

Achim Nonnenmacher
Mathematisches Institut, Universität Tübingen
Auf der Morgenstelle 10, D-72076 Tübingen, Germany
anonnenmacher@gmx.de

Othmar Koch
Institute for Analysis and Scientific Computing, Vienna University of Technology
Wiedner Hauptstraße 8–10, A-1040 Wien, Austria
othmar@othmar-koch.org

Christian Lubich
Mathematisches Institut, Universität Tübingen
Auf der Morgenstelle 10, D-72076 Tübingen, Germany
lubich@na.uni-tuebingen.de

For the low rank approximation of time-dependent data matrices and of solutions to matrix differential equations, an increment-based computational approach is proposed and analyzed. In this method, the derivative is projected onto the tangent space of the manifold of rank-r matrices at the current approximation. With an appropriate decomposition of rank-r matrices and their tangent matrices, this yields nonlinear differential equations that are well-suited for numerical integration. The error analysis compares the result with the pointwise best approximation in the Frobenius norm. It is shown that the approach gives locally quasi-optimal low rank approximations. Numerical experiments illustrate the theoretical results. Finally, we show the relationship of our low rank approximation for finite matrices with model reduction techniques in quantum dynamics.

Inner iterations in eigenvalue solvers

Yvan Notay
Université Libre de Bruxelles, Service de Métrologie Nucléaire
50 Av. F.D. Roosevelt, 1050 Brussels, Belgium
ynotay@ulb.ac.be

We consider inverse iteration-based eigensolvers, which require at each step solving an "inner" linear system. We assume that this linear system is solved by some (preconditioned) Krylov subspace method. In this framework, several approaches are possible, which differ by the linear system to be solved and/or the way the preconditioner is used. This includes methods such as inexact shift-and-invert, inexact Rayleigh quotient iteration, Jacobi-Davidson, generalized Davidson and generalized preconditioned inverse iteration. In this talk, we discuss and compare them, focusing on the evolution of the "outer" convergence (towards the desired eigenpair) according to the numerical effort spent in inner iterations. This gives some advantage to Jacobi-Davidson, and we further introduce simplified variants of this method that perform similarly while being somewhat cheaper. We also introduce a new variant of the shift-and-invert and Rayleigh quotient iteration methods, which cures their main weakness by making a cleverer use of the preconditioner. For each of these approaches, we show how the evolution of the outer convergence may be estimated with parameters readily computable during inner iterations. We also propose a simple test that allows to detect the point from where the outer convergence stagnates because one is close enough to the exact solution of the linear system. We discuss how a practical stopping test may be build on this basis, and we conclude by demonstrating some software.

Preconditioning block Toeplitz matrices

Dimitrios Noutsos
Department of Mathematics, University of Ioannina
Gr-45110 Ioannina, Greece
dnoutsos@cc.uoi.gr

Thomas Huckle
Institut für Informatik, TU München, Boltzmannstr. 3, D-85748 Garching bei München, Germany
huckle@in.tum.de

We investigate the spectral behaviour of Block Toeplitz matrices with small non-Toeplitz blocks. This also gives information on ill-conditioned Toeplitz Schur complement matrices and Toeplitz normal equations. These matrices are much harder to analyse and have a different behaviour than scalar Toeplitz matrices. Based on the connection between Toeplitz and Hankel matrices we derive some negative results on eigenvalue clustering also for preconditioned equations. Furthermore, we identify Block Toeplitz matrices that are easy to solve by the preconditioned conjugate gradient method. We derive some useful inequalities that give information on the location of the preconditioned systems.

A family of inertially arbitrary patterns that is not potentially nilpotent

Dale Olesky
Department of Computer Science
University of Victoria
P.O. Box 3055, Victoria, BC V8W 3P6, Canada
dolesky@cs.uvic.ca

In-Jae Kim, Pauline van den Driessche
Department of Mathematics and Statistics, University of Victoria

An n by n sign pattern $\mathcal{S}$ is inertially arbitrary if each ordered triple (n_1, n_2, n_3) of nonnegative integers with $n_1 + n_2 + n_3 = n$ is the inertia of some matrix in $Q(\mathcal{S})$, the sign pattern class of $\mathcal{S}$. A new family of sign patterns $\mathcal{G}_n$ $(n \geq 4)$ is presented that is proved to be inertially arbitrary for all odd $n = 2k + 1$. However, it is shown that $\mathcal{G}_{2k+1}$ is not potentially nilpotent and thus not spectrally arbitrary. To prove that $\mathcal{G}_{2k+1}$ allows each inertia with $n_3 \geq 1$, a novel method based on the Implicit Function Theorem is used, while matrices in $Q(\mathcal{G}_{2k+1})$ with inertias having $n_3 = 0$ are constructed by a recursive procedure from those of smaller order. Some properties of the coefficients of the characteristic polynomial of an arbitrary matrix having certain fixed inertias are derived, and are used to show that $\mathcal{G}_5$ and $\mathcal{G}_7$ are minimal inertially arbitrary sign patterns.

Numerical integration needs linear algebra

Alexander Ostermann
Institut für Mathematik, Universität Innsbruck
Technikerstraße 25, A-6020 Innsbruck, Austria
alexander.ostermann@uibk.ac.at

Numerical time integration of stiff differential equations cannot be realised with fully explicit schemes. This is due to well-known stability requirements. The same is true for partial differential equations of evolution type. Therefore, numerical integrators for such problems require some implicitness. As a consequence, they need (time consuming) solutions of systems of (non)linear equations. For partial differential equations, these systems can get very large so that direct solutions are usually not feasible. Iterative methods with efficient preconditioning are then inevitable.
In recent years, a new class of so-called exponential integrators has been introduced for the numerical solution of stiff problems. These integrators which are based on the exponential and related functions are able to solve stiff problems explicitly. However, they require evaluations of certain analytic functions of large matrices, involving the actual time step of the integrator. For some problems, this can be implemented cheaply by fast Fourier techniques. In other cases, Krylov subspace methods may be a good choice.
Nevertheless, many practical questions remain open. For example, let the matrix vary (even slightly) in time or suppose that the method has to change the size of the time step due to accuracy requirements. Then it is still unknown how to obtain a good update of the required matrix functions from previously computed evaluations. Appropriate answers to such questions, however, are indispensable for an efficient implementation of the methods.

The search for the nearest defective matrix

Michael L. Overton
Courant Institute of Mathematical Sciences, New York University
New York, NY 10012, USA
overton@cs.nyu.edu

Rafikul Alam, Shreemayee Bora
Indian Institute of Technology Guwahati
Guwahati 7810390, Assam, India
{rafik, shbora}@iitg.ac.in

Ralph Byers
University of Kansas
Lawrence, KS 66045, USA
byers@math.ku.edu

For a given $n \times n$ matrix A, its Wilkinson distance $w(A)$ is the distance to the nearest defective matrix, or equivalently, the nearest matrix with a multiple eigenvalue. Recent work characterizes the spectral and Frobenius norm Wilkinson distance as the supremum over all ϵ for which the ϵ-pseudospectrum of A has n distinct components. Consequently, the distance $w(A)$ may also be characterized in terms of a lowest saddle point z_w of the function $f(z) = \sigma_n(A - zI)$, where σ_n denotes smallest singular value and z ranges over the complex plane, but no algorithm guaranteed to compute z_w is known. Two cases occur: the case $\sigma_{n-1}(A - z_w I) > \sigma_n(A - z_w I)$, when f is smooth at z_w (the saddle point is smooth), and the case $\sigma_{n-2}(A - z_w I) > \sigma_{n-1}(A - z_w I) = \sigma_n(A - z_w I)$, when f is nonsmooth at z_w (the saddle point is nonsmooth). This paper describes a local iteration based on Newton's method that, for all practical purposes, converges quadratically to z_w, seamlessly handling both cases at once without discriminating between them. Geometrical and implementation issues are discussed along with several numerical examples. Various known bounds are used to obtain good starting points for the iteration, resulting in a heuristic algorithm to compute $w(A)$ and return a nearest defective matrix. A MATLAB code implementing all these features is freely available.

Structured Hölder condition number of multiple defective eigenvalues

María Peláez
Departamento de Matemática Aplicada – Universidad Carlos III de Madrid
Avenida de la Universidad 30, 28911 Leganés, Spain
mpelaez@math.uc3m.es

Julio Moro
Departamento de Matemática Aplicada – Universidad Carlos III de Madrid
Avenida de la Universidad 30, 28911 Leganés, Spain
jmoro@math.uc3m.es

Let A and B be square matrices. We are interested in the sensitivity of the perturbed eigenvalues of the matrix $A + \varepsilon B$ when the matrix B preserves the structure of A. We use a definition of Hölder condition number based on the first-order Lidskii expansions of perturbed eigenvalues. We compare the unstructured and structured Hölder condition numbers for certain classes of structured matrices, including complex symmetric, persymmetric, skew-symmetric, and skew-Hermitian matrices.

Gyroscopic stabilization and indefinite damped mechanical systems

Christian Pommer
Department of Mathematics, Technical University of Denmark
Matematiktorvet, bygning 303, DK-2800 Kgs. Lyngby, Denmark
c.pommer@mat.dtu.dk

An important issue is how to modify a given unstable matrix in such a way that the resulting matrix is stable. We investigate in general under which condition a matrix $M + A$ is stable,where M is an arbitrary matrix and $A = -A^*$ is skew-Hermitian. We show that if $\mathrm{trace}(M) > 0$ it is always possible to find a class of feasibel skew-Hermitian matrices A depending on the choise of M. The theory can be applied to dynamical systems of the form

$$\ddot{x}(t) + (\, d\, D + g\, G\,)\, \dot{x}(t) + K\, x(t) = 0 \quad , \tag{1}$$

where $G = -G^T$ is a skew symmetric gyrocopic matrix, $D = D^T$ is a symmetric indefinite damping matrix and $K = K^T > 0$ is a positive definite stiffness matrix. d and g are scaling factors used to control the stability of the system. It is quite astonnishing that when the damping matrix D is indefinite the system can under certain conditions be stable even if there are no gyroscopic forces G present The Lyapunov matrix equation is used to predict the stabilty limit for pure dissipative systems as well as for dissipative systems with gyroscopic stabilization.

Algebraic multigrid smoothing property of Kaczmarz's relaxation for general rectangular linear systems

Constantin Popa
Faculty of Mathematics and Computer Science, "Ovidius" University Constantza
Mamaia-Blvd. 124, RO-900527 Constantza, Romania
cpopa@univ-ovidius.ro

The basic concepts - *smoothing* and *approximation* properties - and corresponding results from Algebraic Multigrid (AMG, for short) have been first introduced by A. Brandt in a Technical Report (1983) at Weizmann Institute, Israel for (at least) square invertible (sparse) linear systems of the form

$$Ax^* = b. \tag{1}$$

In particular, he proved the *smoothing* property for the classical relaxation schemes Jacobi, Gauss-Seidel and Kaczmarz for the system (1). For the later, Brandt's original result is the following.

Theorem 1 *Let $e, r, \bar{e}$ be the error and residual before and the error after the following Kaczmarz relaxation step for* (1)

$$x^{k,0} = x^k,\ x^{k,i} = x^{k,i-1} - \frac{\left(x^{k,i-1}, A_i\right) - b_i}{\|A_i\|^2} A_i, i = 1, \ldots, n,\quad x^{k+1} = x^{k,n},\quad \forall k \geq 0, \tag{2}$$

respectively be defined by

$$e = x - x^*,\ r = Ae = Ax - b,\ \bar{e} = \bar{x} - x^*. \tag{3}$$

Then, the following smoothing property holds

$$\|\bar{e}\|^2 \ \leq\ \|e\|^2 - \gamma_0 \|Dr\|^2, \tag{4}$$

where $D = \mathrm{diag}(\frac{1}{\|A_1\|}, \ldots, \frac{1}{\|A_n\|})$ and γ is a positive constant depending only on the entries of A.

Note. In (2) A_i denotes the i-th row of A, whereas $(\cdot, \cdot)$ and $\|\cdot\|$ stand for the euclidean scalar product and norm, respectively.
Unfortunately, many important real world problems (as e.g. image reconstruction in Computerized Tomography) give rise at systems of the form (1) which are rectangular, rank-defficient, sparse and (usually) inconsistent. For such systems fast iterative solvers, as AMG are requested (e.g. in real medical applications). In this paper we prove *smoothing* properties as (4) for such systems. Beside the proofs, which are derived for this general case, the main difference with respect to the classical previously mentioned approach, consist in the definition of the errors and residuals from (3). In the general case, they are defined by

$$e = x - (P_{N(A)}(x^0) + x_{LS}),\ r = Ae = Ax - P_{R(A)}(b),\ \bar{e} = \bar{x} - (P_{N(A)}(x^0) + x_{LS}), \tag{5}$$

where $P_{N(A)}, P_{R(A)}$ are the orthogonal projections onto the corresponding subspaces associated to A, x_{LS} is the (unique) minimal norm solution of the (consistent or not) system of the type (1) and x^0 is the initial approximation used in the Kaczmarz-like algorithms. With these elements we first extend Brandt's result (4) for general rectangular, but consistent systems like (1). In the inconsistent case we prove, for the Kaczmarz Extended relaxation (developed by the author in a previuos paper) the following inequality of the form (4)

$$\|\bar{e}\|^2 \ \leq\ \|e\|^2 - \frac{\gamma}{2}\|Dr\|^2 + 2\|D\tilde{y}\|^2, \tag{6}$$

where $\tilde{y}$ is a vector generated during the Kaczmarz Extended algorithm, which goes to 0 as the number iterations tends to infinity.

On acceleration methods for approximating the matrix exponential

Marina Popolizio
Dipartimento di Matematica, Università di Bari
Via E. Orabona, 4, I-70125 Bari, Italy
popolizio@dm.uniba.it

Valeria Simoncini
Dipartimento di Matematica, Università di Bologna
Piazza di Porta San Donato, 5, I-40127 Bologna, Italy
valeria@dm.unibo.it

The recent interest in acceleration procedures for the numerical approximation of the matrix exponential inspired this work. By starting from the approach of Hochbruck and van den Eshof, we present new theoretical results that allow us to derive an effective strategy for the selection of the acceleration parameter.
Computational aspects will also be addressed.

References

[1] J. van den Eshof, M. Hochbruck, *Preconditioning Lanczos approximations to the matrix exponential*, to appear in SIAM J. Sci. Comp.

Preconditioners for mixed formulation of groundwater flow problems

Catherine E. Powell
University of Manchester
School of Mathematics, University of Manchester, Sackville Street, Manchester, M60 1QD, UK
c.powell@manchester.ac.uk

The simple deterministic Darcy flow problem, in which the permeability coefficients are assumed to be known explicitly, has received much attention in the literature over the last decade. Linear algebra techniques for tackling it are now mature and considered to be state of the art. Mixed finite element methods are recognised to be an invaluable tool for discretisation and give rise to symmetric and indefinite linear systems of equations. Users are faced with a plethora of solution methodologies, all of which have pros and cons in the context of solving flow problems. Transformations to positive definite systems are popular as well as augmented Lagrangian approaches. However, solving the original indefinite system using minimal residual schemes is not problematic provided care is taken in constructing a preconditioner. We review, first, some popular preconditioning schemes for solving the deterministic Darcy flow problem and discuss their robustness and computational costs. We then focus on the parameter-free solution of the full saddle-point problem using straight-forward, non-nested minimum residual iterations. Two block-diagonal preconditioners are constructed by exploiting the well-known fact that the underlying variational problem is well-posed in two distinct pairs of function spaces. The first relies on the availability of a specialised multigrid approximation to a weighted $H(div)$ operator. The second is a simpler black-box method, the key tools for which are diagonal scaling for a weighted mass matrix and an algebraic multigrid V-cycle applied to a sparse approximation to a Schur complement matrix. We end the talk with a brief look at how existing preconditioning strategies for the deterministic problem can be extended to tackle the more realistic, stochastic formulation of the Darcy flow problem.

Pseudospectra of matrix polynomials and their boundaries

Panayiotis Psarrakos
Department of Mathematics, National Technical University of Athens
ppsarr@math.ntua.gr

Joint work with Lyonell Boulton and Peter Lancaster

Consider the *matrix polynomial* $P(\lambda) = A_m\lambda^m + \cdots + A_1\lambda + A_0$, where $A_j \in \mathbb{C}^{n\times n}$ with $\det A_m \neq 0$, and λ is a complex variable. The *spectrum* of $P(\lambda)$ is defined by $\sigma(P) = \{\lambda \in \mathbb{C} : \det P(\lambda) = 0\}$. We are interested in the spectra of perturbations of $P(\lambda)$ of the form $P_\Delta(\lambda) = (A_m + \Delta_m)\lambda^m + \cdots + (A_1 + \Delta_1)\lambda + A_0 + \Delta_0$, where the matrices $\Delta_j \in \mathbb{C}^{n\times n}$ $(j = 0, 1, \ldots, m)$ are arbitrary.
For a given $\varepsilon > 0$ and a given set of nonnegative weights $\mathbf{w} = \{w_0, w_1, \ldots, w_m\}$ with at least one nonzero element, the *ε-pseudospectrum* of $P(\lambda)$ is defined (with respect to the spectral norm) by

$$\sigma_{\varepsilon,\mathbf{w}}(P) = \{\lambda \in \mathbb{C} : \det P_\Delta(\lambda) = 0,\ \|\Delta_j\|_2 \leq \varepsilon\, w_j,\ j = 0, 1, \ldots, m\}\,.$$

The parameters $w_0, w_1, \ldots, w_m$ allow freedom in how perturbations are measured. Pseudospectra provide important insights into the sensitivity of eigenvalues under perturbations and have several applications. If we denote by $s_{\min}(\cdot)$ the minimum singular value of a complex matrix and consider the polynomial $w(\lambda) = w_m\lambda^m + \cdots + w_1\lambda + w_0$, then we have that

$$\sigma_{\varepsilon,\mathbf{w}}(P) = \{\lambda \in \mathbb{C} : s_{\min}(P(\lambda)) \leq \varepsilon w(|\lambda|)\}\,.$$

Since the leading coefficient A_m is nonsingular, for sufficiently small $\varepsilon > 0$, $\sigma_{\varepsilon,\mathbf{w}}(P)$ consists of no more than nm bounded connected components, each one containing a single (possibly multiple) eigenvalue of $P(\lambda)$. On the other hand, as the parameter ε increases, $\sigma_{\varepsilon,\mathbf{w}}(P)$ enlarges, and for ε large enough, it is no longer bounded. Here, we study the boundary of $\sigma_{\varepsilon,\mathbf{w}}(P)$ and its relation with multiple eigenvalues of associated perturbations of $P(\lambda)$. Furthermore, we obtain an upper bound for the number of connected components of $\sigma_{\varepsilon,\mathbf{w}}(P)$ and a lower bound for the distance between $\sigma_{\varepsilon,\mathbf{w}}(P)$ and a given point $\mu \notin \sigma_{\varepsilon,\mathbf{w}}(P)$. Illustrative examples are also given.

Convergence issues of some classes of iterative aggregation-disaggregation methods in computing Markov chains

Ivana Pultarová
Department of Mathematics, Faculty of Civil Engineering, Czech Technical University in Prague
Thákurova 7, 166 29 Praha 6, Czech Republic
ivana@mat.fsv.cvut.cz

The iterative aggregation-disaggregation (IAD) methods provide a multilevel approach to finding a stationary probability distribution vector of a finite discrete time Markov chain represented by an irreducible stochastic matrix B. The methods consist of combining some basic (fine scale) iteration with a solution of a certain coarse scale problem [2, 3, 4].
We can use various methods as for the basic iteration, e.g. block Jacobi or block Gauss-Seidel methods [4]. In the present work we concentrate on the IAD methods where the basic iteration is represented by several steps of the power method or more generally, when the basic iteraton corresponds to the multiplication by a polynomial p in the matrix B [3]. The present work introduces some new results on convergence properties of these classes of the IAD methods.
The existence and uniqueness of the fixed point is proved for $p(t) = \alpha t + 1 - \alpha$ and $\alpha \in \langle 0, 1\rangle$. It is shown that when $p(t) = \alpha t + 1 - \alpha$ and $\alpha \in (0, 1)$ then the algorithms converges in local sense for any irreducible matrix B. In the case $p(t) = t$, the local convergence is obtained for any irreducible matrix B with a positive diagonal. In addition to it, when B contains at least one positive row, then the asymptotic convergence factor can be estimated.
New examples of divergence of the IAD methods are introduced for $p(t) = t$ or t^2. We introduce a surprising phenomenon, that for some irreducible matrices B the IAD method converges at least locally when $p(t) = t$ while it diverges even in local sense when $p(t) = t^2$. This phenomenon can be generalized and it shows limitations of the IAD methods with power iteration steps as basic iterations.
The introduced theory is based strictly on the sparsity structure of the stochastic matrix B in contrast to the quantitative analysis of the convergence properties of the IAD methods for nearly completely reducible Markov chains [4]. We use an approach similar to that described in [1].
In order to indicate the boundary line between the local convergence and divergence in local sense, some new questions and hypothesis are formulated.

References

[1] I. Ipsen and S. Kirkland, *Convergence analysis of an improved PageRank algorithm*, Technical Report, North Carolina State University, 2004.

[2] I. Marek and P. Mayer, *Convergence theory of some classes of iterative aggregation-disaggregation methods for computing stationary probability vectors of stochastic matrices*, Linear Algebra and its Application, 2003.

[3] I. Marek and I. Pultarová, *A note on local and global convergence analysis of iterative aggregation-disaggregation methods*, to be published in Linear Algebra and its Application.

[4] W. Stewart, *Introduction to the Numerical Solution of Markov Chains*, Princeton University Press, 1994.

Element-based preconditioners for problems in geomechanics

Alison Ramage
Department of Mathematics, University of Strathclyde
Livingstone Tower, 26 Richmond Street, Glasgow G1 1XH, Scotland
A.Ramage@strath.ac.uk

At the heart of a nonlinear finite element analysis for geotechnical engineering problems, in common with many other areas of structural engineering and elastostatics, lies the solution of algebraic nonlinear equilibrium equations. These equations are usually solved with well-established incremental or iterative solution techniques (such as the modified Euler or Newton-Raphson methods), reducing the problem to a sequence of linear systems involving the structure stiffness matrix and load vector which must be solved for the nodal displacements. Because of the size and sparsity of the matrices which occur in practice, solution of these systems consumes most of the computing resources required (in terms of both CPU time and memory) for a finite element analysis. Iterative solvers and preconditioners therefore play important roles. In this talk we review the use of a range of established element-based preconditioning methods for linear elastic and elasto-plastic problems and compare their performance with a new element-based method which appears to offer a significant improvement in performance.

Stochastic eigen-inference from large empirical covariance matrices

N. Raj Rao
Department of Electrical Engineering and Computer Science,
Massachusetts Institute of Technology
77 Massachusetts Avenue, Cambridge, MA 02139, USA
raj@mit.edu

Alan Edelman
Department of Mathematics,
Massachusetts Institute of Technology
77 Massachusetts Avenue, Cambridge, MA 02139, USA
edelman@math.mit.edu

Inferring the eigenvalues of the *true* covariance matrix from the eigenvalues of the *empirical* covariance matrices is an important component of many scientific and engineering applications. This is a particularly challenging problem in high-dimensional settings when there is insufficient data. This is often referred to as the curse of high dimensionality.
Faster eigenvalue computation techniques will not help because the lack of data distorts/blurs the sample eigen-spectrum – thus the eigenvalues of the empirical covariance matrix will not, in any visually discernible way, resemble the eigenvalues of the true covariance matrix. In such a scenario, one needs to exploit any hidden structure to improve estimation and reverse the distortion/blurring.
We present a stochastic eigen-inference algorithm that accomplishes precisely that. The algorithm is particularly suited for statistical applications where the sample covariance matrix may be adequately modelled as random. We present a simple algorithm that exploits the fact that as the size of the such random matrices gets large, they get increasingly non-random in a predictable manner. We turn the curse of high dimensionality into a blessing by demonstrating the relevance and practical utility of this algorithm for a broad class of matrices.

Orthogonal polynomials, quadrature, and ill-posed problems

Lothar Reichel
Department of Mathematics, Kent State University
Kent, OH 44242, USA
reichel@math.kent.edu

The solution of large-scale ill-posed problems has recently received considerable attention. Many of the available solution methods are based on Tikhonov regularization. This talk describes Tikhonov regularization methods that exploit the connection between orthogonal polynomials, Gauss quadrature, and Lanczos bidiagonalization to determine a suitable value of the regularization parameter and an associated approximate solution. Methods for unconstrained and constrained regularization problems will be discussed.

Decomposition methods for ill-posed problems

Lothar Reichel
Department of Mathematics, Kent State University
Kent, OH 44242, USA
reichel@math.kent.edu

New decomposition methods for the solution of linear discrete ill-posed problems are presented. For large-scale problems, which are solved by an iterative method, decomposition methods split the solution space into a Krylov subspace that is determined by the iterative method and an auxiliary subspace, which can be chosen to help represent pertinent features of the solution. We discuss various choices of the auxiliary space. Decomposition is well suited for use with the standard iterative methods, such as GMRES, RRGMRES, and LSQR. Small and medium sized problems are often solved by truncated singular value decomposition (TSVD) of the matrix. We also describe how TSVD can be enhanced by decomposition of the solution subspace.

Maps that preserve orthogonality on real and complex projective spaces

Leiba Rodman
College of William and Mary
Department of Mathematics, College of William and Mary, Williamsburg, VA 23187-8795, USA
lxrodm@math.wm.edu

Peter Šemrl
University of Ljubljana
Department of Mathematics, University of Ljubljana, Jadranska 19, SI-1000, Ljubljana, Slovenia
peter.semrl@fmf.uni-lj.si

Let F be the field of real numbers or the field of complex numbers and let $D, E \in F^{n\times n}$ be invertible matrices, $n \geq 3$. The matrices D and E induce indefinite inner products on F^n. We study maps on the projective space $\mathbb{P}(F^n)$ into itself that send D-orthogonal one-dimensional subspaces of F^n (elements of the projective space) to E-orthogonal one-dimensional subspaces. It is proved that under the assumption of bijectivity such a map T preserves (D, E)-orthogonality if and only if it preserves (D, E)-orthogonality in both directions. In this case it is induced by a linear or conjugate-linear transformation on F^n that is (D, E)-unitary up to a multiplicative constant. The existence of (D, E)-unitary and (D, E)-antiunitary maps is discussed. Examples are given showing the indispensability of the dimension and the bijectivity assumption.

Numerical stability of iterative methods

Miroslav Rozložník
Institute of Computer Science, Czech Academy of Sciences
Pod vodárenskou věží 2, CZ-18207 Prague 8, Czech Republic
miro@cs.cas.cz

Christopher C. Paige
School of Computer Science, McGill University
Montreal, Quebec, Canada, H3A 2A7
paige@cs.mcgill.ca

Julien Langou
The University of Tennessee, Department of Computer Science
1122 Volunteer Blvd., Knoxville, TN 37996-3450, USA
langou@cs.utk.edu

Iterative methods for linear systems are often the time-critical component in the solution of large-scale problems in computational science and engineering. In recent years a large amount of work has been devoted to Krylov subspace methods which are among the most widely used iterative schemes. Significantly less attention, however, has been paid to their numerical stability. Rounding errors occurring in finite-precision implementations of iterative methods can have two main effects on their numerical behavior. They can delay the rate of convergence given by the (theoretical) properties of a solved system and there is a limitation to the accuracy of computed approximate solution which does not decrease below a certain level (called usually maximum attainable or limiting accuracy). A comprehensive survey of error analysis for stationary iterative methods was given by Higham in his book, who addressed the problem how small can be the forward or backward error for various classical schemes. Several results for Krylov subspaces methods on both the accuracy of computed solution and the rate of convergence have been presented in last decades. Based on the pioneering papers of Paige on the (symmetric) Lanczos process, Greenbaum, Strakoš and others worked on a rounding error analysis of the CG method. These articles gave rise to the main result on the delay of convergence (due to rounding errors) which is essentially given by the rank-deficiency of the computed basis vectors. The maximal attainable accuracy of various iterative schemes with short recurrences has been analyzed by Greenbaum, van der Vorst, Strakoš, Gutknecht and others. In this contribution we will review the main results on the numerical behavior of the GMRES method, the most widely known and used representative of nonsymmetric Krylov subspace methods. This method consists of constructing the basis of associated Krylov subspace and then solving the transformed Hessenberg least squares problem at each iteration step. In the talk we analyze different computational variants of the Arnoldi process used in the orthogonalization part of GMRES, including its Householder (HH), classical (CGS) and modified (MGS) Gram-Schmidt implementation. We will examine how important is the orthogonality of computed Arnoldi vectors and to what extent its has an influence on the accuracy of different implementations of GMRES. In particular, we show that there is an important relation to the relative backward error, which gives us the link between the loss of orthogonality in the MGS Arnoldi and the convergence of GMRES. Using this result we will prove that the most usual MGS-GMRES implementation is backward stable. This theoretically justifies the observed fact that the linear independence of Arnoldi vectors in MGS-GMRES is effectively maintained until the convergence to the level of limiting accuracy. Based on a recently obtained bound for the loss of orthogonality in the CGS process, an analogous statement can be formulated also for CGS-GMRES. Presented results lead to important conclusions about the practical use of the GMRES and other iterative methods.

The work of the speaker was supported by the project 1ET400300415 within the National Program of Research "Information Society".

Challenging linear algebra problems in DFT and time-dependent DFT

Yousef Saad
University of Minnesota, dept of computer science and engineering
4-190 EECS bldg, 200 Union st. SE, Minneapolis, MN 55455, USA

The matrix problems encountered in electronic srtucture calculations can be quite challenging and require a combination of effective algorithms, start of the art computational resources, and perhaps most importantly, a good inderstanding of the underlying intrinsic linear algebra problem. In this talk, we will examine a few of these problems and will attempt to show why it is important not to just use a black-box software approach when dealing with some of these problems.
In the case of eigenvalue calculations (see also the talk by this author in the minisymposium "Eigenvalues in Computational Chemistry") it is important to de-emphasize individual eigenvectors and instead consider that the problem is to compute a large invariant subspace [3, 7]. It is also important to take into account the nonlinear (SCF) loop. In Time-Dependent Density Functional Theory, major gains in computational time can be made by focussing on tailoring an FFT-based Fast-Poisson-Solver [6, 4] for the problem. In this particular case, using planewave bases is crucial and one cannot do as well with real space techniques. Another calculation which provides a good illustration of the importance of understaning the underlying mathematical problem is when computing large dielectric matrices. While at first sight the calculation would appear to require the inversion of a large dense matrix, a closer examination reveals that the matrix to be inverted is a small rank perturbation of a diagoanal, so the inversion can be trivially carried out with a Sherman-Morrisson type formula.
Finally, we will discuss a question of paramount importance, namely that of computing charge densities without resorting to (partial) diagonalization. The charge density is the fundamental unknown on which the Kohn-Sham equations are based. It is traditionally obtained from the eigenvectors of a Hamiltonin which in turn depends on the charge density. However, the charge density can be (in principle) obtained without eigenfunctions. The problem can be viewed as that of computing the diagonal of a spectral projector, see, e.g. [5]. Some of the approaches explored in this direction will be briefly discussed [5, 1, 2, 3].

References

[1] C. Bekas, E. Kokiopoulou, and Y. Saad, *An estimator for the diagonal of a matrix*, Tech. Rep. umsi-2005-082, Minnesota Supercomputer Institute, University of Minnesota, Minneapolis, MN, 2005. submitted.

[2] ———, *Polynomial filtered Lanczos iterations with applications in density functional theory*, Tech. Rep. umsi-2005-117, Minnesota Supercomputer Institute, University of Minnesota, Minneapolis, MN, 2005. submitted.

[3] C. Bekas, Y. Saad, M. L. Tiago, and J. R. Chelikowsky, *Computing charge densities with partially reorthogonalized Lanczos*, Computer Physics Communications, (2006). To appear.

[4] W. R. Burdick, Y. Saad, L. Kronik, M. Jain, and J. Chelikowsky, *Parallel implementations of time-dependent density functional theory*, Computer Physics Communications, 156 (2003), pp. 22–42.

[5] L. O. Jay, H. Kim, Y. Saad, and J. R. Chelikowsky, *Electronic structure calculations using plane wave codes without diagonlization*, Comput. Phys. Comm., 118 (1999), pp. 21–30.

[6] E. Lorin de la Grandmaison, S. B. Gowda, Y. Saad, M. L. Tiago, and J. R. Chelikowsky, *Efficient computation of the coupling matrix in time-dependent density functional theory*, Computer Physics Communications, 167 (2005), pp. 7–22.

[7] Y. Zhou, Y. Saad, M. L. Tiago, and J. R. Chelikowsky, *Self-consistent-field calculations with Chebyshev filtered subspace iteration*, Tech. Rep. umsi-2005-yyy, Minnesota Supercomputer Institute, University of Minnesota, Minneapolis, MN, 2005.

Diagonalization algorithms in real space methods for electronic structure

Yousef Saad
University of Minnesota, dept of computer science and engineering
4-190 EECS bldg, 200 Union st. SE, Minneapolis, MN 55455, USA

Density Functional Theory (DFT) is a successful technique used to determine the electronic structure of matter which is based on a number of approximations. It converts the original n-particle problem into an effective one-electron system, resulting in a coupled one-electron Schrödinger equation and a Poisson's equation. This coupling is nonlinear and rather complex. It involves a charge density ρ which can be computed from the wavefunctions ψ_i, for all occupied states. However, the wavefunctions ψ_i are the solution of the eigenvalue problem resulting from Schrödinger's equation whose coefficients depend nonlinearly on the charge density. This gives rise to a non-linear eigenvalue problem which is solved iteratively. The challenge comes from the large number of eigenfunctions to be computed for realistic systems with, say, hundreds or thousands of electrons. We will discuss a parallel implementation a finite difference approach for this problem [7, 4, 2, 3, 6, 5] with an emphasis on diagonalization. We will illustrate the techniques with our in-house code, called PARSEC. This code has evolved over more than a decade as features were progressively added and the diagonalization routine, which accounts for the biggest part of a typical execution time, was upgraded several times. We found that it is important to consider the problem as one of computing an invariant subspace in the non-linear context of the Kohn-Sham equations [8, 1]. This viewpoint leads to considerable savings as it de-emphasizes the accurate computation of individual eigenvectors and focuses instead on the subspace which they span.

References

[1] C. Bekas, Y. Saad, M. L. Tiago, and J. R. Chelikowsky, *Computing charge densities with partially reorthogonalized Lanczos*, Computer Physics Communications, (2006). To appear.

[2] J. Chelikowsky, S. Öğüt, X. Jing, K. Wu, A. Stathopoulos, and Y. Saad, *Atomic and electronic structure of germanium clusters at finite temperature using finite difference methods*, Mat. Res. Soc. Symp. Proc., 408 (1997).

[3] J. R. Chelikowsky, L. Kronik, I. Vasiliev, M. Jain, and Y. Saad, *Using real space pseudopotentials for the electronic structure problem*, in Handbook for numerical analysis, volume X, C. L. Bris and P. G. Ciarlet, eds., Elsevier Science, 2003, pp. 613–635.

[4] J. R. Chelikowsky and Y. Saad, *Electronic Structure of Clusters and Nanocrystals*, American Scientific, 2005. In Press.

[5] J. R. Chelikowsky, N. Troullier, and Y. Saad, *The finite difference pseudo-potential method: electronic structure calculations without a basis*, Phys. Rev. Letters, 72 (1994), pp. 1240–1243.

[6] Y. Saad, A. Stathopoulos, J. Chelikowsky, K. Wu, , and S. Öğüt, *Solution of large eigenvalue problems in electronic structure calculations*, BIT, 36 (1996), pp. 563–578.

[7] A. Stathopoulos, S. Öğüt, Y. Saad, J. Chelikowsky, and H. Kim, *Parallel methods and tools for predicting materials properties*, Computing in Science and Engineering, 2 (2000), pp. 9–18.

[8] Y. Zhou, Y. Saad, M. L. Tiago, and J. R. Chelikowsky, *Self-consistent-field calculations with Chebyshev filtered subspace iteration*, Tech. Rep. umsi-2005-yyy, Minnesota Supercomputer Institute, University of Minnesota, Minneapolis, MN, 2005.

ADI shift parameter computation for large scale algebraic Riccati and Lyapunov equations arising in the LQR problem for parabolic PDEs

Jens Saak
Fakultät für Mathematik, Technische Universität Chemnitz
D-09107 Chemnitz, Germany
jens.saak@mathematik.tu-chemnitz.de

Hermann Mena
Departamento de Mathemática, Escuela Politécnica Nacional del Ecuador
Ladrón de Guevara E11-253 Quito - Ecuador
hmena@server.epn.edu.ec

Peter Benner
Fakultät für Mathematik, Technische Universität Chemnitz
D-09107 Chemnitz, Germany
benner@mathematik.tu-chemnitz.de

We study the selection of shift parameters for the alternating directions implicit (ADI) algorithm for the solution of Lyapunov and Riccati equations arising in linear quadratic regulator (LQR) problems for parabolic partial differential equations (PDEs). This leads to a rational minimax problem which has been considered by many authors. Since one needs to know the complete complex spectrum its optimal solution is not computable for the large scale systems arising from finite element discretization of PDEs. Therefore several alternatives for computing suboptimal parameters are discussed and compared for numerical examples.

Block preconditioners in PDE-constrained optimization

Ekkehard W. Sachs
Virginia Tech, ICAM and Dept. of Mathematics
Blacksburg, VA 24060, USA
and
Universität Trier, FB IV - Mathematik
D-54296 Trier, Germany
sachs@uni-trier.de

Optimization problems with partial differential equations gained considerable importance and attention in the recent years. The application span from classical optimal control to modern applications like optimal shape design, mathematical finance and image processing.
In all these problems an efficient numerical algorithm is among the top priorities in the research efforts. Careful choices of the discretization scheme from the pde point of view is important as well as special attention to the particular optimization routine which is used. In almost all approaches, however, linear systems of Karush-Kuhn-Tucker type have to be solved repeatedly.
The size of the underlying problems as discretized pde-constrained optimization problems causes the use of iterative solvers for the KKT systems. The efficiency of iterative solvers like CG or MINRES is heavily dependent upon the clever choice of preconditioners.
In this talk we emphasize the differences of KKT systems stemming from PDE-constrained optimization problems versus the saddle point problems arising in solving discretized PDEs. We review these aspects first from an infinite dimensional perspective. After that we look at discretized versions of the problems and their mesh dependence.
We concentrate on the use of block preconditioners and review several approaches in the literature. We compare some of the variants in terms of efficiency and analysis.

New results about GMRES convergence

Hassane Sadok
L.M.P.A, Université du Littoral
50 rue F. Buisson BP699, F-62228 Calais Cedex, France
sadok@lmpa.univ-littoral.fr

Mohammed Bellalij
Université de Valenciennes
Le Mont Houy, F-59313 Valenciennes Cedex, France
Mohammed.Bellalij@univ-valenciennes.fr

In this talk, we will derive new bounds for the GMRES method of Saad and Schultz, for solving linear systems. We will first give a simple formula for the norm of the residual of GMRES based only on the Krylov matrix. This formula allows us to generalize the well known result on the convergence behavior of GMRES when the matrix has a full set of eigenvectors. The explicit formula of the residual norm of the GMRES when applied to a normal matrix, which is a rational function, is given in terms of eigenvalues and of the components of the eigenvector decomposition of the initial residual. By minimizing this rational function over a convex subset, we obtain the sharp bound of the residual norm of the GMRES method applied to normal matrices, even if the spectrum contains complex eigenvalues. Hence we completely characterize the worst case GMRES for normal matrices. We use techniques from constrained optimization rather than solving the classical min-max problem (problem in polynomial approximation theory).

Order reduction of second order systems with proportional damping

Behnam Salimbahrami
Lehrstuhl für Regelungstechnik, Technische Universität München
Boltzmannstr. 15, D-85748 Garching bei München, Germany
salimbahrami@tum.de

Boris Lohmann and Rudy Eid
Lehrstuhl für Regelungstechnik, Technische Universität München, Garching bei München, Germany
{eid,lohmann}@tum.de

Jan G. Korvink and Evgenii B. Rudnyi
Institute for Microsystem Technology, Albert Ludwig University, Freiburg, Germany
{rudnyi,korvink}@imtek.uni-freiburg.de

A common result of modeling (for instance by finite element methods) in some fields like Electrical Circuits and Micro-Electro-Mechanical Systems (MEMS) is a large number of *second* order differential equations. The most practical solution to handle such large scale models is offered by order reduction. In reduced order modeling of such systems, not only the behavior of the original system should be approximated in the frequency range of interest, but also it is advisable to preserve the structure of the original model. We consider the second order system of the form,

$$\begin{cases} \mathbf{M}\ddot{\mathbf{z}}(t) + \mathbf{D}\dot{\mathbf{z}}(t) + \mathbf{K}\mathbf{z}(t) = \mathbf{G}\mathbf{u}(t), \\ \mathbf{y}(t) = \mathbf{L}\mathbf{z}(t). \end{cases} \tag{1}$$

The matrices $\mathbf{M}$, $\mathbf{D}$ and $\mathbf{K}$ are called mass, damping and stiffness matrices, respectively.
In [3] and others, the Krylov subspaces were extended to the so called *Second Order Krylov Subspaces* to reduce second order systems by applying a projection directly to the original model and matching the moments.
In many second order systems, the damping matrix is modeled as a linear combination of the mass and stiffness matrices ($\mathbf{D} = \alpha\mathbf{M} + \beta\mathbf{K}$) which is known as proportional or Rayleigh damping. For this special set of second order systems, the reduced order modeling using a second order Krylov method may be simplified. In [2, 1], it is proposed to calculate the projection matrices by neglecting the damping matrix, and then apply the projection to the system with damping. The main problem in these references is that the moment matching property is only proved for the case that $\mathbf{D} = \mathbf{0}$ and the quality of the damped reduced systems is only demonstrated empirically through examples and not based on moment matching.
In this talk, we first show that for systems with a proportional damping matrix, the damping matrix does not contribute to the projection matrices. It is proved that, the corresponding subspaces are equal to some standard Krylov subspaces, independent of the values of α and β and consequently independent of $\mathbf{D}$.
This fact simplifies the reduction procedure and allow us to use the standard numerical algorithms implemented for moment matching in state space that are generally more robust than the second order Krylov algorithms.
It should be noted in order reduction of proportionally damped systems by applying a direct projection, the property of proportional damping is preserved in the reduced system. In other words, if $\mathbf{D} = \alpha\mathbf{M}+\beta\mathbf{K}$ then $\mathbf{D}_r = \alpha\mathbf{M}_r + \beta\mathbf{K}_r$ in the reduced model.

References

[1] J. S. Han, E. B. Rudnyi, and J. G. Korvink, *Efficient optimization of transient dynamic problems in mems devices using model order reduction*, J. Micromech. Microeng., 15 (2005), pp. 822–832.

[2] E. B. Rudnyi, J. Lienemann, A. Greiner, and J. G. Korvink, *mor4ansys: Generating Compact Models Directly from ANSYS Models*, in Proc. NSTI-Nanotech, vol. 2, Boston, USA, 2004, pp. 279–282.

[3] B. Salimbahrami and B. Lohmann, *Order Reduction of Large Scale Second Order Systems Using Krylov Subspace Methods*, Linear Algebra and its Appl., in press, (submitted 2003).

A comparison between different bases in image reconstruction from nonuniform samples

Reginaldo J. Santos
Departamento de Matemática, Universidade Federal de Minas Gerais
Av Antonio Carlos, 6627 30123-900 Belo Horizonte-MG, Brazil
regi@mat.ufmg.br

Alvaro R. De Pierro
Applied Mathematics Department-IMECC, University of Campinas
Campinas-SP Brazil
alvaro@ime.unicamp.br

In [2] Gröchenig and Strohmer used Fourier bases to reconstruct images from nonuniform samples. They suggested "Other reconstruction methods, for instance, methods based on local Fourier bases or wavelets, have to be developed to handle the restoration of large areas of missing pixels."
We present two methods based on the application of Conjugate Gradients using D6 wavelet and Uniform Splines bases. We give an efficient way of computing the products Ax and A^*y for each basis.
Our method is different from that one presented in [1] for Splines. They used a multigrid algorithm that is formed by using classical iterators (damped-Jacobi or Gauss-Seidel) as building blocks.
We present an experimental comparison of image reconstruction using Fourier, Wavelet and Uniform Splines bases from nonuniform samples.

References

[1] M. Arigovindan, M. Sühling, P. Hunziker, and M. Unser, *Variational image reconstruction from arbitrarily spaced samples: A fast multiresolution spline solution*, IEEE Transactions on Image Processing, 14 (2005), pp. 450–460.

[2] K. Gröchenig and T. Strohmer, *Numerical and theoretical aspects of non-uniform sampling of band-limited images*, in Nonuniform Sampling: Theory and Practice, F. Marvasti, ed., Plenum Pr, New York, 2001.

Fast Runge-Kutta approximation of inhomogeneous parabolic differential equations

Achim Schädle
Department Numerical Analysis and Modelling
Konrad-Zuse-Zentrum für Informationstechnik Berlin
Takustr. 7, D-14195 Berlin-Dahlem, Germany
schaedle@zib.de

The result of N steps of an implicit Runge-Kutta time discretization of an inhomogeneous linear parabolic differential equation is computed, up to accuracy ε, by solving only

$$\mathcal{O}\left(\log(N \log \frac{1}{\varepsilon}\right)$$

linear systems of equations. In this talk the algorithm will be derived and analysed.

Combinatorial approaches to the solution of saddle-point problems in large-scale interior-point optimization

Olaf Schenk
Department of Computer Science, University of Basel
Klingelbergstr. 50, 4056 Basel, Switzerland
olaf.schenk@unibas.ch

Interior-point methods are among the most efficient approaches for solving large-scale nonlinear programming problems. At the core of these methods, highly ill-conditioned symmetric saddle-point problems have to be solved. We present combinatorial methods to preprocess these matrices in order to establish more favorable numerical properties for the subsequent solution process. Our approach is based on symmetric weighted matchings and is used in direct and incomplete factorization methods [1, 2] where the pivoting is restricted to static data structures. This technique can be seen as an alternative to the more traditional threshold pivoting techniques. We demonstrate the competitiveness of this approach within an interior-point method on a large set of test problems from the CUTE and COPS sets, as well as large optimal control problems based on partial differential equations. The largest nonlinear optimization problem solved has more than 12 million variables and 6 million constraints [3].

References

[1] M. Hagemann and O. Schenk, *Weighted matchings for the preconditioning of symmetric indefinite linear systems*, SIAM Journal of Scientific Computing. To appear.

[2] O. Schenk and K. Gärtner, *On fast factorization pivoting methods for sparse symmetric indefinite systems*, Technical Report CS-2004-012, Department of Computer Science, University of Basel, 2004. Submitted.

[3] O. Schenk, A. Wächter, and M. Hagemann, *Combinatorial approaches to the solution of saddle-point problems in large-scale parallel interior-point optimization*, Technical Report CS-2005-009, Department of Computer Science, University of Basel, 2005. Submitted.

The fast evaluation of φ-functions for exponential integrators for stiff PDEs

Thomas Schmelzer
Oxford University Computing Laboratory
Wolfson Building, Parks Road, Oxford, OX1 3QD, England
thoms@comlab.ox.ac.uk

Lloyd N. Trefethen
Oxford University Computing Laboratory
Wolfson Building, Parks Road, Oxford, OX1 3QD, England
lnt@comlab.ox.ac.uk

The fast and stable evaluation of the matrix exponential and closely related φ-functions makes exponential integrators attractive for solving stiff parabolic PDEs. Using rational uniform approximations we evaluate those functions on the negative real line - a case of particular interest in practice. By solving as few as three shifted linear systems with a direct or iterative method we can get as much as six digits of accuracy. The efficient and fast construction of the rational function is done by a Carathéodory-Féjer approximation following work by Trefethen and Gutknecht. Based on those results we discuss an efficient implementation of exponential integrators and compare its performance with more standard schemes such as linearly implicit Runge-Kutta methods. We discuss a nonlinear heat equation.

Application and extension of generalized partial Procrustes analysis for the visualization of molecular conformations

Johannes Schmidt-Ehrenberg
Visualisierung und Datenanalyse – Zuse Institut Berlin (ZIB)
Takustr. 7, D-14195 Berlin, Germany
schmidt-ehrenberg@zib.de

Molecular conformation analysis approximates probability densities in a molecule's position space by sets of molecular geometries. The visualization of these sets implies the problem of finding an optimal superposition by applying rigid transformations to the single geometries.
Assuming a set of N geometries of a molecule with M atoms, we describe a single geometry by a $3 \times M$ coordinate matrix X_i $(1 \leq i \leq N)$. A rigid transformation can be expressed as a left multiplication of a 3×3 rotation matrix R_i and the addition of a translation vector t_i to every column of the matrix:

$$X_i \longmapsto X_i^* = R_i X_i + t_i \, 1_M^{\mathsf{T}}, \qquad R_i \in SO(3), t_i \in \mathbb{R}^3,$$

where 1_M is the constant M-dimensional vector with component value 1.
Generalized partial Procrustes analysis (GPPA) defines optimal superposition by the minimal sum of squared component differences between all pairs of transformed geometries

$$O_{GPPA} = \sum_{1 \leq i < j \leq N} \operatorname{tr} \left[\left(X_i^* - X_j^* \right) \left(X_i^* - X_j^* \right)^{\mathsf{T}} \right]$$

and specifies an iterative algorithm for the determination of R_i, t_i $(1 \leq i \leq N)$ minimizing O_{GPPA}. The objective function O_{GPPA} is based on the model assumption that the molecular geometries are scattered around one central form, i.e. all geometries have approximately the same form. However, molecular conformation analysis has to deal with geometries of substantially differing forms. In this case the results of GPPA are not suitable for visualization.
The talk will explain GPPA, its relevance for the visualization of molecular conformations and the problems of their combination. Further, an extension of GPPA will be presented that considers previously gathered knowledge about geometrical differences between the metastable conformations of a molecule.

The linear algebra world of 50 years ago

Hans Schneider
Mathematics Department, University of Wisconsin
480 Lincoln Drive, Madison, WI 53706-1313, USA
hans@math.wisc.edu

I present some egocentric and autobiographical remarks on the travails and rewards of an old fashioned linear algebraist in the age of Bourbaki.

Krylov subspace methods for computing stationary probability distributions of CTMCs

Olaf Schneider
Institut für Wissenschaftliches Rechnen (IWR), Forschungszentrum Karlsruhe
Hermann-von-Helmholtz-Platz 1, D-76344 Eggenstein-Leopoldshafen, Germany
olaf.schneider@iwr.fzk.de

The computation of the stationary probability distribution of a continuous time Markov chain (CTMC) requires the solution of a homogeneous system of linear equations $A\boldsymbol{x} = \boldsymbol{b}$ whose underlying matrix A is singular. Several Krylov subspace methods, like QMR and GMRES, have been proposed for the solution of this equations (cf. [3, 4, 2]).
The mathematical analysis of these methods (cf. [1] and the references therein) has shown that it is advantageous to regard abstract approximations based on minimal residuals (MR) and orthogonal residuals (OR). We present a concise overview of the theory of iterative MR and OR subspace correction methods applied to singular systems, specialize this approach to Krylov subspaces, and explain why the problems originated from Markov chains are well suited for restarted Krylov methods.
The investigation of abstract MR and OR methods requires the distinction of regular and singular breakdowns. A regular breakdown, which is the common termination behavior of the method for a nonsingular operator, may also occur if A is not invertible but the linear system of equations is still consistent. Singular breakdowns are an exceptional behavior which can be only observed for singular operators with certain right-hand sides. If a singular breakdown occurs the MR approximation is not uniquely determined and the OR approximation fails to exist. The specialization to Krylov subspaces allows to decide a-priori which kind of termination behavior is expected. We regard the decomposition of the initial residual $\boldsymbol{r}_0$ into the components in the range and the nullspace of A^d, where d denotes the index of the operator A, i. e. the dimension of the largest Jordan block corresponding to the eigenvalue 0. The method breaks down singularly if and only if $\boldsymbol{r}_0 \notin \mathcal{R}(A^d)$.
The analysis of the underlying transition semigroup of a CTMC yields that the infinitesimal generator is a singular M-matrix with index 1. For an initial approximation $\boldsymbol{x}_0$ chosen such that $\boldsymbol{x}_0 \notin \mathcal{R}(A)$ a Krylov method converges to a nontrivial stationary distribution without singular breakdowns. This is an application of our more general result: A Krylov method yields a solution of the linear equation for any initial approximation if and only if the index of A is not greater than 1 and the system of equations is consistent.

References

[1] M. Eiermann and O. G. Ernst, *Geometric aspects of the theory of Krylov subspace methods*, Acta Numerica, (2001), pp. 251–312.

[2] R. W. Freund and M. Hochbruck, *On the use of two QMR algorithms for solving singular systems and applications in Markov chain modeling*, Numerical Linear Algebra with Applications, 1 (1994), pp. 403–420.

[3] Y. Saad, *Projection methods for the numerical solution of Markov chain models.*, in Numerical Solution of Markov Chains, W. J. Stewart, ed., Marcel Dekker, New York, 1991, pp. 455–471.

[4] ———, *Preconditioned Krylov subspace methods for the numerical solution of Markov chains*, in Computations with Markov Chains, W. J. Stewart, ed., Kluwer Academic, Dordrecht, 1995, pp. 49–64.

Structured Kronecker forms for palindromic matrix pencils

Christian Schröder
Technische Universität Berlin
Straße des 17. Juni 136, D-10623 Berlin, Germany
schroed@math.tu-berlin.de

Consider the so-called *palindromic eigenvalue problem*

$$Ax = \lambda A^* x \tag{1}$$

for a given square matrix $A \in \mathbb{C}^{n,n}$. The goal is to derive a canonical form under those transformations that preserve all eigenvalues as well as the palindromic structure. Congruence transformations satisfy both conditions, as

$$P^*(A - \lambda A^*)P\tilde{x} = 0 \;\Leftrightarrow\; \big(P^*AP - (P^*AP)^*\big)\tilde{x} = 0.$$

We present a canonical form under congruence for a square matrix A that reveals all spectral infomation of the problem (1), thus induces a structured Kronecker form for palindromic matrix pencils.
Using this canonical form we establish a relation between palindromic matrix pencils and symplectic matrices. The latter play a dominant role in the solution of discrete time optimal control problems.

The use of total least squares data fitting in the shape from moments problem

Mieke Schuermans
Katholieke Universiteit Leuven, ESAT-SCD
Kasteelpark Arenberg 10, 3001 Leuven, Belgium
mieke.schuermans@esat.kuleuven.be

Philippe Lemmerling, Sabine Van Huffel
Katholieke Universiteit Leuven, ESAT-SCD
Kasteelpark Arenberg 10, 3001 Leuven, Belgium
philippe.lemmerling,sabine.vanhuffel@esat.kuleuven.be

Lieven De Lathauwer
ETIS (CNRS, ENSEA, UCP)
UMR 8051, Cergy-Pontoise, France
delathau@ensea.fr

In this talk, we discuss the problem of recovering the vertices of a planar polygon from its measured complex moments. Because the given, measured moments can be noisy, the recovered vertices are only estimates of the true ones. The literature offers many algorithms for solving such an estimation problem. We will restrict our discussion to the Total Least Squares (TLS) data fitting models Hankel TLS (HTLS) and Structured TLS (STLS), the matrix pencil method Generalized Pencil of Function (GPOF) and a tensor-based algorithm. We show the close link between the HTLS and the GPOF methods and use the HTLS method to compute starting values for the STLS method. We compare the statistical accuracy of these four methods on simulated data. The data matrix formed by the given sequence of complex moments will be Hankel structured. We will show that the vertices estimated via the STLS algorithm are more accurate than the ones estimated via the pencil method because the STLS method preserves the Hankel structure of the matrix. Through experiments it will become clear that the parameter accuracy may improve by arranging the data sequence in a higher-order tensor and estimating the model parameters via a multilinear generalization of the SVD.

The local C-numerical range and its relation to gradient flows in quantum information

Thomas Schulte-Herbrüggen
Department of Chemistry, Technical University Munich, TUM
Lichtenbergstr. 4, D-85747 Garching, Germany
tosh@ch.tum.de

Gunther Dirr and Uwe Helmke
Mathematical Institute, University of Würzburg
Am Hubland, D-97074 Würzburg, Germany
{dirr,helmke}@mathematik.uni-wuerzburg.de

Steffen J. Glaser
Department of Chemistry, Technical University of Munich
Lichtenbergstrasse 4, D-85747 Garching, Germany
glaser@ch.tum.de

The task of maximising the trace function $\mathrm{Re}\,\mathrm{tr}\{C^\dagger UAU^\dagger\}$ over the entire unitary group relates to the C-numerical range of A. In view of quantum computation with N qubits this means $U \in SU(2^N)$.
Here we restrict the focus to the N-fold tensor product $U \in SU(2) \otimes \cdots \otimes SU(2)$ corresponding to what is known as the subgroup of *local* operations in quantum information. This translates into a highly non-convex optimization problem. In order to study its geometry, we introduce a new entity, the *local* C-numerical range $W_{\mathrm{loc}}(C, A)$. We present examples illustrating its intricate geometry, which is neither star-shaped nor simply connected in contrast to the conventional C-numerical range.
The equivalence of finding bounds on the size of $W_{\mathrm{loc}}(C, A)$ and maximizing the above trace function is established. We then describe *(i)* gradient flow and *(ii)* Jacobi-type methods to tackle the optimization task. Explicit step-size selections ensuring convergence are given.
We present a number of numerical examples and relate them to solving current problems in quantum information processing.

A half-step Jacobi–Davidson-like update formula of order $(3+\sqrt{5})/2 \approx 2.61$ for nonnormal matrices

Hubert Schwetlick
Institut für Numerische Mathematik, Technische Universität Dresden
D-01062 Dresden, Germany
schwetlick@math.tu-dresden.de

Kathrin Schreiber
Institut für Numerische Mathematik, Technische Universität Dresden
D-01062 Dresden, Germany
schreib@math.tu-dresden.de

In [1], a Generalized Rayleigh Quotient Iteration (GRQI) has been proposed for computing a simple eigenvalue λ_* of a nonnormal matrix $A \in \mathbb{C}^{n\times n}$ and the corresponding right and left normalized eigenvectors x_* and y_*, resp.. This GRQI can be reformulated as follows, cf. [2]: Given the current generalized Ritz triplet $(u_k, v_k, \theta_k) = (u, v, \theta)$ with $u^H u = v^H v = 1$ and $\theta = v^H Au / v^H u$ where $u \approx x_*$, $v \approx y_*$, the subsequent triplet $(u_{k+1}, v_{k+1}, \theta_{k+1}) = (u_+, v_+, \theta_+)$ is defined by

$$\begin{aligned}
&\texttt{Step 1.1:}\ \text{Solve}\quad (I - vv^H)(A-\theta I)(I-uu^H)s = -(A-\theta I)u \quad \text{for}\quad s \perp u\\
&\texttt{Step 1.2:}\ \text{Set}\quad u_+ = (u+s)/\|u+s\|\\
&\texttt{Step 2.1:}\ \text{Solve}\quad (I - uu^H)(A-\theta I)^H(I-vv^H)t = -(A-\theta I)^H v \quad \text{for}\quad t \perp v\\
&\texttt{Step 2.2:}\ \text{Set}\quad v_+ = (v+t)/\|v+t\|, \quad \theta_+ = v_+^H A u_+ / v_+^H u_+
\end{aligned}$$

Now we modify GRQI by taking always the newest available approximations to x_*, y_*, λ_* as follows:

$$\begin{aligned}
&\texttt{Step 1.1':}\ \text{Solve}\quad (I - vv^H)(A-\theta I)(I-uu^H)s = -(A-\theta I)u \quad \text{for}\quad s \perp u\\
&\texttt{Step 1.2':}\ \text{Set}\quad u_+ = (u+s)/\|u+s\|, \quad \theta_{1/2} = v^H Au_+ / v^H u_+\\
&\texttt{Step 2.1':}\ \text{Solve}\quad (I - u_+u_+^H)(A-\theta_{1/2} I)^H(I-vv^H)\,t = -(A-\theta_{1/2} I)^H v \quad \text{for}\quad t \perp v\\
&\texttt{Step 2.2':}\ \text{Set}\quad v_+ = (v+t)/\|v+t\|, \quad \theta_+ = v_+^H A u_+ / v_+^H u_+
\end{aligned}$$

Define $\xi_k = \angle(\operatorname{span}\{u_k\}, \operatorname{span}\{x^*\})$, $\eta_k = \angle(\operatorname{span}\{v_k\}, \operatorname{span}\{y^*\})$. Then, if $\xi_0, \eta_0 \le \varepsilon_0$ with $\varepsilon_0 > 0$ sufficiently small, the half-step method is well defined, and one has

$$\sin\xi_{k+1} \le C_1 \sin\xi_k \sin\eta_k, \quad \sin\eta_{k+1} \le C_2 \sin\xi_k \sin^2\eta_k,$$

with constants $C_1, C_2 > 0$. Hence, we have $\sin\xi_k \to 0$ Q-superlinearly and $\sin\eta_k \to 0$ Q-superquadratically. Moreover, the above estimates imply that $\sin\xi_k, \sin\eta_k \to 0$ with R-order $(3+\sqrt{5})/2 = \varrho\left(\begin{smallmatrix}1 & 1\\ 1 & 2\end{smallmatrix}\right) \approx 2.61$. The convergence of $\{\theta_k\}$ is characterized by $|\,\theta_{k+1} - \lambda^*| \le C_3 |\,\theta_k - \lambda^*|^2 \sin\eta_k$, i.e., $\theta_k \to \lambda^*$ even Q-superquadratically.
Note that for the original GRQI one has only $\sin\xi_{k+1} \le C_4 \sin\xi_k \sin\eta_k$, $\sin\eta_{k+1} \le C_4 \sin\xi_k \sin\eta_k$, i. e., (alternating) Q-quadratic convergence, and $|\theta_{k+1} - \lambda^*| \le C|\theta_k - \lambda^*|^2$, see [2].
The original and modified GRQI updates have some advantages over the the standard Jacobi-Davidson update in case of strongly nonnormal matrices with almost orthogonal x_* and y_*, i.e., in case of very sensitive eigenvalues, cf. [2] for a comparison of asymptotic condition numbers and numerical examples.
When solving the projected systems by `gmres` with `ilu` preconditioning, the transposed of the preconditioner used in Step 1.1' may be used for solving the system in Step 2.1', too.
The new updates has also successfully been used with Jacobi-Davidson-style trial space expansion.

References

[1] H. Schwetlick and R. Lösche, *A generalized Rayleigh quotient iteration for computing simple eigenvalues of nonnormal matrices*, Z. Angew. Math. Mech., 80 (2000), pp. 9–25.

[2] H. Schwetlick and K. Schreiber, *Generalized Rayleigh-Quotient-Iteration revisited: Jacobi-Davidson-like reformulation and asymptotic condition numbers*. Manuscript, in preparation for publication, 2006.

Refinement algorithms for bandwidth reduction

Jennifer Scott
Rutherford Appleton Laboratory
Chilton, Didcot, Oxfordshire OX11 0QX, UK
j.a.scott@rl.ac.uk

John Reid
Rutherford Appleton Laboratory
Chilton, Didcot, Oxfordshire OX11 0QX, UK
j.k.reid@rl.ac.uk

Since the 1960s, a large number of algorithms for reducing the bandwidth of sparse symmetric matrices have been proposed. The most well-known and most widely used is that of Cuthill and McKee [1]. Much less attention has been paid to the unsymmetric case. Recently, we have considered how variants of the Cuthill-McKee algorithm can be used to reduce the bandwidth of sparse matrices with an unsymmetric sparsity pattern [3].
In this talk, we review the hill-climbing algorithm of Lim, Rodrigues and Xiao [2]. This is a local refinement algorithm for reducing the bandwidth of symmetric matrices. We propose a simple extension to the unsymmetric case. We also adapt the node centroid algorithm of Lim *et. al.* for unsymmeric matrices. Using a range of practical problems, we illustrate how using these refinement techniques in combination with our unsymmetric Cuthill-McKee algorithms can give significant further bandwidth reductions.

References

[1] E. Cuthill and J. McKee, *Reducing the bandwidth of sparse symmetric matrices*, in Proceedings of the 24th National Conference of the ACM, Brandon Systems Press, 1969.

[2] A. Lim, B. Rodrigues, and F. Xiao, *A centroid-based approach to solve the bandwidth minimization problem*, Proceedings of the 37th Hawaii international conference on system sciences, IEEE, (2004).

[3] J. Reid and J. Scott, *Reducing the total bandwidth of a sparse unsymmetric matrix*, Technical Report RAL-TR-2005-001, Rutherford Appleton Laboratory, 2005.

Eigendecompositions for low-rank updates of the stochastic Google matrix

Teresa Selee
North Carolina State University, Department of Mathematics
255 Harrelson Hall, Raleigh, NC 27695, USA
tmselee@ncsu.edu

Ilse Ipsen
North Carolina State University, Department of Mathematics
255 Harrelson Hall, Raleigh, NC 27695, USA
ipsen@ncsu.edu

The search engine Google determines the order in which to display the results of a websearch, to a large extent, from the so-called PageRank vector. The PageRank vector contains a ranking for every web page; this ranking reflects the importance of the web page. The PageRank vector is computed as the stationary distribution of an extremely large stochastic matrix, the Google matrix.
The Google matrix G is a convex combination of two stochastic matrices, $G = \alpha(H + E_1) + (1 - \alpha)E_2$. The matrix H models the webgraph and contains a row and a column for every webpage. Web pages that have no outgoing links, such as pdf files, correspond to zero rows in H. One adds a rank-one matrix E_1, so that $H + E_1$ is a stochastic (possibly reducible) matrix. The matrix E_2 also has rank one and is used to manually adjust the PageRank. The PageRank vector contains, for every webpage, the probability that a surfer arrives at that page. The PageRank vector can be expressed in terms of E_2 and an eigendecomposition of $H + E_1$.
In order to model surfing behavior more realistically, we allow matrices E_1 and E_2 to be of higher rank. We will determine expressions for the PageRank vector in terms of E_1, E_2 and an eigendecomposition of H. The resulting expressions will show the effect of different choices of E_1 and E_2 on the PageRank vector.
We begin with a discussion on finding the diagonal and Jordan forms for general rank-one matrices using low-rank updates to the identity matrix. We then extend this to higher-rank matrices, and finally specialize this to the Google matrix.

The subspace projected approximate matrix (SPAM) iterative diagonalization method

Ron Shepard
Theoretical Chemistry Group, Chemistry Division, Argonne National Laboratory
Argonne, IL 60439 USA
shepard@tcg.anl.gov

Subspace methods for the solution of the real symmetric matrix eigenproblem ($\mathbf{Hv} = \lambda\mathbf{v}$) are widely used in chemistry applications, from electronic structure methods to the study of molecular vibrations. The Davidson, Jacobi-Davidson, and Lanczos methods are among the most familiar methods of this type. In a subspace method, the eigenvector of interest is expanded in some subspace of basis vectors. The approximate eigenvector and eigenvalue are computed within this subspace, a new correction vector is determined in some manner, that new vector is appended to the subspace, and the process is continued until sufficient accuracy is attained. This general subspace approach is applicable to the lowest or highest eigenpair, to the lowest few or the highest few eigenpairs, and to interior eigenpairs. This general approach requires only matrix-vector products and scalar products involving the basis vectors, so it does not suffer from complications of loss of sparsness due to fill-in, and the individual matrix elements may be accessed in any convenient order. In many chemistry applications, the individual matrix elements are not computed and stored at all; the necessary matrix-vector products are computed in some "operator" form from simpler underlying data structures. In some important chemistry applications the matrix dimensions exceed 10^9, and these iterative methods are the only practical approach. The main disadvantage of the subspace approach is that each matrix-vector product can be very expensive (i.e. measured in terms of wall time, arithmetic effort, internal and external storage requirements, or I/O or communication requirments). Particularly when slow convergence is observed, it is of great benefit to reduce the number of these expensive products that must be computed.

The Subspace Projected Approximate Matrix (SPAM) method (see *J. Comput. Physics* **172**, 472–514 (2001)) addresses this issue of computational cost. An approximation $\mathbf{H}^{(1)}$ to the matrix $\mathbf{H}^{(0)} \equiv \mathbf{H}$ of interest is used to accelerate convergence of the original subspace problem. There are several possible recursive and iterative implementations of this method. In our implementation we have chosen an iterative approach in which the contributions from exact and approximate matrix-vector products are combined in order to produce a sequence of intermediate approximations to the eigenpair(s) of interest. Computational efficiency is improved when the cost of an approximate product $\mathbf{w}^{(1)} = \mathbf{H}^{(1)}\mathbf{x}$ is less than the cost of the corresponding exact product $\mathbf{w}^{(0)} = \mathbf{H}^{(0)}\mathbf{x}$. In effect, the goal of this approach is to substitute cheap products for expensive products, and to substitute linear combinations of several basis vectors associated with these cheap products with individual basis vectors associated with the expensive products. In principle, if $\mathbf{H}^{(1)} = \mathbf{H}^{(0)}$, then convergence will be achieved with a single exact product; in practice, if $\mathbf{H}^{(1)}$ is a sufficiently good approximation, then convergence can be achieved with as few as one or two exact products. If an approximate product involving $\mathbf{H}^{(1)}$ is still relatively expensive, then additional even-cheaper and less accurate approximations may be introduced ($\mathbf{H}^{(2)}$, $\mathbf{H}^{(3)}$,...) with the goal of reducing the number of $\mathbf{w}^{(1)}$ products that must be computed. This leads to a multilevel SPAM approach in which sets of products from all levels of approximation are mixed together to produce the sequence of approximate eigenpairs. The matrix approximation sequence may be chosen based on numerical thresholds, screening, operator approximation, tensor-product approximation, basis approximation, expansion truncation, finite-difference order, or any other physical- or numerical-based approach that is appropriate for the target application. That is, unlike a traditional style black-box method, the SPAM approach allows the scientist or engineer to use his knowledge and experience within his application area to accelerate the convergence of his numerical eigenproblem.

This talk discusses the SPAM method, some chemical applications that use the method, some outstanding problems with the method, and some possible directions for future work.

This work was supported by the U.S. Department of Energy by the Office of Basic Energy Sciences, Division of Chemical Sciences, and by the Office of Advanced Scientific Computing Research, Mathematical, Information, and Computational Science Divison, under contract W-31-109-ENG-38.

Bifurcation analysis and computational study of a multi-component Bose–Einstein condensate

Shih-Feng Shieh
Department of Mathematics, National Tsinghua University
Department of Mathematics, National Tsinghua University, Hsinchu, 300, Taiwan
sfshieh@am.nthu.edu.tw

In this talk, we first prove that the solution curve of ground/positive bound states of a two component Bose-Einstein condensate(BEC) undergoes supercritical pitchfork bifurcations at some finite values of the inter-component scattering length. The ground state solutions bifurcate into two symmetric solutions with respect to some suitable axis on the symmetric domain, when a two-component BEC has equal intra- and inter-component scattering lengths.[2] Second, we propose a Gauss-Seidel-type iteration(GSI)for the computation of energy states of a multi-component BEC. We prove that the GSI method converges locally and linearly to a solution of a nonlinear algebraic eigenvalue problem if and only if the associated minimized energy functional problem has a strictly local minimum. Numerical experience shows that the GSI converges globally within 10 to 20 steps [1].

References

[1] S. M. Chang, W. W. Lin, and S. F. Shieh, *Gauss-Seidel-type Methods for Energy States of A Multi-Component Bose-Einstein Condensate*, J. Comp. Phy., 202 (2005), pp. 367–390.

[2] Y. C. Kuo, W. W. Lin, and S. F. Shieh, *Bifurcation Analysis of a Two-Component Bose-Einstein Condensate*, Physica D, 211 (2005), pp. 311–346.

Rigidity in finite-element matrices:
sufficient conditions for the rigidity of structures and substructures

Gil Shklarski
School of Computer Science, Tel-Aviv University
Tel-Aviv 69978, Israel
shagil@tau.ac.il

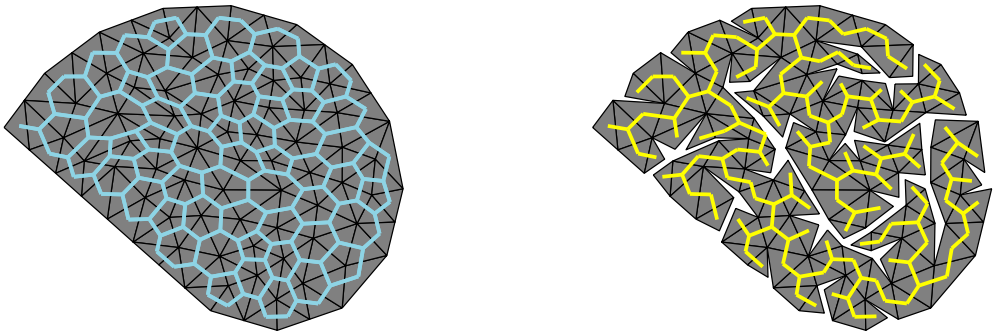
Sivan Toledo
School of Computer Science, Tel-Aviv University
Tel-Aviv 69978, Israel
stoledo@tau.ac.il

We present an algebraic-combinatorial theory of rigidity for finite-element matrices and apply this theory to two important problems: determining whether a finite-element matrix represents a rigid structure, and determining whether a matrix representing a structure and a matrix representing a substructure have the same range and null space. We address these problems by providing simple sufficient conditions for rigidity and null-spaces equality, and by providing linear-time algorithms (assuming bounded element degrees) to test these conditions. We also propose a sparsification technique for finite-element matrices that can be used to construct preconditioners for them.

Our results employ three new technical tools, one combinatorial and two algebraic. One algebraic tool is a purely-algebraic definition of the rigidity relationships between two rank-deficient matrices. The other algebraic tool is a definition of a finite-element matrix A as a sum $A = \sum_e A_e$ of symmetric semi-definite matrices that all satisfy a certain condition. The combinatorial tool is a graph, called the *rigidity graph*, that represents the rigidity relationships between the terms A_e of the finite-element matrix A. These tools may be applicable to the solution of other problems involving finite-element matrices.

We make several contributions to support preconditioning of finite-element matrices. First, we provide a reasonable definition of what a finite-element matrix is: a sum of element matrices whose null spaces are derived from a single global null space. Second, we provide a graph model of finite-element matrices, and propose graph algorithms for sparsifying the coefficient matrix A. Third, we provide simple combinatorial conditions that allow us to show that the range and null space of the sparsified matrix (the preconditioner) are the same as those of the original matrix A.

We also present a novel sparsification algorithm for finite-element matrices, called fretsaw extension, which is based on our new theory. The figure below illustrates this sparsification. We show that our algorithm generates preconditioners that have the same range and null-space as the original matrix, and that the preconditioners can be factored with essentially no fill. Numerical experiments suggest that this technique can serve as a basis for a new class of preconditioners.



(a) Linear elastic elements in two dimensions. The elements of the original structure and its rigidity graph in blue.

(b) The fretsaw extension of this structure.

Regularization for the total least squares problem

Diana M. Sima, Sabine Van Huffel
Department of Electrical Engineering, ESAT-SCD, Katholieke Universiteit Leuven
Kasteelpark Arenberg 10, 3001, Leuven, Belgium
{diana.sima,sabine.vanhuffel}@esat.kuleuven.be

We survey regularization methods that can be designed in the context of the total least squares problem. In the ordinary linear least squares regression, an incompatible (overdetermined) linear system $Ax \approx b$ is solved by computing the smallest 2-norm correction on b such that the system becomes compatible. In the total least squares problem, smallest (Frobenius norm) corrections on A and b are computed in order to make the system compatible.
When the system $Ax \approx b$ is a *discrete ill-posed problem*, the least squares or total least squares methods yield unreliable solutions for x, dominated by noise in the data or by numerical approximation errors. This happens because discrete ill-posed problems have an intrinsic sensitivity, which is shown through the fact that the singular values of A decay without gap towards zero (or towards a "noise level").
Regularization techniques are used for many years as a way of stabilizing the computation of least squares solutions in discrete ill-posed problems. Truncated singular value decomposition and Tikhonov regularization are two of the most known methods.
Regularization was also introduced in the total least squares context in the last decade. Truncated total least squares and quadratically constrained or penalty-based regularized total least squares techniques received attention in [3, 4, 7, 6, 2, 1].
In this talk, we focus on recent developments concerning regularization for the total least squares problem, which encompass the following items:

- recasting the truncated total least squares problem as a truncated *core problem*, where the term of "core problem," originating in [5], refers to a reduced, but essential part of a linear system, which can be computed from the SVD or another orthogonal decomposition, such as the bidiagonalization of A;
- efficient globally convergent numerical optimization methods for quadratically constrained or Tikhonov regularized total least squares;
- specially adapted model selection methodologies (with emphasis on *generalized cross validation*) for choosing the most appropriate regularization hyperparameters – the *truncation level* in the truncated total least squares framework and the *regularization parameter* in the regularized total least squares context.

References

[1] A. Beck and A. Ben-Tal, *On the solution of the Tikhonov regularization of the regularized total least squares problem*, SIAM Journal on Optimization. To appear.

[2] A. Beck, A. Ben-Tal, and M. Teboulle, *Finding a global optimal solution for a quadratically constrained fractional quadratic problem with applications to the regularized total least squares*, SIAM Journal on Matrix Analysis and Applications. To appear.

[3] R. D. Fierro, G. H. Golub, P. C. Hansen, and D. P. O'Leary, *Regularization by truncated total least squares*, SIAM Journal on Scientific Computing, 18 (1997), pp. 1223–1241.

[4] G. H. Golub, P. C. Hansen, and D. P. O'Leary, *Tikhonov regularization and total least squares*, SIAM Journal on Matrix Analysis and Applications, 21 (1999), pp. 185–194.

[5] C. C. Paige and Z. Strakoš, *Core problems in linear algebraic systems*, SIAM Journal on Matrix Analysis and Applications. To appear.

[6] R. Renaut and H. Guo, *Efficient algorithms for solution of regularized total least squares*, SIAM Journal on Matrix Analysis and Applications, 26 (2005), pp. 457–476.

[7] D. M. Sima, S. Van Huffel, and G. H. Golub, *Regularized total least squares based on quadratic eigenvalue problem solvers*, BIT Numerical Mathematics, 44 (2004), pp. 793–812.

Fast solution of KKT systems arising from a model control problem

Rene Simon
Special Research Program SFB F013 Johannes Kepler Universität Linz
Altenbergerstr. 69, A-4040 Linz, Austria
rene.simon@sfb013.uni-linz.ac.at

In optimal control we treat optimization problems with a partial differential equation as a constraint. Considering the first-order necessary conditions, we have to solve a large scale, sparse, symmetric, but indefinite system – KKT system. It can be partitioned into the following block structure:

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}$$

On the basis of a simple model optimal control problem, we will discuss efficient iterative methods for solving such saddle point systems. Two classes of methods are considered. The first idea is to use a multigrid approach to the whole system. Then we have to construct an appropriate smoothing operation. In the theoretical analysis we show, that an approximation property and a smoothing property are fulfilled, which are the main ingredients of proving the optimal convergence of the multigrid iteration.
As a second approach, we consider Krylov space methods; either methods for indefinite systems like MINRES or SYMMLQ, or the conjugate gradient method in a proper inner product. In both cases, efficient preconditioners are needed. In typical optimal control problems, the PDE is hidden in the off-diagonal block matrix B. In this case, the main issue is the construction of an efficient preconditioner for the Schur complement. Assuming, we have a good preconditioner $\tilde{B}$ for B, we want to discuss under which conditions $\tilde{B}^2$ is an appropriate preconditioner for B^2.

Krylov subspace methods for linear systems and matrix equations

Valeria Simoncini
Dipartimento di Matematica, Università di Bologna
Piazza di Porta S. Donato 5, I-40127 Bologna, Italy
valeria@dm.unibo.it

Krylov subspace methods are widely used for the iterative solution of a large variety of linear systems of equations. The underlying key idea of projecting the problem onto a subspace of smaller dimension has also been exploited to enhance convergence, by coupling the regular Krylov subspace with a specifically chosen subspace of small dimension. In this talk we review some recent developments in Krylov subspace methods, and discuss the use of projection-type approaches for solving challenging problems such as polynomial or parametrized systems, and other linear matrix equations.

Effective subspace expansion for computing eigenpairs

Gerard L. G. Sleijpen
Mathematical Institute, Utrecht University
Budapestlaan 6, NL-3508 TA Utrecht, The Netherlands
sleijpen@math.uu.nl

A parameter free ADI-like method for the numerical solution of large scale Lyapunov equations

D. C. Sorensen
Department of Computational and Applied Mathematics, Rice University,
Houston, Texas 77251-1892, USA
sorensen@rice.edu

An algorithm is presented for constructing a numerical solution to a large scale Lyapunov equation in low rank factored form. The algorithm is based upon a synthesis of an approximate power method and an alternating direction implicit (ADI) method. The former is parameter free and tends to be efficient in practice but there is little theoretical understanding of its properties. The ADI method has a well understood convergence theory, but the method relies upon selection of shift parameters and a poor shift selection can lead to very slow convergence in practice. The algorithm presented here uses an approximate power method iteration to obtain a basis update and then constructs an appropriate re-weighting of this basis to provide a factorization update that satisfies ADI-like convergence properties.

Nearly-linear time algorithms for solving symmetric, diagonally-dominant linear systems, and approximating Fiedler vectors for spectral partitioning

Daniel A. Spielman
Department of Computer Science, Yale University
51 Prospect Street, New Haven, CT 06511, U.S.A.
spielman@cs.yale.edu

Shang-Hua Teng
Department of Computer Science, Boston University
111 Cummington St., Boston, MA 02215, U.S.A.
steng@cs.bu.edu

We use subgraph preconditioners to solve symmetric, diagonally-dominant linear systems to precision ϵ in time $m \log^{O(1)} m \log(1/\epsilon)$, where m is the number of non-zero entries in the matrix. Using this algorithm, we can approximate the Fiedler vectors needed for spectral partitioning in a similar amount of time.
Our algorithm applies subgraph preconditioners in a multilevel fashion. These preconditioners improve upon the subgraph preconditioners first introduced by Vaidya [2]. On input a diagonally-dominant symmetric matrix A and $k < n$, our algorithm constructs in time $m \log^{O(1)} n$ a preconditioner B of A with $2(n - 1 + k)$ non-zero off-diagonal entries such that

$$\kappa_f(A, B) \leq (n/k) \log^{O(1)} n. \tag{1}$$

We will explain how to construct these preconditioners for planar systems, and thereby achieve an $O(m \log^2 m)$ time algorithm for their solution.
The construction of our preconditioners for general symmetric, diagonally-dominant matrices is more complicated, and is best understood by just considering Laplacian matrices of graphs. In the first stage of our construction, we build preconditioners by augmenting Laplacians of low-stretch spanning trees. Using the more recent constructions of low-stretch spanning trees [1], we obtain preconditioners with $2(n - 1 + k^2)$ non-zero off-diagonal entries such that

$$\kappa_f(A, B) \leq O((n/k) \log^2 n).$$

To reduce the number of non-zero off-diagonal entries, we introduce a procedure called sparsification. Our sparsification procedure takes as input an arbitrary n-by-n Laplacian matrix M and produces a matrix N with $n \log^{O(1)} n$ non-zero entries such that

$$\kappa_f(M, N) \leq 2.$$

The sparsification procedure also runs in time $n \log^{O(1)} n$. We obtain the preconditioners satisfying (1) by using the sparsification procedure to reduce the number of non-zero off-diagonal entries in the preconditioner obtained from the augmented spanning trees.

References

[1] K. Dhamdhere, A. Gupta, and H. Räcke, *Improved embeddings of graph metrics into random trees*, in SODA '06: Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm, New York, NY, USA, 2006, ACM Press, pp. 61–69.

[2] P. M. Vaidya, *Solving linear equations with symmetric diagonally dominant matrices by constructing good preconditioners.* Unpublished manuscript UIUC 1990. A talk based on the manuscript was presented at the IMA Workshop on Graph Theory and Sparse Matrix Computation, October 1991, Minneapolis., 1990.

An optimal solver to a KKT-system arising from local stress constrained topology optimization

Roman Stainko
Special Research Program SFB F013 (sponsored by FWF), Johannes Kepler University Linz,
Altenbergerstr. 69, A-4040 Linz, Austria.
roman.stainko@sfb013.uni-linz.ac.at

We consider topology optimization with local stress constraints. In the basic formulation we have a PDE-constrained optimization problem, where the finite element and design analysis are solved simultaneously. Here we introduce a new relaxation scheme based on a phase–field method. The starting point of this relaxation is a reformulation of the constraints of the optimization problem involving only linear and 0–1 constraints. The 0–1 constraints are then relaxed and approximated by a Cahn-Hillard type penalty in the objective functional. As the corresponding penalty parameter decreases to zero, it yields convergence of minimizers to 0–1 designs. A major advantage of this kind of relaxation scheme opposed to standard approaches is a uniform constraint qualification that is satisfied for any positive value of the penalization parameter.
After the relaxation we end up with a large-scale optimization problem with a high number of linear inequality constraints. Discretization is done by usual finite elements and for solving the resulting large-scale programming problems an interior-point method is used. Numerical experiments based on different stress criteria attest the success of the new approach.
Most of the computing time in interior-point methods is actually spent in solving linear saddle point problems arising from the primal-dual optimality equations, i.e. the perturbed KKT conditions. Multigrid methods certainly belong to the most efficient methods for solving large-scale systems, especially arising from discretized PDEs. The construction of such methods for saddle point problems is still a challenge. One of the most important ingredients of an efficient multigrid method is a proper smoother. In this talk we consider a multiplicative Schwarz-type smoother. Each iteration step of this smoother consists of the solution of several small local saddle points. Numerical experiments confirm the optimal linear complexity of the resulting solver.

PRIMME: PReconditioned Iterative MultiMethod Eigensolver, a robust, efficient and flexible Hermitian eigenvalue software

Andreas Stathopoulos
Department of Computer Science, College of William and Mary
Box 8795, Williamsburg, VA 23187-8795, USA
andreas@cs.wm.edu

James McCombs
Department of Computer Science, College of William and Mary
Box 8795, Williamsburg, VA 23187-8795, USA
mccombjr@cs.wm.edu

In recent years, many preconditioned eigensolvers have emerged that perform well in many, but not all applications. Moreover, in the related software, robustness has taken a secondary role behind efficiency. In view of the above, our group has developed a robust, efficient multi-method software called PRIMME. PRIMME is based on a Davidson-type main iteration, but it implements various techniques such block, locking, CG-type restarting, and adaptive inner-outer iterations, that allow it to transform to any current eigenvalue method. More than thirty features are controllable and tunable by the user, but a complete set of defaults requires only two to be set. The software runs both on parallel and sequential machines and can use the optimized BLAS and LAPACK libraries of the target machines. Additionally, PRIMME has the ability to sense both the computing environment and the problem solved, and adapt its choice of parameters accordingly.
In this talk, we present an overview of PRIMME, its robustness and efficiency, and show how JDQMR, one of the supported methods, can improve even on the Lanczos/ARPACK methods for Laplacian-type problems that are common in many applications.

The Gatlinburg Symposia and their influence on the development of numerical linear algebra

Josef Stoer
Institut für Mathematik, Universität Würzburg

About 1960, A.S. Householder initiated the biennial "Gatlinburg Symposia on Numerical Linear Algebra". They are named after the resort of Gatlinburg, Tennessee, where the early meetings took place; later on they shifted to various other locations in North America and also in Europe. Responsible for their program is the "Gatlinburg Committee"; the first consisted of A.S. Householder, J.H. Wilkinson, W. Givens, G.E. Forsythe, P. Henrici, and F.L. Bauer. There are only invited lectures and there were, at least initially, no parallel sessions. The list of participants contains many known names and reads like a "Who is who" in the area of Numerical Linear Algebra, so that it is not surprising that the Symposia had a tremendous influence on its development, both with respect to theory and the design of reliable and efficient algorithms: Such well-known software packages like LINPACK, LAPACK, EISPACK, and SPARSEPACK had their root in many critical discussions during Gatlinburg Symposia and the close cooperation of scientists attending these meetings.

A Krylov-Schur like method for quadratic eigenvalue problems with Hamiltonian symmetry

Martin Stoll
Computing Laboratory, University of Oxford
Wolfson Building, Parks Road, Oxford, OX1 3QD, England
martin.stoll@comlab.ox.ac.uk

Peter Benner
Mathematik in Industrie und Technik, Technische Universität Chemnitz
Reichenhainer Str. 41, 09126 Chemnitz, Germany
benner@mathematik.tu-chemnitz.de

Second-order eigenproblems with Hamiltonian spectral symmetry arise, for instance, from the stability analysis of gyroscopic systems or acoustic fluid-structure interaction. We show how a Krylov-Schur type algorithm based on the symplectic Lanczos process can be applied to such problems. The Krylov-Schur approach facilitates the introduction of deflation, locking and purging strategies for the implicitly restarted symplectic Lanczos process originally introduced by Benner and Faßbender in 1997. We demonstrate the efficiency of the method for several eigenproblems arising from the Finite Element models in the application context mentioned above.

Lanczos tridiagonalization and Golub-Kahan bidiagonalization: ideas, connections and impact

Zdeněk Strakoš
Institute of Computer Science, Academy of Sciences
Pod Vodárenskou věží 2, 182 07 Praha 8, Czech Republic
strakos@cs.cas.cz

The Lanczos algorithm was introduced in the middle of the last century in the context of computing eigenvalues and solving linear algebraic systems, see [2, 3]. It can be viewed in many different ways — for example as a (partial) reduction of the given real symmetric matrix to tridiagonal form. The closely related Golub-Kahan bidiagonalization was introduced in 1965 in the context of calculating the singular value decomposition [1], and applies to a general rectangular matrix. Both algorithms have influenced the theory and practice of scientific computing immensely. We will not survey the multitude of related algorithms and their implementations here. Instead, this presentation will concentrate on several mathematical ideas which link together both different periods, and different disciplines.

We start with interpolatory quadratures, continued fractions, orthogonal polynomials and the problem of moments — all inherited from the works of Gauss, Jacobi, Christoffel, Stieltjes, Chebyshev, Markov and many others. These deeply resonated many decades later with the analytically and/or computationally oriented developments of Krylov, Akhieser, Krein, Gantmakher, Lanczos, Stiefel, Henrici, Rutishauser and Vorobyev, just to name a few. Together, these prepared the scene for the modern formulation and analysis of the computational methods that are related to the subject of our contribution.

The list of authors who contributed significantly to the development and understanding of linear algebraic and/or eigenvalue solvers based on the Lanczos tridiagonalization and the Golub-Kahan bidiagonalization would be long. Various successful applications, descriptions of convergence based on the relationship to Gauss quadrature, and rounding error analyses which have led to our understanding of the behavior in finite precision arithmetic, certainly belong among the great achievements of numerical linear algebra in the second half of the 20th century. Related algorithms been have widely used (sometimes without noticing the relationship) in Statistics, Engineering, and the Sciences; with a very important role in Computational Physics and Computational Quantum Chemistry. From more recent developments it is worth mentioning, *e.g.*, the bidiagonalization-based methods for solving ill-posed problems, and the core problem formulation in errors-in-variables modelling. When recalling key papers in the area, one frequently encounters the names of Björck, Paige, Saunders, Parlett, van der Vorst and Greenbaum. In the area of computing orthogonal polynomials and Gauss-type quadratures, the state-of-the-art is strongly influenced by the encyclopedic work of Gautschi. For decades Golub has played an instrumental role in originating many developments in, and in making connections between, the areas of orthogonal polynomials and algebraic solvers, especially through the use of moments and quadrature.

In our contribution we recall some basic historical and interdisciplinary links. We close by giving examples of seemingly unrelated results. These demonstrate how a classical analytic view can be useful in approaching very recently formulated open problems, and vice versa. These open problems are motivated by computational issues — such as rates of convergence, and effects of rounding errors.

This work has been supported by the National Program of Research under project 1ET400300415.

References

[1] G. Golub and W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, J. Soc. Indust. Appl. Math. Ser. B Numer. Anal., 2 (1965), pp. 205–224.

[2] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Research Nat. Bur. Standards, 45 (1950), pp. 255–282.

[3] ———, *Solution of systems of linear equations by minimized-iterations*, J. Research Nat. Bur. Standards, 49 (1952), pp. 33–53.

A review of algebraic multigrid with emphasis on its industrial application

Klaus Stüben
Fraunhofer Institute SCAI
Schloss Birlinghoven, D-53754 St. Augustin, Germany
stueben@scai.fraunhofer.de

The development of algebraic multigrid (AMG) already started in the early nineteen-eighties, driven by the attempt to automate and generalize classical multigrid for the efficient solution of elliptic partial differential equations. While such "geometric" multigrid solvers operate on a pre-defined hierarchy of grids, algebraic multigrid solvers can be directly applied to (sparse) matrix equations, $Au = f$, without explicitly referring to geometry and without requiring any pre-defined hierarchy. In particular, based only on the concept of strength of connectivity between variables (e.g., defined by the size and sign of matrix coefficients) and by exploiting the Galerkin-principle, AMG directly mimics geometric multigrid on a hierarchy of linear systems of equations which is constructed fully automatically.
The original AMG approach is effectively restricted to particular classes of problems, an important one being the class of linear algebraic systems with matrices which are close to rowsum zero M-matrices. Problems like this widely occur in connection with discretized scalar elliptic PDEs of second order. In such cases, the original AMG is very mature and can handle millions of variables much more efficiently than any one-level method and is especially suited for unstructured grids. Although geometric multigrid solvers, when available, are generally still faster than their algebraic counterpart, the strengths of AMG are its robustness and ease of use, its applicability in complex geometric situations (in particular, in the case of unstructured grids in 2D or 3D) and its applicability to even solve certain problems which are beyond the reach of geometric multigrid (in particular, problems with no geometric or continuous background at all).
In spite of its potential, it took until around 1995 before there was a remarkable increase of interest in AMG and, more generally, in algebraically oriented hierarchical methods, both in science and applications. Among other reasons, this interest was fed by the increasing geometrical complexity of applications which, technically, limited the immediate use of alternate fast solvers such as those based on geometric multigrid. Another reason was the steadily increasing demand for efficient "plug-in" solvers. In industrial software development, this demand was driven by increasing problem sizes which made clear the limits of the classical one-level solvers still used in most industrial software packages. For instance, CFD applications in the car industry involve very complicated flow regions. Flows through heating and cooling systems, vehicle underhood flows, flows around complete car bodies, or flows within passenger compartments are computed on a regular basis. Large complex meshes, normally unstructured, are used to model such situations. Requirements on the achievable accuracy are ever increasing, leading to finer and finer meshes. Meshes of the order of 10-100 million cells are required for today's simulation purposes.
Fostered by this situation, sophisticated extensions of the original AMG method have been developed aiming at increasing its range of applicability. In particular, substantial progress has been achieved towards the efficient treatment of elliptic systems of PDEs. Several other possibilities to generalize AMG have been investigated, research on various new and related approaches has started and there is still an ongoing rapid development. In spite of this, however, there is no unique and best approach yet. Each method seems to have its range of applicability but all of them may fail to be efficient in certain other applications.
The "quality" of AMG's coarsening process (including the interpolation and the coarse-level operators) is crucial for obtaining fast and (nearly) h-independent convergence. Generally, the more effort is put into this construction, the faster the convergence can be. But, unfortunately, the required numerical work may increase even faster. That is, from a practical point of view, a major problem in designing efficient AMG algorithms is the tradeoff between convergence, memory requirement and numerical work. Keeping the balance between these three aspects is the ultimate goal of any practical algorithm.
In this presentation, we give a review on "classical" AMG approaches and some of their major developments. Various coarsening strategies as well as generalizations towards the treatment of systems of PDEs will be discussed. We will present concrete results for selected cases to demonstrate the efficiency and potential of AMG. A special focus of the presentation will be on those developments which have appeared to be particularly important for AMG's acceptance in industrial simulation.

Balanced truncation model reduction for optimal control problems

Tatjana Stykel
Institut für Mathematik, Technische Universität Berlin
Strasse des 17.Juni 136 D-10623 Berlin, Germany
stykel@math.tu-berlin.de

J. C. de los Reyes
Departamento de Matemática, EPN Quito
Ladrón de Guevara E11-253, Quito, Ecuador
juancadelosreyes@gmx.net

We will discuss the application of model order reduction to optimal control problems governed by nonlinear time-dependent partial differential equations. The main idea of the method considered here consists in a spatial discretization of the optimality system and application of a balanced truncation model reduction method to the semidiscretized linear adjoint equation. The convergence analysis of the descent method will be provided. We will also present several numerical examples to demonstrate the properties of the proposed method.

Efficient updates in Bregman's algorithm optimizing for von Neumann matrix divergence with optional range space constraint

Mátyás Sustik
University of Texas at Austin – Department of Computer Sciences
1 University Station C0500, Austin, TX, 78758, USA
sustik@cs.utexas.edu

Inderjit S. Dhillon
University of Texas at Austin – Department of Computer Sciences
1 University Station C0500, Austin, TX, 78758, USA
inderjit@cs.utexas.edu

There has been recent interest in matrix nearness problems when the measure of nearness is a Bregman matrix divergence. Example applications include clustering, kernel learning, finding the nearest correlation matrix, see [2, 1]. In this talk we analyze the projection step of Bregman's algorithm which can be used to solve the underlying optimization problem. We consider the von Neumann and the Burg matrix divergencies. In case of the von Neumann matrix divergence the matrix update amounts to the calculation of $exp(log(A) + \xi cc^*)$, where the matrix A is positive definite and cc^* is a rank-one update. We show how to do the updates efficiently on a factored form of the matrix. We extend the definitions from the domain of positive definite matrices to positive semi-definite matrices and show that our factored form based algorithm performs well in the low rank case. The talk is concluded with discussion on how to use our methods to find the nearest correlation matrix in. This is joint work with Brian Kulis.

References

[1] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh, *Clustering with bregman divergences*, Journal of Machine Learning Research, 6 (2005), pp. 1705–1749.

[2] I. S. Dhillon and J. Tropp, *Matrix nearness problems using Bregman divergences*. Submitted for publication, 2006.

Ideal GMRES and the polynomial numerical hull for a Jordan block

Petr Tichý
Institut für Mathematik, Technische Universität Berlin
Straße des 17. Juni 136, D-10623 Berlin, Germany
tichy@math.tu-berlin.de

Jörg Liesen
Institut für Mathematik, Technische Universität Berlin
Straße des 17. Juni 136, D-10623 Berlin, Germany
liesen@math.tu-berlin.de

The convergence analysis of the GMRES method for solving linear algebraic systems $Ax = b$ is still an open area of research. If A is normal, then GMRES solves a certain approximation problem on the spectrum of A. For nonnormal matrices, the situation is much less clear. A possible approach for investigating the GMRES convergence behavior in the nonnormal case is to concentrate on the *ideal* GMRES approximation

$$\Phi_k^A \equiv \min_{p \in \pi_k} \|p(A)\| ,$$

that represents a bound on the *worst-case* GMRES residual norm. Two open research problems arise. First, for which classes of nonnormal matrices can the value Φ_k^A be identified with the worst-case GMRES residual norm? Second, how can the value of Φ_k^A be evaluated or estimated?
In this talk we address both these two questions for a very simple nonnormal matrix, namely an n by n Jordan block J. Under some assumptions, we show that ideal and worst-case GMRES are identical at steps k and $n-k$ such that k divides n, and we derive explicit expressions for the $(n-k)$th ideal GMRES approximation. Furthermore, we extend previous results in the literature by showing new results about the radii of the polynomial numerical hulls of Jordan blocks. Using these, we discuss the tightness of the lower bound on the ideal GMRES approximation that is derived from the radius of the polynomial numerical hull of J.
This presentation is based on our paper [1].

References

[1] P. Tichý and J. Liesen, *GMRES convergence and the polynomial numerical hull for a Jordan block*, submitted to Linear Algebra and its Applications, 2005.

On structured condition numbers and backward errors in scalar product spaces

Françoise Tisseur
School of Mathematics, The University of Manchester
Sackville Street, Manchester, M60 1QD, UK
F.Tisseur@manchester.ac.uk

Stef Graillat
Laboratoire LP2A, Université de Perpignan
52, avenue Paul Alduy, F-66860 Perpignan Cedex, France
graillat@univ-perp.fr

Michael Karow, Daniel Kressner
Institut für Mathematik, MA 4-5, TU Berlin
Str. des 17. Juni 136, D-10623 Berlin, FRG
{karow,kressner}@math.tu-berlin.de

We investigate structured condition numbers and backward errors for

- the solution to a linear system,
- matrix inversion,
- distance to singularity of a matrix,
- the eigenvalues of structured matrices.

Particular attention is paid to linear and nonlinear structures that form Lie algebras, Jordan algebras and automorphism groups of a scalar product that is both orthosymmetric and unitary. These include, for example, complex symmetric, pseudo-symmetric, persymmetric, skew-symmetric, Hamiltonian, unitary, complex orthogonal and symplectic matrices. We show that there is little or no difference between structured and unstructured condition numbers for the solution to a linear system, matrix inversion and distance to singularity for matrices in Lie and Jordan algebras. Hence, for these classes of matrices, the usual unstructured perturbation analysis is sufficient.
The structured and unstructured eigenvalue condition numbers are equal for structures in Jordan algebras. For Lie algebras, the effect on the eigenvalue condition number of incorporating structure varies greatly with the structure.
Structured eigenvalue and linear system condition numbers for nonlinear structures in automorphism groups can be much smaller than the corresponding unstructured condition number. For these nonlinear structured we derive bounds and computable expressions for the structured condition numbers.
Structured backward errors for approximate solutions of linear systems are also considered. Conditions are given for the structured backward error to be finite. We prove that for Lie and Jordan algebras, whenever the structured backward error is finite it is within a small factor of the unstructured one. The same conclusion holds for orthogonal and unitary structures but cannot easily be extended to other matrix groups. This work extends and unifies earlier analyses. For details see [1] and [2].

References

[1] M. Karow, D. Kressner, and F. Tisseur, *Structured eigenvalue condition numbers*, Numerical Analysis Report No. 467, Manchester Centre for Computational Mathematics, Manchester, England, Apr. 2005.

[2] F. Tisseur and S. Graillat, *Structured condition numbers and backward errors in scalar product spaces*, Numerical Analysis Report No. 473, Manchester Centre for Computational Mathematics, Manchester, England, Sept. 2005.

Talbot quadratures and rational approximations

Lloyd N. Trefethen
Oxford University Computing Laboratory
Wolfson Bldg., Parks Road, Oxford OX1 3QD, UK
LNT@comlab.ox.ac.uk

J. A. C. Weideman
Dept. of Applied Mathematics, University of Stellenbosch
Private Bag X1, Matieland 7602, South Africa
weideman@sun.ac.za

Thomas Schmelzer
Oxford University Computing Laboratory
Wolfson Bldg., Parks Road, Oxford OX1 3QD, UK
thoms@comlab.ox.ac.uk

Many computational problems related to functions of matrices and operators can be solved with the aid of contour integrals containing e^z in the the integrand: examples include inverse Laplace transforms, special functions, parabolic PDEs, and ETD-type methods for stiff PDEs such as reaction-diffusion equations. One approach to the numerical quadrature of such integrals is to apply the trapezoid rule on a Hankel contour defined by a suitable change of variables. Optimal parameters for three classes of such contours have recently been derived by Weideman: (a) parabolas, (b) hyperbolas, and (c) cotangent contours, following Talbot in 1979. The convergence rates for these optimized quadrature formulas are very fast: roughly $O(3^{-N})$, where N is the number of sample points or function evaluations. On the other hand, convergence at a rate apparently about twice as fast, $O(9.28903^{-N})$, can be achieved by using a different approach: best supremum-norm rational approximants to e^z for $z \in (-\infty, 0]$, following Cody, Meinardus and Varga in 1969 (and computed in practice by the method of Carathéodory–Fejér approximation, based on the singular values of a Hankel matrix of Chebyshev coefficients). We review these methods and show how the quadrature formulas can be interpreted as rational approximations and the rational approximations as quadrature formulas. The strengths and weaknesses of the different approaches are discussed.

Bounds on the solution to the Lyapunov equation with a general stable matrix

Ninoslav Truhar
Department of Mathematics, University of Osijek, Croatia
Trg Ljudevita Gaja 6, HR-31000 Osijek, Croatia
ntruhar@mathos.hr

Krešimir Veselić
LG Mathematische Physik, Fernuniversität Hagen, Germany
Universitätsstraße 21, D-58084 Hagen, Germany
kresimir.veselic@fernuni-hagen.de

We present some new estimates for the eigenvalue decay of the Lya- punov equation $AX + XA^T = B$ with a low rank right-hand side B. The new bounds show that the right-hand side B can greatly influence the eigenvalue decay rate of the solution. This suggests a new choice of the ADI-parameters for the iterative solution. The advantage of these new parameters is illustrated on second order damped systems with a low rank damping matrix. We will also present the new perturbation bound for the solution X of the Lyapunov equation with the general matrix A.

Hadamard scaling of matrix quadratic equations

Vassilios A. Tsachouridis
Centre for Mathematical Science, School of Engineering and Mathematical Sciences, City University London
Northampton Square, EC1V 0HB London, UK
vassilios.tsachouridis@ieee.org

A scaling framework for general quadratic algebraic matrix equations is presented. The presented method is based on the assignment of predetermined values to the coefficients and the unknowns of the matrix equations using Hadamard multiplication. The presented methods are implemented for the special case of scaling algebraic matrix Riccati equations, where the numerical solution of ill conditioned problems is considered via the Newton algorithm.

P-matrices and the principal minor assignment problem

Michael Tsatsomeros
Washington State University
Department of Mathematics, Washington State University Pullman, WA 99164-3113, USA
tsat@wsu.edu

Challenges and some results related to the Principal Minor Assignment Problem will be discussed, especially as they pertain to P-matrices and matrices of specific combinatorial structure.

Null-space methods for solving saddle-point problems

Miroslav Tůma
Institute of Computer Science, Academy of Sciences of the Czech Republic
Pod vodárenskou věží 2, 18207 Praha 8, Czech Republic
tuma@cs.cas.cz

Mario Arioli
Numerical Analysis Group, Rutherford Appleton Laboratory
Chilton, Didcot, Oxfordshire, OX11 0QX, UK
m.arioli@rl.ac.uk

Jiří Maryška
Technical University in Liberec
Hálkova 6, 461 17 Liberec, Czech Republic
jiri.maryska@vslib.cz

Miroslav Rozložník
Institute of Computer Science, Academy of Sciences of the Czech Republic
Pod vodárenskou věží 2, 18207 Praha 8, Czech Republic
miro@cs.cas.cz

The goal of this contribution is to discuss null-space methods for solving saddle-point problems. We will consider symmetric indefinite saddle-point problems arising from a particular mixed finite element discretization of the stationary fluid flow problem.
We propose solution techniques based on computation of the null-space basis of the whole or of a part of the left lower off-diagonal block in the system matrix. Such solution techniques are often motivated by the need to solve a sequence of such systems with the same mesh but different material properties. In this case, subsequent systems can reuse some information from previous members of the sequence.
The general framework of null-space methods covers a couple of approaches. On one hand, a fundamental cycle null-space basis of the whole off-diagonal block based on the spanning tree of an associated graph can be constructed. Note that at the continuous level, such basis can be obtained from curls of appropriate vector potentials. It is shown that this basis may be theoretically rather ill-conditioned. Our experiments confirm this. Nevertheless, the resulting projected system is symmetric positive definite and so the conjugate gradient method can be applied. On the other hand, the orthogonal null-space basis of its sub-block that enforces continuity over faces can be easily constructed. The projected system then remains indefinite and the preconditioned minimal residual method for a reduced system can be used. Rates of convergence in these cases are discussed. Efficiencies of preconditioned solvers based on these reductions are compared in numerical experiments.
The null-space method is typically not considered as the first method of choice for solving large and sparse saddle-point problems. One of our aims is to show that this approach can be better than a Schur complement-based solution strategy for systems with ill-conditioned left-upper block of the saddle-point matrix which corresponds, in our application, to the tensor of hydraulic permeability.
Partially supported by the project 1ET400300415 within the National Program of Research “Information Society”.

A geometric level set method for transcendental equations from chemical engineering with multiple roots and bifurcation

Frank Uhlig
Auburn University, Department of Mathematics and Statistics
Auburn University, Auburn, AL 36849-5307, USA
uhligfd@auburn.edu

(Joint work with Chadia Affane, Auburn University, `affanac@auburn.edu` and
Said Elnashaie, University of British Columbia, `nashaie@chml.ubc.ca`)

When trying to solve a single transcendental equation $f(x, a) = 0$ that depends on one parameter a, we have found the standard root finders such as bisection and Newton's method lacking within the parameter range of a with multiple roots.
Such equations occur naturally in chemical and biological engineering problems associated with continuously stirred tank reactors (CSTR), fluidized bed catalytic reactors etc., where the occurrence of multiple steady states and the bifurcation behavior have to be understood for proper plant operation.
Here we propose, implement, and test a level set method for solving parameterized transcendental equations with bifurcation. We create the elevation matrix data $Z = f(x, a) \in R^{n,n}$ for x and a data vectors in R^n and draw the level zero contour curve for the associated surface $z = f(x, a) \in R^3$ via data interpolation. The method is very simply implemented via MATLAB's built-in `contour` command. It works exceedingly well when there are multiple solutions x for one a and is much faster than when using standard root finders that generally fail do give useful information near the bifurcation points.

(Our level set method, invented originally for transcendental equations has been modified to find the roots of polynomials geometrically. [ILAS 2006 → Amsterdam])

Solution strategies for stochastic finite element discretizations

Elisabeth Ullmann
Institut für Numerische Mathematik und Optimierung – Fakultät für Mathematik und Informatik,
Technische Universtät Bergakademie Freiberg
Akademiestraße 6, D-09599 Freiberg, Germany
ullmann@math.tu-freiberg.de

Michael Eiermann, Oliver G. Ernst
Institut für Numerische Mathematik und Optimierung – Fakultät für Mathematik und Informatik,
Technische Universtät Bergakademie Freiberg
Akademiestraße 6, D-09599 Freiberg, Germany
{eiermann,ernst}@math.tu-freiberg.de

Various physical phenomena are mathematically modeled by partial differential equations together with appropriate boundary conditions. Their approximate solution can be computed efficiently and accurately for instance by the finite element method. On the other hand, input data such as material parameters, boundary conditions and loads contain measurement or modelling errors that may outweigh discretization errors coming from the numerical simulation.
In recent years it has become popular to take account of this problem in the model itself by treating all input data as spatially dependent random variables, i.e., random fields. Consequently the randomness is propagated to the output data and allows to quantify the uncertainty of the output.
From the mathematical point of view the task is to solve a stochastic partial differential equation. We consider elliptic boundary value problems with stochastic coefficients or right hand side that are equivalent to stochastic variational problems. In analogy to the case of deterministic input data, one discretises such problems by means of the stochastic finite element method (SFEM) [3].
The basic principle is to treat the spatial and stochastic parts of the problem separately. Ansatz and test functions are elements of tensor product spaces that contain solely deterministic or stochastic shape functions. As a consequence, the discretization of the spatial part is independent of that of the stochastic part. On the other hand the total number of degrees of freedom (DOFs) is exactly the number of DOFs of the deterministic problem multiplied by the number of stochastic DOFs. Therefore we require efficient solvers for the large linear systems that arise from SFEM.
We discuss solution strategies for the Galerkin system that take advantage of the special structure of the system matrix. For stochastic coefficients linear in a finite set of independent random variables we employ Krylov subspace recycling techniques [4] after having decoupled the SFEM stiffness matrix [1, 2].

References

[1] I. Babuška, R. Tempone, and G. E. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM Journal on Numerical Analysis, 42 (2004), pp. 800–825.

[2] M. Eiermann, O. G. Ernst, and E. Ullmann, *Computational aspects of the stochastic finite element method*, To appear in Computing and Visualization in Science.

[3] R. Ghanem and P. Spanos, *Stochastic finite elements: a spectral approach*, Springer-Verlag, New York, 1991.

[4] M. Parks, E. de Sturler, G. Mackey, D. Johnson, and S. Maiti, *Recycling Krylov subspaces for sequences of linear systems*, Technical Report UIUCDCS-R-2004-2421(CS), University of Illinois, March 2004.

Potentially nilpotent and spectrally arbitrary even cycle sign patterns

Pauline van den Driessche
Department of Mathematics and Statistics, University of Victoria
Victoria, BC, Canada, V8W3P4
pvdd@math.uvic.ca

Brad Bingham
Department of Computer Science, University of British Columbia
binghamb@cs.ubc.ca

Dale Olesky
Department of Computer Science, University of Victoria
dolesky@cs.uvic.ca

An $n \times n$ sign pattern $\mathcal{S}_n$ is potentially nilpotent if there is a real matrix having sign pattern $\mathcal{S}_n$ and characteristic polynomial x^n. A new family of sign patterns $\mathcal{C}_n$ with a cycle of every even length is introduced and shown to be potentially nilpotent by explicitly determining the entries of a nilpotent matrix with sign pattern $\mathcal{C}_n$. These nilpotent matrices are used together with a Jacobian argument to show that $\mathcal{C}_n$ is spectrally arbitrary, i.e., there is a real matrix having sign pattern $\mathcal{C}_n$ and characteristic polynomial $x^n + \sum_{i=1}^{n}(-1)^i \mu_i x^{n-i}$ for any real μ_i. Some results and a conjecture on minimality of these spectrally arbitrary sign patterns are given.

Reminiscences in developments in applied linear algebra

Richard S. Varga
Institute for Computational Mathematics
Kent State University, Kent, OH 44242, USA
varga@math.kent.edu

This talk will discuss the following topics:

1. Classical Iterative Methods, for solving systems of linear equations (Jacobi, Gauss-Seidel, Southwell, etc.).
2. Young's Property A (cyclic of index 2) Method for the SOR (successive overrelaxation) iterative method and its generalizations.
3. Applications of M-matrix and H-matrix theory in linear algebra, plus various kinds of splittings and their generalizations.
4. Application of tools from approximation theory and complex function theory in the solution of linear systems of linear equations.
5. Alternating Direction Implicit (ADI) iterative methods for solving linear systems of linear equations.
6. Multigrid Methods in solving systems of llnear equations.
7. Direct Methods (e.g., the Lanczos method) for solving linear systems, which, because of rounding errors, become useful iterative methods.
8. New Directions and Applications of the above theories.

On the computation of the null space of a sparse matrix

Xavier Vasseur
CERFACS
42 avenue Gaspard Coriolis, F-31057 Toulouse cedex 1, France
Xavier.Vasseur@cerfacs.fr

Serge Gratton
CERFACS
42 avenue Gaspard Coriolis, F-31057 Toulouse cedex 1, France
Serge.Gratton@cerfacs.fr

Computing a basis for the null space of a large and sparse matrix A with m rows and n columns is often a key step of some algorithms in e.g. structural analysis, fluid mechanics, constrained optimization, image reconstruction or electrical engineering.
In this talk we will first review various computational procedures that have been proposed when A is rectangular (case $m < n$ [1, 2, 3] and $m \geq n$ [4]). These approaches are often based either on a sparse Gaussian elimination procedure with pivoting or on a sparse QR factorization applied to A. We will then present an algorithm that combines a sparse orthogonal scheme with an iterative method to compute a null space basis of A and show some preliminary numerical results.

References

[1] M. W. Berry, M. T. Heath, I. Kaneko, M. Lawo, R. J. Plemmons, and R. C. Ward, *An algorithm to compute a sparse basis of the null space*, Numerische Mathematik, 47 (1985), pp. 483–504.

[2] T. F. Coleman and A. Pothen, *The null space problem I. Complexity*, SIAM J. Algebraic and Discrete Methods, 7 (1986), pp. 527–537.

[3] ———, *The null space problem II. Algorithms*, SIAM J. Algebraic and Discrete Methods, 8 (1987), pp. 544–563.

[4] C. Gotsman and S. Toledo, *On the computation of null spaces of sparse rectangular matrices*, (2005). Submitted to SIAM Journal on Matrix Analysis and Applications.

An SVD approach to identifying meta-stable states of Markov chains

Elena Virnik
TU Berlin, MA 462, Institut für Mathematik, Sekretariat MA 4-5
D-10623 Berlin, Germany
virnik@math.tu-berlin.de

Being one of the key tools in conformation dynamics, the identification of meta-stable states of Markov chains has been subject to extensive research in recent years. Previous work on this topic usually involves the computation of the eigenvalue cluster close to one, as well as the corresponding eigenvectors and the stationary probability distribution of the stochastic matrix. As a possible alternative, we present an SVD approach to identifying meta-stable states of a stochastic matrix, where we only have to calculate the second largest singular vector. We outline some theoretical background and discuss the advantages of this strategy. Some simulated and real numerical examples illustrate the effectiveness of the algorithm.

Positive descriptor systems

Elena Virnik
TU Berlin, MA 462, Institut für Mathematik, Sekretariat MA 4-5
D-10623 Berlin, Germany
virnik@math.tu-berlin.de

We consider linear time-invariant control systems of the form

$$\begin{aligned} E\dot{x} &= Ax + Bu,\ x(t_0) = x_0 \\ y &= Cx, \end{aligned}$$

where $x \in C(\mathbb{R}_+, \mathbb{R}^n)$ is the state vector function, $u \in C(\mathbb{R}_+, \mathbb{R}^m)$ the input vector function, $y \in C(\mathbb{R}_+, \mathbb{R}^p)$ the output vector function and $E, A \in \mathbb{R}^{q\times n}$, $B \in \mathbb{R}^{q\times m}$, $C \in \mathbb{R}^{p\times n}$ are constant coefficient matrices. In applications such as pharmacokinetics or drug design, biological or chemical systems are modelled by descriptor systems, in which the state x describes concentrations, populations of species, or numbers of cells. The solution is then a nonnegative vector function. Hence, the numerical methods for the control or simulation should respect this special structure. This leads to a number of problems. The application of the classical control theory is not possible, since methods for, e.g., control design, model reduction etc. not necessarily preserve the positivity properties of the system. We consider differential-algebraic equations since in practice the dynamics of the system are usually coupled with algebraic constraints such as capacities or concentration bounds. This leads to further difficulties. We will give an introduction to positive descriptor systems and discuss several control theoretical topics in the positive setting.

Recent research around the MRRR algorithm

Christof Vömel
Computational Research Division, Lawrence Berkeley National Lab
1 Cyclotron Road, MS 50F-1650, Berkeley, CA 94720, USA
cvoemel@lbl.gov

This talk gives an overview about recent research and development around the algorithm of Multiple Relatively Robust Representations (MRRR, or MR^3) for the symmetric eigenvalue problem [3, 4, 5]. In our presentation, we focus on the following performance-related aspects:

1. **The use of IEEE arithmetic.** Bidiagonal factorizations by the differential qds algorithms [8] play a central role in the MRRR algorithm.. Their implementation relies on IEEE-754 arithmetic to avoid branched tests in loops [2]. An invalid operation ∞/∞, denoted by NaN ('Not a Number') is allowed, but the computation needs to be repeated by a slower variant that avoids the NaN. We discuss benefits and possible architecture-related penalties [7].

2. **Parallelization.** The MRRR algorithm naturally allows for a parallel computation with perfect $\mathcal{O}(n^2/p)$ memory per processor, (p denoting the number of processors used). We describe our new parallel eigensolver PDSYEVR [1] and compare it with existing ScaLAPACK and PLAPACK eigensolvers. The relationship between parallelization and computing subsets of eigenpairs is also discussed [6].

3. **Vectorization and multisection.** We discuss possibilities to improve the sequential and parallel performance of MRRR by multisection based on vectorization or shared memory parallelism. This work is still in its early stages but preliminary experiments indicate that important performance gains are possible.

This is joint work with D. Antonelli, Y. Bai, J. Demmel, I. Dhillon, T. Katagiri, O. Marques, B. Parlett, J. Riedy, and R. Ward.

References

[1] D. Antonelli and C. Vömel, *LAPACK working note 168:* PDSYEVR. *ScaLAPACK's parallel MRRR algorithm for the symmetric eigenvalue problem*, Tech. Rep. UCBCSD-05-1399, University of California, Berkeley, 2005.

[2] J. W. Demmel and X. S. Li, *Faster numerical algorithms via exception handling*, IEEE Trans. Comp., 43 (1994), pp. 983–992. (Also: LAPACK Working Note 59).

[3] I. S. Dhillon, *A New $O(n^2)$ Algorithm for the Symmetric Tridiagonal Eigenvalue/Eigenvector Problem*, PhD thesis, University of California, Berkeley, California, 1997.

[4] I. S. Dhillon and B. N. Parlett, *Multiple representations to compute orthogonal eigenvectors of symmetric tridiagonal matrices*, Linear Algebra and Appl., 387 (2004), pp. 1–28.

[5] ———, *Orthogonal eigenvectors and relative gaps*, SIAM J. Matrix Anal. Appl., 25(3) (2004), pp. 858–899.

[6] O. A. Marques, B. N. Parlett, and C. Vömel, *LAPACK working note 167: Subset computations with the MRRR algorithm*, Tech. Rep. UCBCSD-05-1392, University of California, Berkeley, 2005.

[7] O. A. Marques, E. J. Riedy, and C. Vömel, *Lapack working note 172: Benefits of IEEE-754 features in modern symmetric tridiagonal eigensolvers*, Tech. Rep. UCBCSD-05-1414, University of California, Berkeley, 2005.

[8] B. N. Parlett, *Acta Numerica*, Cambridge University Press, 1995, ch. The new qd algorithms, pp. 459–491.

Accelerating the parallel band edge state calculation of a semiconductor quantum dot

Christof Vömel
Computational Research Division, Lawrence Berkeley National Lab
1 Cyclotron Road, MS 50F-1650, Berkeley, CA 94720, USA
cvoemel@lbl.gov

Reducing large gyroscopic eigenproblems by automated multi-level substructuring

Heinrich Voss
Institute of Numerical Simulation, Hamburg University of Technology
Schwarzenbergstrasse 95, D-21073 Hamburg, Germany
voss@tuhh.de

Kolja Elssel
Institute of Numerical Simulation, Hamburg University of Technology
Schwarzenbergstrasse 95, D-21073 Hamburg, Germany
elssel@tuhh.de

Simulating numerically the sound radiation of a rolling tire requires the solution of a very large and sparse gyroscopic eigenvalue problem. Taking advantage of the automated multi–level substructuring (AMLS) method it can be projected to a much smaller gyroscopic problem, the solution of which however is still quite costly since the eigenmodes are non–real and complex arithmetic is necessary. This paper discusses the application of AMLS to huge gyroscopic problems and the numerical solution of the AMLS reduction. A numerical example demonstrates the efficiency of AMLS.

An inverse eigenvalue problem for a symmetric arrowhead matrix

Nils Wagner
Institut für Angewandte und Experimentelle Mechanik, Universität Stuttgart
Pfaffenwaldring 20, D-70550 Stuttgart, Germany
nwagner@mecha.uni-stuttgart.de

We consider an inverse eigenvalue problem for a symmetric arrowhead matrix [5]. Such matrices arise in the description of a primary mechanical oscillator connected to continuously distributed secondary oscillators [1, 3, 4]. The equations of motion for such a system results in a second-order system with a symmetric positive definite mass matrix M and a symmetric positive definite stiffness matrix K. The associated generalized eigenvalue problem is reduced to a standard eigenvalue problem with a mass-normalized stiffness matrix $\tilde{K}$. Then the following inverse eigenvalue problem is studied here: Construct a physically realizable $\tilde{K}$ such that the associated eigenvalue problem has the specified eigenvalues λ^*. We benefit from a secular equation and an interlacing property [2] which results from the special structure of $\tilde{K}$.

References

[1] A. Carcaterra and A. Akay, *Transient energy exchange between a primary structure and a set of oscillators: Return time and apparent damping*, J. Acoust. Soc. Am., 115 (2004).

[2] O. E. Livne and A. Brandt, *n roots of the secular equation in $o(n)$ operations*, SIAM J. Matrix Anal. Appl., 24 (2002).

[3] J.-M. Mencik and A. Berry, *A theoretical formulation of the response of a master structure coupled with elastic continuous fuzzy subsystems with discrete attachments*, Journal of Sound and Vibration, 280 (2005).

[4] R. J. Nagem, I. Veljkovic, and G. Sandri, *Vibration damping by a continuous distribution of undamped oscillators*, Journal of Sound and Vibration, 207 (1997).

[5] D. P. O'Leary and G. W. Stewart, *Computing the eigenvalues and eigenvectors of symmetric arrowhead matrices*, Journal of Computational Physics, 90 (1990).

Constraint preconditioning and Schilders factorization in PDE-constrained optimization

Andy Wathen
Oxford University
Oxford University Computing Lab, Parks Road, Oxford, OX1 3QD, UK
wathen@comlab.ox.ac.uk

PDE-Constrained Optimization gives rise to very large scale saddle-point (KKT) systems like other types of constrained problems. Preconditioning of the resulting indefinite linearised systems is generally necessary for acceptable convergence of iterative methods.
In this talk we will discuss the attractive theoretical feature of so-called 'constraint preconditioners' - preconditioners which include the constraint blocks of the original problem: these would be forward and adjoint solves of the PDEs in the PDE-Constrained Optimization setting - and describe various effective ways to implement them via the 'Schilders Factorization'. By using different approximations to the various blocks in this factorization one is able to describe a number of possible practical approaches and thus balance the speed of iterative convergence with the cost of preconditioner application.

Some of this is joint work with Sue Dollar, Nick Gould and Wil Schilders.

The reduction to Hamiltonian Schur form explained

David Watkins
Department of Mathematics
Washington State University
Pullman, WA, 99164-3113, USA
watkins@math.wsu.edu

The Linear-Quadratic Gaussian problem of control theory can be solved by computing the stable invariant subspace of a Hamiltonian matrix. For 25 years it has been an open problem to produce a backward-stable algorithm that solves the problem in a structure-preserving way by computing the Hamiltonian Schur form. Recently Chu, Liu, and Mehrmann proposed such an algorithm. We will present an alternate derivation of their algorithm that explains it in terms of swapping blocks of eigenvalues in the skew-Hamiltonian matrix obtained by squaring the Hamiltonian matrix.

Robust Perron cluster analysis in conformation dynamics

Marcus Weber
Zuse Institute Berlin (ZIB)
Takustraße 7, D-14195 Berlin, Germany
weber@zib.de

The talk is about Robust Perron Cluster Analysis which is a useful clustering method for various applications in computational life sciences [1, 3]. In molecular conformation dynamics one is interested in a "soft" clustering of the Boltzmann-distributed position states of a molecule into dynamically metastable parts. The basis for this clustering is a stochastic transition matrix P for a given discretization of the position space. A simple linear transformation maps the Perron Cluster eigenvectors of P to the desired almost characteristic membership vectors. Robust Perron Cluster Analysis solves an optimization problem, in which "maximizing metastability" is formulated as a convex objective function with linear non-negativity and partition-of-unity constraints [2]. We will explain this optimization approach and present some new results about the simplex structure of the input eigenvectors. From these results we derive how to use Robust Perron Cluster Analysis for general geometrical cluster problems.

References

[1] P. Deuflhard and M. Weber, *Robust Perron cluster analysis in conformation dynamics*, Lin. Alg. Appl., 398 (2005), pp. 161–184.

[2] M. Weber, *Meshless Methods in Conformation Dynamics*, PhD thesis, Free University Berlin, Verlag Dr. Hut Munich, 2006.

[3] M. Weber and S. Kube, *Robust Perron Cluster Analysis for Various Applications in Computational Life Science*, in Computational Life Sciences, First International Symposium, CompLife 2005, Konstanz, Germany, September 25-27, 2005, Proceedings, M. R. Berthold, R. C. Glen, K. Diederichs, O. Kohlbacher, and I. Fischer, eds., vol. 3695 of Lecture Notes in Computer Science, Springer, 2005, pp. 57–66.

Condition numbers for structured least squares problems

Yimin Wei
School of Mathematical Sciences, Fudan University
Shanghai, 200433, P.R. of China
ymwei@fudan.edu.cn

Sanzheng Qiao
Department of Computing and Software, McMaster University
Hamilton, Ontario, L8S 4L7, Canada
qiao@mcmaster.ca

Wei Xu
Department of Computing and Software, McMaster University
Hamilton, Ontario, L8S 4L7, Canada
xuw5@mcmaster.ca

This paper studies the normwise perturbation theory for structured least squares problems. The structures under investigation are symmetric, persymmetric, skewsymmetric, Toeplitz and Hankel. We present the condition numbers for structured least squares problems.

Preconditioning interior point methods for PDE optimization and identification

Martin Weiser
Zuse Institute Berlin
Takustr. 7, D-14195 Berlin, Germany
weiser@zib.de

Anton Schiela
Zuse Institute Berlin
Takustr. 7, D-14195 Berlin, Germany
schiela@zib.de

In this talk we will introduce function space oriented interior point methods for PDE constrained optimization and work out two algorithmic variants: control reduction for simple control constrained problems and control discretization for identification problems with H^1-regularization. The specific structure of the KKT systems arising in these algorithmic variants are pointed out, and block preconditioners for the different cases are proposed and analyzed.
Numerical examples, both artificially and from identification problems arising in the cancer therapy regional hyperthermia, are given in order to illustrate the practical effectiveness of the preconditioners.

A MRRR algorithm for the bidiagonal SVD

Paul R. Willems
Fachbereich C – Mathematik und Naturwissenschaften, Bergische Universität Wuppertal
Gauß-Str. 20, D-42097 Wuppertal, Germany
willems@math.uni-wuppertal.de

For computing the singular value decomposition (bSVD) of a real square bidiagonal matrix $B \in \mathbb{R}^n$ the standard approach is to reduce the problem to the symmetric tridiagonal eigenproblem (tSEP). To this end, one can use the so-called *normal equations* B^TB and BB^T or the Golub-Kahan-Matrix T_{GK}, which is a symmetric tridiagonal matrix of dimension $2n$ resulting from permuting the Jordan-Wielandt-Form of B.
For tSEP, a valuable theoretical improvement has been made by the quite recent discovery of the algorithm of Multiple Relatively Robust Representations (MRRR, or MR^3) by Dhillon and Parlett [1, 2, 3], which can be understood as a variant of inverse iteration [7]. For a symmetric tridiagonal matrix of dimension n, the MRRR algorithm allows to compute k accurate eigenpairs in $\mathcal{O}(kn)$ time and has the additional benefit of being naturally suited for parallelization.
As Großer pointed out in his thesis [4], due to numerical problems it is not sufficient simply to apply MRRR to T_{GK} or B^TB and BB^T in a black box fashion in order to solve the bSVD. To overcome these problems, Großer and Lang proposed so called *coupling relations* [5, 6] to adapt the MRRR algorithm, without spoiling the SVD of B.
In an ongoing research project we are working on the refinement and extension of this approach. So far the result can be favorably compared to the Divide & Conquer algorithm [8]. Further research focuses on optimization, theoretical foundation and simplification of the method. It is planned that the algorithm will be incorporated into the next release of the LAPACK library.
The first part of this talk will give a gentle overview of the MRRR algorithm and the coupling approach for the bSVD. The second part will delve deeper into theoretical background and techniques for practical realisation of the algorithm. Finally, we will present up to date results of numerical tests comparing our implementation to the QR and Divide & Conquer implementations from LAPACK.

Acknowledgements

This presentation covers joint work with Bruno Lang. We wish to thank James Demmel, Osni Marques, Beresford Parlett and Christof Vömel for letting us profit from the not yet published improvements in the MRRR algorithm, its implementation and its testing environment.

References

[1] I. S. Dhillon, *A new $O(n^2)$ algorithm for the symmetric tridiagonal eigenvalue/eigenvector problem*, PhD thesis, University of California, Berkeley, 1997.

[2] I. S. Dhillon and B. N. Parlett, *Multiple representations to compute orthogonal eigenvectors of symmetric tridiagonal matrices*, Linear Algebra and its Applications, 387 (2004), pp. 1–28.

[3] ———, *Orthogonal eigenvectors and relative gaps*, SIAM Journal on Matrix Analysis and Applications, 25 (2004), pp. 858–899.

[4] B. Grosser, *Ein paralleler und hochgenauer $O(n^2)$ Algorithmus für die bidiagonale Singulärwertzerlegung*, PhD thesis, Bergische Universität Wuppertal, Fachbereich Mathematik, Wuppertal, Germany, 2001. In German.

[5] B. Grosser and B. Lang, *An $O(n^2)$ algorithm for the bidiagonal SVD*, Linear Algebra and its Applications, 358 (2003), pp. 45–70.

[6] ———, *On symmetric eigenproblems induced by the bidiagonal SVD*, SIAM Journal on Matrix Analysis and Applications, 26 (2005), pp. 599–620.

[7] I. C. F. Ipsen, *Computing an eigenvector with inverse iteration*, SIAM Review, 39 (1997), pp. 254–291.

[8] P. R. Willems, B. Lang, and C. Vömel, *Computing the bidiagonal SVD using multiple relatively robust representations*. To appear in SIAM Journal on Matrix Analysis and Applications. (Also as LAPACK Working Note #166.).

Computing accurately ordered PageRank scores

Rebecca S. Wills
Department of Mathematics, North Carolina State University
Box 8205, Raleigh, NC 27695, USA
rmwills@ncsu.edu

To determine the order in which query results are displayed, Google's search algorithm calculates a PageRank score for each webpage. These scores correspond to elements of the stationary distribution for a large Markov chain based on the hyperlink structure of the Web. Computing the unique stationary distribution requires extensive time and computing resources, so Google calculates successive approximations to the stationary distribution using an iterative method, such as the power method. The elements of the final approximation are the PageRank scores used by Google. Determining when to stop the iterations requires deciding when the computed approximation is good enough. A popular criterion for terminating iterations is based on the residual norm. We present reasons to consider termination criteria based on the ordering of approximated scores instead. We look at the ordering of elements of the stationary distributions corresponding to the Markov chains of simple directed graphs, and we show that the number of iterations required to achieve accurately ordered PageRank scores can be much higher or much lower than the number of iterations required using a residual norm. We also show that a comparison of ordered scores for two successive iterates is not an effective termination criterion since the ordering of scores is not stable, meaning that correct ordering can occur in a certain iteration and be destroyed in the next. We conjecture that the iteration count to achieve a correct ordering depends on the connectivity of the graph, in particular, the out degree and the number and size of the strongly connected components.

Structured matrices and polynomial interpolation

Iwona Wróbel
Faculty of Mathematics and Information Science, Warsaw University of Technology
Pl. Politechniki 1, 00-661 Warsaw, Poland
wrubelki@op.pl

Alicja Smoktunowicz
Faculty of Mathematics and Information Science, Warsaw University of Technology
Pl. Politechniki 1, 00-661 Warsaw, Poland
smok@mini.pw.edu.pl

The problem of constructing efficient algorithms for polynomial interpolation involves matrices with special structure. Numerical algorithms should exploit the structure of these matrices. We give a detailed componentwise error analysis of certain algorithms, e.g. Goertzel's algorithm, which is frequently used in the interpolation and approximation problems and in signal processing. It is implemented in Matlab, in the Signal Processing Toolbox. However, Goertzel's algorithm is much less numerically accurate than the Fast Fourier Transform (FFT). In order to improve accuracy we propose a modification of Goertzel's algorithm based on the divide-and-conquer techniques.
We also present new backward stable algorithms for construction of the Newton divided differences and for evaluation of the Lagrange representation of interpolating polynomial.
Numerical tests in Matlab, performed for various distributions of interpolation points, such as equally spaced, Chebyshev and fast Leja points, demonstrate the computational advantages of proposed algorithms.

On the use of sparse direct solver in a projection method for generalized eigenvalue problems using numerical integration

Yusaku Yamamoto
Department of Computational Science & Engineering, Nagoya University
Furo-cho, Chikusa, Nagoya, Aichi, 464-8603, Japan
yamamoto@na.cse.nagoya-u.ac.jp

Takamitsu Watanabe
Department of Computational Science & Engineering, Nagoya University
Furo-cho, Chikusa, Nagoya, Aichi, 464-8603, Japan
taka-w@na.cse.nagoya-u.ac.jp

We consider the solution of the generalized eigenvalue problem

$$A\mathbf{x} = \lambda B\mathbf{x}, \tag{1}$$

where A and B are sparse symmetric matrices and B is positive definite. Such problem arises in structural analysis and quantum chemistry and we often need eigenvalues within a specified interval.
Recently, Sakurai and Sugiura [2] proposed a new algorithm that is especially suited for this purpose. Their method projects the original problem to a smaller generalized eigenvalue problem associated with the wanted eigenvalues using the complex integration of a resolvent function $f(z) = \mathbf{u}^H(zB - A)^{-1}\mathbf{v}$, where $\mathbf{u}$ and $\mathbf{v}$ are suitably chosen vectors. The integration is computed by the trapezoidal rule and most of the work is spent for evaluating the function $f(z)$ at the sample points. This involves solving multiple linear equations $(z_iB - A)\mathbf{x}_i = \mathbf{v}$, where z_i's are the sample points. Since each equation can be solved independently, this method has large grain parallelism.
In the work of Sakurai et al., the COCG solver with incomplete (complex) Cholesky preconditioner has been used to solve these equations. However, it has been observed that the number of iterations varied significantly from one point to another, leading to load imbalance among the processors. In addition, if one wants to use more processors than the number of sample points, one has to parallelize the solver itself. However, this type of solver is difficult to parallelize.
In this study, we apply a sparse direct solver to the solution of $(z_iB - A)\mathbf{x}_i = \mathbf{v}$. Because the coefficient matrix is complex symmetric, theoretically, pivoting is necessary to ensure numerical stability [1]. However, we used a straightforward extension of the sparse Cholesky factorization to complex symmetric matrices and omitted pivoting. This makes it possible to solve the equation for each sample point in roughly the same time and also permits us to parallleize the solver itself using ordering algorithms such as the nested dissection. To check numerical stability, after the complex Cholesky factor has been computed, we calculate the growth factor and verify that it is not too large.
We applied our method to several generalized eigenvalue problems taken from the Harwell-Boeing Matrix Library and from real molecular orbital applications. We found that in most cases, the growth factor was modest and the linear equations can be solved safely without pivoting. The computed eigenvalues agreed with those computed using the (complex) Gaussian elimination with partial pivoting to at least eight digits. In the presentation the parallel performance of our method will also be reported.

References

[1] G. H. Golub and C. Van Loan, *Matrix Computations, 3rd Edition*, The Johns Hopkins University Press, Baltimore, MA, 1996.

[2] T. Sakurai and H. Sugiura, *A projection method for generalized eigenvalue problems using numerical integration*, Journal of Computational and Applied Mathematics, 159 (2003), pp. 119–128.

A constrained minimization algorithm for electronic structure calculation

Chao Yang
Computational Research Division
Lawrence Berkeley National Laboratory
1 Cyclotron Rd, Berkeley, CA, USA
cyang@lbl.gov

Juan Meza, Lin-Wang Wang
Computational Research Division
Lawrence Berkeley National Laboratory
1 Cyclotron Rd, Berkeley, CA, USA
JCMeza@lbl.gov, LWWang@lbl.gov

One of the fundamental problems in electronic structure calculation is to determine electron orbitals associated with the minimum total energy of large atomistic systems. The total energy minimization problem is often formulated as a nonlinear eigenvalue problem and solved by an iterative scheme called Self Consistent Field (SCF) iteration. In this talk, a new direct constrained optimization algorithm for minimizing the Kohn-Sham (KS) total energy functional is presented. The key ingredients of this algorithm involve projecting the total energy functional into a sequences of subspaces of small dimensions and seeking the minimizer of total energy functional within each subspace. The minimizer of the projected energy functional not only provides a search direction along which the KS total energy functional decreases but also gives an optimal "step-length" to move along this search direction. Due to the small dimension of the projected problem, the minimizer of the projected energy functional can be computed by several different methods. These methods will be examined and compared in this talk. Numerical examples will be provided to demonstrate that this new direct constrained optimization algorithm can be more efficient and robust than the SCF iteration.

Linear algebra problems for simulating the railway vehicle-track dynamic interactions

Akiyoshi Yoshimura
School of Computer Science, Tokyo University of Technology
1404-1, Katakura, Hachioji, Tokyo, Japan
yosimura@cc.teu.ac.jp

Linear algebra problems for simulating the railway vehicle-track dynamic interactions are described. Firstly a mathematical linear algebraic structure of the model based on the finite element method is developed. In this representation, the relation between the running vehicle and the exciting force internally generated by the contact of the wheel and the rail is formulated precisely and as a vehicle moves on the track, a time-invariant part and a time-depending part in the algebraic expression are clearly perceived. A time-variant part is simply expressed by the equation $\boldsymbol{x}\boldsymbol{x}^T$ using specially defined vector $\boldsymbol{x}$, and occupies only small space while they move in the large-scale sparse symmetric positive definite matrix. Then, under these circumstances, to execute efficiently numerical simulations in the time domain, various numerical methods of explicit and/or implicit schemes such as Newmark method , Runge-Kutta method etc., are comparatively investigated from the point of view of stability and accuracy.
Lastly a special model of the forced vibration experiments of vehicle-track interaction excited by the actuator is studied. In this case a differential-algebraic equation is obtained as a result of simulating the experiment. A method of obtaining the numerical stable solution for this ia shown.

A singular homogeneous quasilinear system of partial differential equations

M. Ziad
Department of Mathematics and Statistics, Sultan Qaboos University, Oman
mziad@squ.edu.om

A system of the type $\mathbf{A}\mathbf{w}_t + \mathbf{B}\mathbf{w}_x + \mathbf{C}\mathbf{w}_y + \mathbf{D}\mathbf{w}_z = \mathbf{0}$, where $\mathbf{w} = [f_1, f_2, f_3, f_4, f_5, f_6, f_7, 0, 0, 0]^t$, $f_i = f_i(t, x, y, z)$ and the 10×10 singular matrices, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ are functions of $\mathbf{w}$ only, is considered. The singularity of matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ prohibits to use the existing methods of characteristics. The other methods (if exist) to solve this type of a homogeneous quasilinear system are not well known. Here, the use of linear and Lie algebraic methods is demonstrated to solve such type of systems. A list of solutions of the system is provided.

Applications and algorithms for the polar decomposition of a matrix

Krystyna Ziętak
Wrocław University of Technology, Institute of Mathematics and Computer Science
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
Krystyna.Zietak@pwr.wroc.pl

Beata Laszkiewicz
Wrocław University of Technology, Institute of Mathematics and Computer Science
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
Beata.Laszkiewcz@pwr.wroc.pl

The *polar decomposition* of a complex matrix $A \in C^{m\times n}, m \geq n$, is defined as follows: $A = EH$, where H is a Hermitian positive semidefinite matrix of order n and $E \in C^{m\times n}$ is a *subunitary matrix* (*partial isometry*), $EE^HE = E$. If A is not full rank then E is not unique.
Some approximation properties and applications of the polar factors E and H are well known. The unitary polar factor E of a nonsingular matrix A is the best approximation of A by unitary matrices with respect to all unitarily invariant norms. The approximant of A by means of Hermitian positive semidefinite matrices, with respect to the Frobenius norm, can be expressed by means of the Hermitian polar factor of the real part of A. In the talk we show new applications of polar factors to approximation by subunitary matrices and minimal rank approximation (see [3]).
The polar decomposition can be computed directly from the `SVD` of A. However, an alternative approach, based on iterative methods, is also interesting and convenient. We consider the family of Gander methods [1] and Higham's scaled method [2]. We show how these iterative methods can be applied for computing subunitary and minimal rank approximants of A (see [3]).
The family of Gander methods depends on a parameter f. It is shown that under some assumptions on A and for a suitable parameter f Gander methods can be applied to computing subunitary and minimal rank approximants of A. Unfortunately, these assumptions can not be verified easily in practice, hence Gander methods are not suitable for our purpose in the general case.
We present numerical tests comparing different algorithms for computing subunitary and minimal rank approximants. Additionally we compare the efficiency of the Gander and Higham algorithms for computing unitary polar factor with the efficiency of the `SVD`.

References

[1] W. Gander, *Algorithms for the polar decomposition*, SIAM J. Sci. Stat. Comput., 11 (1990), pp. 1102–1115.

[2] N. J. Higham, *Computing the polar decomposition — with applications*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 1160–1174.

[3] B. Laszkiewicz and K. Ziętak, *Approximation of matrices and family of Gander methods for polar decomposition*, to appear in BIT Numer.Math.

Updated Arnoldi factorization for derogatory and defective matrices

Jan Zítko
Charles University, Faculty of Mathematics and Physics, Department of Numerical Mathematics, Prague
Sokolovská 83, 186 75 Praha 8, Czech Republic
zitko@karlin.mff.cuni.cz

Iveta Hnětynková
Charles University, Faculty of Mathematics and Physics, Department of Numerical Mathematics, Prague
Sokolovská 83, 186 75 Praha 8, Czech Republic
hnetynkova@cs.cas.cz

Let A be an $n \times n$ matrix and let us consider the process which iteratively updates the starting vector in the Arnoldi factorization via shifted QR-iteration. The process yields the sequences of starting vectors, corresponding Krylov subspaces and Hessenberg matrices. The iteration can be organized such that the repetition of the process yields Krylov subspaces approximating an invariant subspace associated with a few wanted (usually small) eigenvalues of A quite well. We present two main results. First we prove that the angels between udated starting vectors and the wanted invariant subspace tends to zero. Then we consider derogatory and defective matrices. We prove that if the Jordan canonical form of the matrix A has more than one block corresponding to the "smallest" eigenvalue with the maximal dimension d then the elements in the position $(d+1, d)$ in the sequence of Hessenberg matrices tend to zero. Numerical examples supporting the theoretical results conclude the paper.

The Spectrum of a Fourier migration matrix pencil

Petr Zizler
Mount Royal College
4825 Mount Royal Gate, Calgary, Alberta, Canada, T3E 6K6
pzizler@mtroyal.ca

Consider a scalar wavefield $\psi(x, z, t)$ where x and z are spacial variables and t is time. Consider the wave equation

$$\nabla^2\psi - \frac{1}{v(x)^2}\frac{\partial^2\psi}{\partial t^2} = 0$$

where the velocity function $v = v(x)$ (a function of the lateral position x only) is assumed to have nonzero terminal velocities in both directions. When Fourier migration techniques are applied to solve the wave equation the following $(2N+1)\times(2N+1)$ hermitian matrix pencil is encountered (up to a sign)

$$A_N(w) = D_N + w^2 L_N \text{ with } w \in \mathbf{R}$$

The matrix L_N is a truncated Laurent matrix of the matrix representation of a certain convolution operator L and D_N is a diagonal matrix of the same size with

$$D_N(n,n) = (n-1-N)^2 \text{ for } n = \{1, 2, \ldots, 2N+1\}$$

The spectrum of the matrix pencil $A_N(w)$ evolves both with the truncation parameter N and the temporal parameter w. Asymptotic results in this direction dictate the accuracy of the migration techniques. We provide spectral asymptotic results for the pencil $A_N(w)$ and show the migration techniques are significantly more accurate if the terminal velocities coincide. Symmetry properties of eigenvectors of the pencil $A_N(w)$ are also presented.

Symmetric indefinite preconditioners for saddle point problems

Walter Zulehner
Institute of Computational Mathematics, Johannes Kepler University Linz
Altenberger Str. 69, A-4040 Linz, Austria
zulehner@numa.uni-linz.ac.at

In this talk we consider large scale sparse linear systems of the typical saddle point form

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix},$$

where A is a symmetric matrix and B is a matrix with full rank. Such problems typically result, e.g., from the discretization of optimization problems with pde-constraints.
A natural property in this context is the positivity of A on the kernel of B only:

$$(Ax, x) > 0 \quad \text{for all } 0 \neq x \in \ker B.$$

Under this assumption the Schur complement of the system does not necessarily exist and several standard approaches, which are well-developed for positive definite A, are not directly applicable.
One way to enforce the positivity, is to use an augmented Lagrangian approach by replacing A by a matrix like $A_\rho = A + \rho\, B^T B$. However, the adjustment of the parameter ρ is a critical issue.
We will present a different approach that is not based on such an explicit augmentation technique. A symmetric and indefinite preconditioner will be constructed and analyzed. Under appropriate conditions the preconditioned saddle point system is symmetric and positive definite with respect to a particular scalar product. Therefore, conjugate gradient acceleration can be used.
If applied to a typical pde-constrained optimization problem it can be shown that the convergence rate for solving the discretized problem with the proposed preconditioned conjugate gradient method is independent of the mesh size of the underlying mesh. Numerical experiments are presented for such a model problem for illustrating the theoretical results.

Author index