IV International Workshop on
Accurate Solution of Eigenvalue Problems
Split, Croatia, June 24-27, 2002

BOOK OF ABSTRACTS

Main Lectures

Singular Value Decay for Solutions to Matrix Equations

by Christopher A. Beattie

The matrix solutions, X, to either Sylvester equations: $AX + XB = G$ or Ricatti equations: $XHX + AX + XB = G$ occasionally allow good low rank approximations when the nonhomogeneous term G itself has low rank. We discuss circumstances where this does (or does not) happen and what consequences are borne to the approximation of invariant subspaces and model reduction and control of large scale dynamical systems.

The Complexity of Accurate Floating Point Computation or Can we do Numerical Linear Algebra in Polynomial Time?

by James W. Demmel, Plamen Koev and Ben Diament

Our goal is *accurate* and *efficient* floating point (FP) computation. Accurate means that we guarantee a relative error less than 1, i.e. that some leading digits are correct, and zero is exact. Efficient means computable in polynomial time in the size of the input. We use the abbreviation CAE to mean "compute accurately and efficiently." More specifically, we would like to answer the following questions.

- What FP expressions can we CAE?
- Are there FP expressions that we cannot CAE?
- For which structured matrices (matrices with FP expressions as entries) can we perform matrix computations accurately and efficiently? The matrix computations we consider include the determinant, inverse, triangular factorization, QR factorization, SVD etc.

The answers to these questions depends strongly on two decisions: The set of numbers permitted to appear in the expressions, and the model of floating point arithmetic. The possible sets of numbers are

- positive real numbers
- all real numbers
- complex numbers

The models of arithmetic are

- The traditional $1 + \delta$ model, but where δ and all numbers are real or complex, as opposed to discrete values represented by bits.
- The usual bit-model, where FP numbers are presented by a pair of integers (e, f) representing $f \cdot 2^e$. However, e is limited in size compared to f in a certain way. We call this the "small exponent" (SE) model.
- The usual bit-model as above, but now the exponent e is allowed to be any integer, not constrained by the size of f. We call this the "large exponent" (LE) model.

It turns out that the answers to our questions depend very strongly on which choice of numbers and which choice of arithmetic models one makes. The *same* problem can change from doable in polynomial time, to exponential time, to impossible depending on our choices.

Depending on our choices, we classify FP expressions according to their factorizability properties to decide whether they can be CAE. In the case of matrix computations, the answer depends on the factorizability properties of minors.

In addition to systematizing the many apparently ad hoc results on high accuracy algorithms produced by the authors and collaborators over the past years, a number of new classes have been identified. For example, by exploiting the combinatorial properties of Schur functions (which appear as the largest irreducible factors of determinants of generalied Vandermonde matrices) we have new efficient and accurate algorithms for matrix computations on totally positive generalized Vandermonde matrices. We have also new SVD algorithms for some polynomial Vandermonde matrices and M-matrices, which will be discussed in a separate talk.

Rank Deficient Problems and Numerical Ranks with Known Structure

by Zlatko Drmač

We discuss two classes of matrix eigenvalue (singular value) problems where rank deficiency is expressed by additional data on input: either by a null space basis or by additive perturbation of given structure.

In certain applications, symmetric semidefinite matrix A, of rank r, is given in implicit form (e.g. natural factor formulation in finite element computation) as $A = FF^T$, where the matrix F is of full column rank. In this way, F exposes the singularity of A and enables elegant computation of its nontrivial eigenelements. Another way to have explicit singularity is to accompany the matrix A with a basis Y of its null space. This is the situation e.g. in computational electromagnetics where explicit null space enforces divergence–free discrete electric field. We study accurate (generalized) eigenvalue and singular value computation in such situations. In particular, we analyze numerical stability of the Cholesky and QR factorizations, and the implicit Shur complement computation. This line of research opens various problems related to *geometric conditions* for accurate eigenvalue and singular value computations. This presentation is the result of joint work with Peter Arbenz.

In our second theme, we analyze the numerical rank of general matrices in presence of structured perturbations of prescribed rank (by static filters). The theoretical and computational tool is the restricted SVD. We describe new algorithm for computing the restricted singular values with high relative accuracy. Some techniques in accurate computations of various forms of the generalized singular value decompositions will be presented. We discuss the structure of an universal algorithm that follows the generalization tree of the O–P–Q–R-SVD's.

Absolute and Relative Inclusion Regions for Matrix Eigenvalues

by Ilse C.F. Ipsen

An 'inclusion region' for matrix eigenvalues is a region in the complex plane guaranteed to contain a certain number of eigenvalues. Gerschgorin discs and residual bounds are well-known examples of inclusion regions, containing respectively, all eigenvalues and at least one.

Here we review Lehmann's inclusion regions for Hermitian matrices and extend them to general, non-normal matrices. We present two types of inclusion regions, absolute (based on absolute distances of the eigenvalues from a given point) and relative (based on relative distances). Absolute inclusion regions are disks whose radii are expressed in terms of singular values of the restriction of the matrix to a subspace, and the condition number of an eigenvector matrix or the departure of the matrix from normality. Analogously, relative inclusion regions are disks whose radii depend on generalized singular values. We show that the absolute inclusion regions are optimal for normal matrices.

We also derive a number of extensions: absolute inclusion regions that are unbounded exteriors of disks; annular exclusion regions that are guaranteed to exclude a given number of eigenvalues; and relative inclusion intervals for eigenvalues of positive definite matrices with a known factorization.

This is joint work with Christopher Beattie.

Accurate Solution of Quadratic Eigenvalue Problems with Structure

by Volker Mehrmann

Numerical methods for the solution of quadratic eigenvalue problems with structure are discussed. We discuss in particular problems with skew Hamiltonian/Hamiltonian structure or perturbed versions of such problems, as they arise in singularity exponents in elasticity problems or in damped gyroscopic systems. For the solution of these problems it is essential to use methods that preserve the structure, since otherwise meaningless may occur.

We discuss a structure preserving linearization of the quadratic problem, an implicitely restarted skew Hamiltonian Arnoldi-like method including a new extraction procedure for the computation of eigenvectors and invariant subspaces of skew-Hamiltonian/Hamiltonian pencils.

We also discuss several application problems.

This is joint work with Thomas Apel, Tsung-Min Hwang, Wen-Wei Lin and David Watkins.

Rational Krylov for Nonlinear Eigenvalues

by Axel Ruhe

This is a rather intricate algorithm that I described in a volume dedicated to my academic mother Vera Nikolaevna Kublanovskaya on the occasion of her 80-th birthday on September 21, 2000. An eigenvalue problem that is nonlinear in the eigenvalue parameter is solved by succesive linearizations with Lagrange interpolation. Each such generalized linear eigenproblem is solved with a rational Krylov approach. An interesting difference, compared to linear eigenproblems, is that now both the numerator and the denominator of the rational function are of interest. The numerators are the matrix evaluated at a sequence of shifts which are updated in each iteration step, in order to converge to the latent roots one by one. When the shift is at a latent root, the solution of a standard eigenvalue problem gives the corresponding latent vector. The denominators are chosen at poles and are changed much more seldom, since each new pole needs a new matrix factorization. I will describe the algorithm and give some examples from nonlinear mechanical systems. It has been implemented by Patrik Hager in his thesis in Structural Mechanics at Chalmers.

Inverse Iteration, Newton and Jacobi-Davidson Type Methods for Computing Eigenvalues and Singular Values

by Hubert Schwetlick

Consider a matrix $A \in \mathbb{R}^{n\times n}$, and let λ_* be an algebraically simple eigenvalue with normalized right and left eigenvector x_*, y_*, resp. Given an approximation (u, μ), $u^*u = 1$, to (x_*, λ_*), we are looking for improved values (u_+, μ_+). The standard approach for obtaining an improved eigendirection u_+ consists in using inverse iteration with shift μ, i. e., in solving $(A - \mu I)u_+ = u\alpha$ (1) with an appropriate scaling factor $\alpha \neq 0$. If the Ritz value $\mu = \mu(u) = u^*Au$ is taken in (1) one obtains the nonstationary Rayleigh quotient iteration. Note that $A - \mu I$ becomes almost singular if μ approaches λ_*. Whereas backward stable direct methods as pivoted Gaussian elimination work well on (1), cf. Peters/Wilkinson [79], iterative solvers may fail so alternative approaches that use system matrices with uniformly bounded inverses are of interest.

One possibility proposed by Unger [50] consist in applying one Newton step $u_+ = u + s$, $\mu_+ = \mu + \alpha$ from (u, μ) to the nonlinear system $Ax - \lambda x = 0$, $w^*x = 1$ (2) in (x, λ) where w is an appropriate scaling vector. This leads to $(A - \mu I)u_+ = u\alpha$, $w^*u_+ = 1$, i. e., the u_+ is, up to scaling, the same as in the inverse iteration (1). An other way is taking $w = u$ and computing the Newton correction (s, α) from the Newton equation $(A - \mu I)s - u\alpha = -(A - \mu I)u =: -r$, $u^*s = 0$ (3). The system matrix $\left[\begin{smallmatrix}(A-\mu I) & -u\\ u^* & 0\end{smallmatrix}\right]$ has a uniformly bounded inverse for sufficiently good (u, μ) so (preconditioned) iterative solvers will work, in general. If, in addition, μ is taken as Ritz value $\mu = u^*Au$ which is equivalent to $u^*r = 0$, multiplying the left equation (3) by u^* leads to $\alpha = u^*(A - \mu I)s$. Putting this into (3) one gets $(I - uu^*)(A - \mu I)(I - uu^*)s = -r$, $u^*s = 0$ (4), which is just the Jacobi-Davidson correction equation, see Sleijpen/Van der Vorst [96]. Hence, the Jacobi-Davidson correction s_{JD} defined as unique solution of (4) is identical with the x-part s_N of the Newton correction (s, α) defined by (3), and working with the projected matrix $(I - uu^*)(A - \mu I)(I - uu^*)$ on the subspace $\{s : u^*s = 0\}$ is equivalent to using the augmented matrix $\left[\begin{smallmatrix}(A-\mu I) & -u\\ u^* & 0\end{smallmatrix}\right]$ on the non restricted product space $\{s, \alpha\}$. Whereas Newton continues with one single approximation $u_+ = u + s$, eventually normalized, and $\mu_+ = \mu + \alpha = u^*Au_+$ or alternatively with $\mu_+^{alt} = \mu(u_+) = u_+^*Au_+$, Jacobi-Davidson expands the current trial space by the new correction s so it contains u_+, too.

In the talk, several extensions of these basic techniques will be discussed, namely:

- Use of more general and different projectors $I - vw^*/w^*v$ with appropriate v, w,
- Working with nonsymmetrically augmented matrices $\left[\begin{smallmatrix}(A-\mu I) & -v\\ w^* & 0\end{smallmatrix}\right]$,
- Use of generalized Ritz-Petrov values,
- Simultaneous approximation of $q \ll n$ possibly clustered eigenvalues and corresponding invariant subspaces by block versions using $I - V(W^*V)^{-1}W^*$ or else $\left[\begin{smallmatrix}(A-\mu I) & -V\\ W^* & 0\end{smallmatrix}\right]$ with $V, W \in \mathbb{R}^{n\times q}$.

Moreover, the relation of these methods to some Newton techniques used in singularity

theory is investigated, and methods of related structure for approximating one or q smallest singular values and corresponding right and left singular vectors/subspaces are discussed.

Based on joint work with G. Timmermann, R. Lösche, U. Schnabel and M. Hochstenbach.

Contributed Talks and Posters

A Posteriori Error Bounds for the Spectral Approximation of Compact Operators on a Hilbert Space

by Oscar F. Bandtlow

An interesting problem of practical importance is the determination of the eigenvalues of an arbitrary compact operator on a Hilbert space. While there exist a number of useful approximation schemes, which effectively reduce the problem to the accurate solution of finite-dimensional eigenvalue problems, less is known about how to find readily computable upper bounds for the error of these approximations.

I shall present an approach to this problem which relies on an extension of a formula of Ostrowski giving an upper bound for the Hausdorff distance of the set of eigenvalues of two matrices.

As an application I shall consider the calculation of the eigenvalues of an operator arising in the theory of dynamical systems, the spectral data of which yields insight into the statistical long-time behaviour of a given dynamical system.

A Parallel $O(n^2)$ Algorithm for the Computation of a Subset of Eigenvectors of a Symmetric Tridiagonal Matrix

by Paolo Bientinesi and Inderjit Dhillon

In 1997 Dhillon & Parlett presented a new $O(n^2)$ algorithm for computing the eigenvectors of a symmetric tridiagonal matrix T ([1]). This algorithm proceeds by computing representations LDL^T of translates of T. Given the eigenvalues λ_i of T, with $\lambda_i \leq \lambda_{i+1},\ i = 1, \ldots, n-1$, the computation path of this algorithm depends on the relative gaps between the eigenvalues, i.e. $relgap(\lambda_i, \lambda_{i+1}) = |\lambda_i - \lambda_{i+1}|/|\lambda_i|$. Whenever this relative gap is large enough ($relgap > \min(10^{-2}, 1/n)$) the i-th eigenvector is directly computed using twisted factorizations; otherwise we are in the presence of a cluster C of eigenvalues and a new representation $LDL^T = T - \sigma I$ needs to be computed. The shift σ is chosen in such a way that (a) the new representation is robust relatively to the eigenvalues in C and (b) at least one of the shifted eigenvalues in C is relatively well separated from the others. The process is iterated if other clusters are encountered. The sequence of computations can be pictorially expressed by a *representation tree* which contains information about how the eigenvalues are clustered (nodes of the tree) and what shifts are used to break a cluster (edges of the tree). If the input matrix T is reducible, then the former algorithm is applied to each of the unreduced blocks B_i.

We tackle the problem of computing the eigenvectors of T or a subset of them in a parallel system with no shared memory. The total number of operations for the sequential algorithm to compute all the eigenvectors is $c(T) = O(\sum_i n_i^2)$ where n_i is the size of B_i and the amount of memory needed is $O(n^2)$. This represents a major limitation for the use of sequential code with very large matrices. Our goal is to achieve good load balancing among processors in terms of operation and memory requirement on matrices which have substantially different characteristics: large number of blocks, large number of clusters, nested clusters, very few clusters, etc. Also, we want to minimize the communication in order to enhance the scalability of the algorithm.

For this we come up with a parallel model where each processor keeps a queue of clusters which contain a superset of the eigenvalues relative to the eigenvectors to be computed locally. Each processor computes N/p eigenvectors (N = number of eigenvectors sought) and allocates only $O(N^2/p)$ memory. The granularity of the computation consists of the blocks for the first step and clusters for all the later stages.

The idea is to locally descend the representation trees (there is one for each block) in a breadth first fashion to fragment the clusters until all the eigenvectors relative to a cluster have to be computed by the local processor. At that stage the sequential algorithm can be invoked. Clusters containing eigenvalues whose eigenvectors are associated to other processors are removed from the queue.

We show numerical results to demonstrate the accuracy and speedup obtained by our parallel algorithm. We also report on the latest experiments performed on the sequential algorithm on very large matrices ($n > 10,000$) that arise as part of finite element models

in automobile design.

References

[1] I. S. Dhillon. *A new $O(n^2)$ Algorithm for the Symmetric Tridiagonal Eigenvalue/Eigenvector Problem*. PhD thesis, Computer Science Division, University of California, Berkeley, California, May 1997.

Sensitivity Analysis of Eigenvalues Based on ϵ-pseudopsectra of Matrices

by Shreemayee Bora and Rafikul Alam

Eigensystems of matrices are generally highly sensitive to perturbation (see, for example, [5], [6]). It is well known that often a small perturbation may cause highly sensitive eigenvalues to move and coalesce. Once coalescence takes place, in the new dispensation, the *successors* of these eigenvalues may acquire a set of entirely different properties by losing their original identities. The magnitude of the smallest perturbation for which an eigenvalue λ moves and coalesces with a neighbouring eigenvalue is referred to as the *dissociation* of the eigenvalue λ from the rest of the spectrum of A and is denoted by $diss(\lambda)$ [2]. Evidently, the value of $diss(\lambda)$ provides a characterisation of the sensitivity of the eigenvalue λ. However, the exact value of $diss(\lambda)$ which depends on the matrix as well as on the underlying vector space is almost impossible to know except for some special cases (see [1]).

We obtain lower bounds of $diss(\lambda)$ which are easy consequences of inclusion theorems for ϵ-pseudospectra for a large class of norms. The lower bounds of $diss(\lambda)$ thus obtained for the ℓ^1 and ℓ^∞ norms may be computed quite cheaply. The lower bound for the ℓ^2 norm is relatively cheaper to compute than the bound due to Demmel included in the software package LAPACK and compares well with it. We illustrate this by means of examples. We also discuss the special case of the 2-by-2 matrix. These bounds involve the notions of separation of matrices namely *sep* and sep_λ (see [2], [3]). We extend these notions to all operator norms and investigate their various properties as well as relations between them. In particular, we prove a conjecture due to Demmel [2] under appropriate assumptions.

References

[1] S. Bora, *A Geometric Analysis of Spectral Stability of Matrices and Operators,* Ph.D. Thesis, Indian Institute of Technology, Guwahati, July 2001.

[2] J. W. Demmel, Computing stable eigendecompositions of matrices, *Linear Algebra & Appl.*, 79(1986), pp.163-193.

[3] G. W. Stewart and J. Sun, *Matrix Perturbation Theory,* Academic Press, San Diego, 1990.

[4] L. N. Trefethen, Pseudospectra of matrices, in D. F. Griffiths and G. A. Watson, Numerical Analysis 1991, Longman, (1992), pp. 234-266.

[5] J. H. Wilkinson, Sensitivity of eigenvalues, *Utilitas Mathematica,* 25(1984), pp. 5-76.

[6] J. H. Wilkinson, Sensitivity of eigenvalues II, *Utilitas Mathematica,* 30(1986), pp. 243-286.

Accurate Computation of Eigenvalues and Eigenvectors of Tridiagonal Matrices

by Inderjit Dhillon

Recent work has seen a substantial advance, in terms of both speed and accuracy, in the computation of eigenvalues and eigenvectors of symmetric tridiagonal matrices. High accuracy has become possible by the use of bidiagonal factors of indefinite tridiagonals and rank-revealing twisted factorizations (both triangular and orthogonal) of nearly singular tridiagonals. High accuracy in intermediate computations leads to substantial speedups in the overall computation. In this talk, I will highlight these advances and also describe the latest LAPACK software that permits these calculations. Results will be presented on a matrix of size 13,000 that arises from a finite element model of an automobile body.

Computing Accurate Eigenvalues and Eigenvectors of Symmetric Quasi-Cauchy Matrices

by Froilán M. Dopico

Quasi-Cauchy matrices generalize Cauchy matrices via multiplication by diagonal matrices on both sides. These matrices appear frequently in the literature as test matrices and also in applications like rational interpolation problems. Demmel has shown in [1] that it is possible to compute singular value decompositions of quasi-Cauchy with high relative accuracy. We extend these results and give two algorithms to compute spectral decompositions of symmetric quasi-Cauchy matrices with high relative accuracy. The first algorithm consists just in assigning the correct signs to the singular values computed by Demmel's algorithm. Reference [2] has proved that this can be accurately done. The second algorithm begins by computing an accurate enough $G\,J\,G^T$ $(J = I_p \oplus (-I_{n-p}))$ factorization of the quasi-Cauchy matrix, adapting an algorithm in [1] to compute accurate componentwise Schur complements of quasi-Cauchy matrices. Then the J-orthogonal Jacobi-like algorithm of Veselić and Slapničar [4], [3] is applied to the matrices G, J. Theory and numerical experiments of these algorithms are presented.

References

[1] J. Demmel, *Accurate singular value decompositions of structured matrices*, SIAM J. Matrix Anal. Appl., 21 (1999), pp. 562-580.

[2] F. M. Dopico, J. M. Molera, J. Moro, *An orthogonal high relative accuracy algorithm for the symmetric eigenproblem*, sent for publication.

[3] I. Slapničar, *Accurate Symmetric Eigenreduction by a Jacobi Method*, PhD Thesis, Fachbereich Mathematik Fernuniversität, Gesamthochschule Hagen, Germany, 1992.

[4] K. Veselić, *A Jacobi eigenreduction algorithm for definite matrix pairs*, Numer. Math., 64 (1993), pp. 241-269.

Iterative Refinement of the Eigensolution Computed by a Fast Method

by Liljana Ferbar

We consider the eigenvalue problem $Hx = \lambda x$, where H is a real symmetric matrix of order n, λ is the eigenvalue of H and x is the corresponding eigenvector.

As we know, using the computers in solving eigenvalue problems has lead to two aspects of research: speed and accuracy. The first aspect of research is finding faster algorithms and the analysis of their speed of convergence. The second is the question of accuracy: how many accurate digits does the computed eigenvalue have? We know that the accuracy of computed eigenvalue depends on the following: whether the matrix "well behaved", that is, do small relative changes in matrix elements cause small relative changes in eigenvalues, and, if this is the case, on the accuracy of the algorithm which is used to compute the eigenvalues. In general, in order to answer the first question, an appropriate perturbation theory for the given type of problem needs to be developed, while the answer to the second question is given by the numerical analysis of the algorithm.

The aim is, therefore, to solve the problem as quickly and as accurately as possible. However, the existing algorithms can not satisfy both of these two requirements, thus we have to find a compromise, which ultimately depends on our priorities.

To satisfy the speed requirement, Drmač and Veselić suggested to first perform the eigenvalue decomposition $H = U\Lambda U^T$ by using some fast method (QR or Divide-and-Conquer), then to compute the Rayleigh matrix $H' = U^T HU$ and finally to refine the solution by applying the Jacobi method to the Rayleigh matrix. They proved that this algorithm is more efficient (faster) than the Jacobi method on the initial matrix H, while giving the similar relative accuracy a sthe Jacobi method. In this talk we show that we get similar accuracy and faster convergence if in the third step we substitute Jacobi method with one of the fast methods and then iteratively repeat the second and the third step until the accuracy requirement is satisfied.

Based on results of a series of numerical experiments we show that the combination of tridiagonalisation by using Householder's reflectors and D&C method gives us the most efficient (fastest) algorithm, taking into account the accuracy requirement.

This is a joint work with Ivan Slapničar.

Approximating Invariant Subspaces of Elliptic Operators

by Luka Grubišić

We compute lower bounds of the Rayleigh–Ritz approximations of a part of the spectrum of an elliptic operator. Furthermore we present bounds for the accompanying Ritz vectors. The bounds include a form of a relative gap between the Ritz values and the part of the spectrum that is "not wanted".

This is a joint work with Krešimir Veselić.

Accurate Approximate Eigenpairs from Subspaces for the Quadratic Eigenvalue Problem

by Michiel Hochstenbach

We consider the quadratic eigenvalue problem $(\lambda^2 A + \lambda B + C)x = 0$. Suppose that u is an approximation to an eigenvector x (for instance obtained by a subspace method), and that we want to determine an approximation to the corresponding eigenvalue λ. The usual approach is to impose the Galerkin condition $r(\theta, u) := (\theta^2 A + \theta B + C)u \perp u$ from which it follows that θ must be one of the two solutions to the quadratic equation $(u^*Au)\theta^2 + (u^*Bu)\theta + (u^*Cu) = 0$. An unnatural aspect of this approach is that if $u = x$, the second solution has in general no meaning. When u is not very accurate, it may not be clear which solution is the best. Moreover, when the discriminant of the quadratic equation is small, the solution may be very sensitive to perturbations in u.

In this talk we therefore examine alternative approximations to λ. We compare three new approaches theoretically and by numerical experiments. The methods are extended to approximations from subspaces and to the polynomial eigenvalue problem.

Improving a Stability of the Three-way Decomposition

by Ilghiz Ibraghimov

In this article we discuss the decomposition of $A_k \in \mathtt{R}^{n_1 \times n_2}, k = 1, \ldots, n_3$, as $A_k \simeq BE\hat{D}_k C^*$ in the Frobenius norm, where $B \in \mathtt{R}^{n_1 \times r}$ and $C \in \mathtt{R}^{n_2 \times r}$ have normalized columns, E and $\hat{D}_k \in \mathtt{R}^{r \times r}$ are diagonal and $\sum_{k=1}^{n_3} \hat{D}_k^2$ is the identity matrix. This decomposition is widely used in the data processing and is the generalization of the singular value decomposition for the 3 dimensional case. We propose a new algorithm for finding B, C, $\hat{D}_k$ and E if A_k and r are given and B, C have full column rank. If A_k have exact decomposition then this algorithm has a linear convergence.

An implementation of the numerical algorithm was developed, several examples were tested and good results obtained [1, 2]. We present the public domain software package based on our new theory (http://www.ilghiz.com).

References

[1] I. V. Ibraghimov. Application of 3-Way Decomposition for Matrix Compression. (2002) *In Press in J. Num. Lin. Alg.*

[2] V. Orekhov, I. Ibraghimov, M. Billeter, (2001), MUNIN: A new Approach to Multi-dimensional NMR Spectra Interpretation, *Journal of Biomolecular NMR*, v. 20(1), 49–60.

Computing Accurate SVDs of Polynomial-Vandermonde Matrices Involving Orthogonal Polynomials

by James Demmel and Plamen Koev

If P is a basis of orthogonal polynomials satisfying a three-term recurrence $P_0(z) = 1, P_1(z) = d_0(z - b_0)$ and $P_{n+1}(z) = d_n(z - b_n)P_n(z) - c_nP_{n-1}(z)$, where $d_i \neq 0$ for all i, then given a vector $x = (x_1, \ldots, x_n)$ the matrix

$$V_P(x) = \begin{bmatrix} P_0(x_1) & P_1(x_1) & \ldots & P_{n-1}(x_1) \\ P_0(x_2) & P_1(x_2) & \ldots & P_{n-1}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ P_0(x_n) & P_1(x_n) & \ldots & P_{n-1}(x_n) \end{bmatrix},$$

is called Polynomial Vandermonde matrix involving othogonal polynomials.

To compute the SVD of $V_P = W \cdot \Sigma \cdot Z$ to high relative accuracy means the following:

- The singular values are computed with small relative error $|\sigma_i - \hat{\sigma}_i| < O(\epsilon)\sigma_i$.
- The angle $\theta(w_i, \hat{w}_i)$ between the true left singular vector w_i and the computed singular vector $\hat{w}_i$ is bounded by $O(\epsilon)/\text{relgap}_i$, where $\text{relgap}_i = \min_{i \neq j}|\sigma_j - \sigma_i|/\sigma_i$ is the relative gap between σ_i and the nearest other singular value. An analogous statement holds for the computed right singular vectors $\hat{z}_i$.

We present a new $O(n^3)$ algorithm for computing an SVD of V_P to high relative accuracy, which is derived as follows

- We use the identity $V_P(x) = C \cdot V_P(y)$, where $C_{ij} = \frac{a_i f_j}{x_i - y_j}$ is a quasi-Cauchy matrix and the constants a_i and f_j are known accurately.
- If $y_1, \ldots, y_n$ are the roots of P_n, then because of the discrete orthogonality property of the orthogonal polynomials we have $V_P(y) = \text{diag}(\lambda_1, \ldots, \lambda_n) \cdot Q$, where λ_i are the Christoffel numbers and Q is orthogonal. Now $V_P(x) = C \cdot V_P(y) = C\Lambda \cdot Q$.
- We can compute the SVD of the quasi-Cauchy matrix $C\Lambda = W \cdot \Sigma \cdot Z'$ to high relative accuracy by using the algorithm of Demmel (1998).
- The SVD of $V_P(x)$ is $V_P(x) = W \cdot \Sigma \cdot (Z'Q)$, which is computed to high relative accuracy.
- We explain how to accurately compute the roots $y_1, \ldots, y_n$ of P_n and evaluate $V_P(y)$ for the Chebyshev and other classes of orthogonal polynomials and what to do if $x_i = y_j$ for some i and j.

Pole Placement and the QR Algorithm

by Daniel Kressner

Numerical linear algebra has played an active role in studying problems related to control theory. More and more, the other way around gains in importance. As another example we discuss the presence of the single input pole placement problem in the multishift QR algorithm. For a given matrix $A \in \mathbb{C}^{n\times n}$ and shifts $\sigma_1, \ldots, \sigma_m \in \mathbb{C}$ the first column of the shift polynomial is defined as

$$x = (A - \sigma_1 I_n)(A - \sigma_2 I_n) \ldots (A - \sigma_m I_n)e_1.$$

The computation of this vector is required to start an iteration of the multishift algorithm. For an efficient implementation it is desirable to use the QR algorithm with $m = 64$ or 128 shifts; even larger values are appreciated for parallelization efforts. We show that computing x amounts to assigning the shifts as poles to a certain linear time-invariant system using state feedback. Unfortunately the latter problem turns out to be generically ill-conditioned for large m, say larger than 15. One consequence is that in this case the multishift QR algorithm suffers from serious convergence difficulties. We discuss the possibilities to avoid such shift blurring phenomena, for example by tracking the condition numbers or by employing some kind of robust pole assignment. Furthermore, extensions to descendants such as the QZ or the periodic QR algorithm are presented.

Asymptotic Behavior of Scaled Iterates by Diagonalization Eigenvalue Methods

by Josip Matejaš and Vjeran Hari

Diagonalization methods for solving eigenvalue and singular value problems have been lately reconsidered for their accuracy properties. The usual measures of advancing of the processes and the corresponding stopping criteria have been replaced by ones which warrant that all output data are computed with possibly highest relative accuracy. For several diagonalization methods such as: symmetric and Hermitian Jacobi methods, J-symmetric eigenvalue method and Kogbetliantz method for computing SVD of triangular matrices, new improved measures of convergence have been found. When divided by suitably defined relative gaps in the set of eigenvalues or singular values, they actually measure relative distance between eigenvalues or singular values and the corresponding diagonal elements of (scaled) almost diagonal matrices. They also appear in some way in perturbation estimates for eigenvectors or singular vectors. For all mentioned methods, these measures have form $\|DHD\|_F$ where D is diagonal matrix which makes diagonals of $|DHD|$ ones.

When one-sided versions of diagonalization methods are used, computing of $\|DHD\|_F$ is costly. It requires around $n^2/2$ dot products. Hence it is of interest to find a quadratic reduction rule for the new measure. When eigenvalues/singular values cluster around zero, classical results are useless, since they use absolute gaps which are tiny. Therefore, beside using the new measure of convergence, one would like to obtain quadratic convergence estimates which use relative gaps instead of absolute gaps. Then such results can be used in predicting the number of sweeps till convergence.

In all obtained results, quadratic reduction of the ratio $\|DHD\|_F/\mathrm{rel\_gap}$ is sharply estimated for scaled diagonally dominant matrices. Additionally, the logarithm of this ratio measures number of correct digits in all diagonal elements as approximations of eigenvalues or singular values. The results hold in the general case of multiple eigenvalues or singular values.

Low Rank Perturbation of Eigenvalues of Arbitrary Matrices

by Julio Moro and Froilán M. Dopico

It is well known that if an arbitrary matrix A is slightly perturbed, then each of its multiple eigenvalues splits typically into a certain number of so-called rings of eigenvalues of the perturbed matrix, say, $A(\varepsilon) = A+\varepsilon B$ with B any matrix and ε a small scalar perturbation parameter. If the perturbation matrix B satisfies certain nondegeneracy conditions, which only involve the eigenvectors of A associated with the eigenvalue λ under scrutiny, then each Jordan block of size k associated with λ gives rise to a group of k eigenvalues which are $O(\varepsilon^{1/k})$ perturbations of λ.

The situation becomes much more complicated if the perturbation matrix B fails to satisfy these genericity conditions. This nongeneric case is still far from being well understood. We address the particular case when B has low rank. More precisely, we assume that the rank r of B is smaller than the geometric multiplicity g of λ. In that case, λ is still an eigenvalue of $A(\varepsilon)$ for any ε and the question is how does the Jordan structure of A change under the perturbation. Which Jordan blocks are destroyed and which (if any) remain as Jordan blocks of $A(\varepsilon)$ for $\varepsilon > 0$?

We show that if $r < g$, then generically the r largest Jordan blocks of A disappear and the rest of the blocks remain unaffected. In fact, we show this is true *not only for small* ε *but for perturbations of any size*, provided the rank of the perturbation matrix B does not exceed the geometric multiplicity of λ.

A Jacobi–Davidson Type Method for a Right Definite Two-parameter Eigenvalue Problem

by Bor Plestenjak and Michiel E. Hochstenbach

We are interested in computing one or more eigenpairs of a right definite two-parameter eigenvalue problem

$$\begin{aligned} A_1 x &= \lambda B_1 x + \mu C_1 x, \\ A_2 y &= \lambda B_2 y + \mu C_2 y, \end{aligned} \tag{1}$$

where A_i, B_i, C_i are given real symmetric $n_i \times n_i$ matrices, $x \in \mathbb{R}^{n_1}$, $y \in \mathbb{R}^{n_2}$, and $\lambda, \mu \in \mathbb{R}$. The condition for right definiteness is that

$$(x^T B_1 x)(y^T C_2 y) - (x^T C_1 x)(y^T B_2 y) > 0$$

for $x, y \neq 0$. A pair (λ, μ) that satisfies (1) for nonzero vectors x, y is called an eigenvalue and the tensor product $x \otimes y$ is the corresponding eigenvector. There exist $n_1 n_2$ linearly independent eigenvectors for problem (1).

We present a Jacobi–Davidson type method for computing selected eigenpairs of (1). The method works even without good initial approximations and is able to tackle large problems that are too expensive for existing methods. In each step we first compute Ritz pairs of the projected smaller problem and then expand the search space using approximate solutions of appropriate correction equations. We present two alternatives for the correction equations, introduce a selection technique that makes it possible to compute more than one eigenpair, and give some numerical results.

One Sided Reduction to Bidiagonal Form, Part II

by Rui Ralha and Beresford Parlett

In the previous workshop of the series one of us has proposed an idea for the bidiagonalization of a matrix A using one-sided transformations. A paper on this matter is to appear in a special issue of LAA. However, the method in its simplest form is not backward stable (although there are matrices for which it is more accurate than the standard two-sided algorithm).

The basic idea is the implicit reduction of the symmetric positive semidefinite matrix $A'A$ to tridiagonal form. This can be achieved using Householder, Givens or fast Givens transformations, on the right side of A, and produces a matrix B whose non-adjacent columns can be made orthogonal to working accuracy. The important issue is how to produce an accurate bidiagonal matrix from B. In exact arithmetic the upper bidiagonal matrix R is delivered by a simple QR decomposition with the Gram-Schmidt method. In finite precision, errors will be introduced when a significant reduction in the norm of a transformed column occurs.

In this work we show how to use twisted factorizations to obtain a stable transformation from B to R.

Hyperbolic/Symplectic QR Factorization – A Reflector Approach

by Sanja Singer and Saša Singer

It is widely known that in the most of the matrix reduction or eigenvalue algorithms we can use either rotations or reflectors. We intend to show that the same holds in cases of hyperbolic/symplectic QR factorization algorithms.

We shall construct necessary J–reflectors and block J–reflectors, according to hyperbolic/symplectic "inner–product" matrices

$$J = \mathrm{diag}(\pm 1) \quad \text{and} \quad J = \mathrm{diag}\left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right).$$

It is interesting that both hyperbolic and symplectic block-reflectors can have the same form

$$H(W) = I - 2W(W^*JW)^+W^*J,$$

which is a generalization of the ordinary block-reflectors constructed by Schreiber and Parlett.

The hyperbolic case is much harder to deal with then the symplectic case. This is a consequence of the essential nonuniqueness of the factorization. The easiest way to apply the reflectors is to "borrow" final form from the Givens-like implementation of the algorithm.

In the symplectic case, we can use only block-reflectors, which follows from $w^*Jw = 0$ for all vectors w, but the reduction, at least to the Bunch form can easily be accomplished.

Accuracy of the Symplectic QR Factorization

by Sanja Singer and Saša Singer

To compute the eigenvalues of a nonsingular skew-symmetric matrix A, a one-sided Jacobi-like algorithm constructed by Pietzsch can be used to enhance accuracy. This algorithm begins by a suitable Cholesky-like factorization of A, $A = G^T JG$, where J is block-diagonal with elementary skew-symmetric matrices of order 2 on the diagonal. In fact, this factorization is closely related to the Bunch factorization of A.

In some applications, A is given implicitly in such a form that its natural Cholesky-like factor G is immediately available, but "tall", i.e., not of full row rank. This factor G is unsuitable for the Jacobi-like process. To avoid explicit computation of A, and possible loss of accuracy, the factor has to be preprocessed by a symplectic QR factorization.

First, we present the Givens-like symplectic QR algorithm which uses elementary symplectic matrices analogous to ordinary rotations. Remarkably, this algorithm requires only one nonunitary transformation per step (a pair of columns), which greatly simplifies the analysis. A sequence of simple post-processing transformations also yields the Bunch factor of A, without explicitly forming A.

We discuss numerical properties and implementation details of both algorithms to obtain the rounding error and perturbation bounds. These bounds fit well into the known relative perturbation theory for skew-symmetric matrices given in factorized form, so the accuracy of computed eigenvalues can also be estimated.

Finally, we present several numerical examples and discuss some open problems.

Relative Residual Bounds for Eigenvalues of Indefinite Hermitian Matrices

by Ninoslav Truhar

Let $H \in \mathbf{C}^{n\times n}$ be a Hermitian matrix, $X \in \mathbf{C}^{n\times m}$ be an orthonormal matrix, and $M = X^*HX$, $R = HX - XM$. Furthermore, let $\lambda_1 \geq \ldots \geq \lambda_n$ and $\mu_1 \geq \ldots \geq \mu_m$, be the eigenvalues of H and M, respectively. Then the eigenvalues of M (Ritz values) approximate some eigenvalues of H with absolute error bound by $\|HX - XM\|$. This work contains bounds for $|\lambda_{i_k} - \mu_k|/|\lambda_{i_k}|$ and $|\lambda_{i_k} - \mu_k|/|\mu_k|$ when H is a Hermitian indefinite possibly singular matrix. The bounds are expressed in terms of sines of the canonical angles between relevant subspaces associated with H and M. The presented linear and quadratic relative residual bounds are proper generalization of existing bounds for Hermitian semidefinite matrices. However, the bounds contain the additional condition numbers of certain J-unitary matrices ($J = \mathrm{diag}(\pm 1)$). Therefore, two special cases in which these condition numbers are of moderate magnitude are pointed out: scaled diagonal dominant matrices and quasi definite matrices.

A Priori Error Bounds for Ritz and Harmonic Ritz Vectors

by Jasper van den Eshof

We discuss error bounds for the Rayleigh-Ritz method for the approximation of the extreme eigenpair of a symmetric matrix. The bounds are expressed in terms of the eigenvalues of the matrix and the angle between the given subspace and the eigenvector and therefore the bounds can be truly called *a priori*. We also discuss a sharp bound. These bounds can be of interest in the analysis of iterative eigensolvers since they improve upon the classical error bounds.

Thereafter, we turn our attention to harmonic Rayleigh-Ritz for approximating interior eigenpairs. Practical experiments suggest that there is always a good eigenvector approximation among the harmonic Ritz vectors. We confirm this by deriving an error bound for the asymptotic situation. However, the theoretical justification of this observation in cases where the shift (almost) equals the eigenvalue is still incomplete. We discuss some of the problems and conjecture an error bound for this situation.

This is joint work with Gerard Sleijpen and the material in this talk is largely presented in [2, 1]

References

[1] Gerard L. G. Sleijpen and Jasper van den Eshof, *On the use of harmonic ritz pairs in approximating internal eigenpairs*, To appear in LAA, Special issue IWASEP III, Hagen.

[2] Gerard L. G. Sleijpen, Jasper van den Eshof, and Paul Smit, *Optimal a priori error bounds for the Rayleigh-Ritz method*, To appear in Math. Comp.

Naive Jacobi Algorithms for General Matrices

by Krešimir Veselić

Known Jacobi like algorithms for non-normal matrices have disadvantages of being less effective than the standard algorithms, even more so, if the matrix is (nearly) defective. The so-called naive Jacobi algorithms work well on nearly (block) diagonal matrices and show favourable relative accuracy besides.