Neural, Parallel, and Scientific Computations 19 (2011) 129-164

AN OVERVIEW ON THE EIGENVALUE COMPUTATION FOR MATRICES

MOHAN K. KADALBAJOO[1] AND ANKIT GUPTA[2]

[1]Department of Mathematics and Statistics, Indian Institute of Technology,
Kanpur - 208016
Email Id: kadal@iitk.ac.in
[2]Institute of Technology, Banaras Hindu University, Varanasi – 221005
Email Id: ankit.gupta.apm08@itbhu.ac.in

Abstract. This paper sketches the research developments in the area of computational methods for solving the eigenvalue problems and how the methods developed relate to each other as they evolve over centuries. This is an attempt to write a complete overview on the research on computational aspects of eigenvalue problem, emphasize the history of methods that still play a role and some of those that no longer are considered to be on the main track but are somehow related to the present techniques in some smaller steps. This contribution brings out the state-of-the-art of the algorithms for solving large-scale eigenvalue problems for both symmetric and nonsymmetric matrices separately, thereby clearly drawing a comparison between the differences in the algorithms in practical use for the two. Some of the methods or modifications to the earlier methods that have been notable at the turn of the 21st century will also be covered through this paper under "Modern Approaches". Also as the standard eigenvalue problem has become better understood, in a numerical sense, progress has been made in the generalized eigenvalue problem and this will briefly cover developments in solving these problems too.

1. INTRODUCTION

The terminology eigenvalue comes from the German word *Eigenvert* which means proper or characteristic value. The concept of eigenvalue first appeared in an article on systems of linear differential equations by the French mathematician d' Alembert in the course of studying the motion of string with masses attached to it at various points.

For squared matrix A satisfying $Ax = \lambda x$, is called an eigenvalue problem. The scalar λ is called an "eigenvalue", and the vector x is called the "eigenvector". Computation of the eigenvalues λ explicitly via the characteristic equation i.e.

$$\det(A - \lambda I) = 0$$

is not feasible except for the lower order square matrices($n \leq 4$) since a general polynomial of order $n > 4$ cannot be solved by a finite sequence of arithmetic operations and radicals (Abel–Ruffini theorem[Wikipedia-1]). There do exist efficient root-finding algorithms for higher order polynomials, however, finding the roots of the characteristic polynomial may well be an *ill-conditioned* problem even when the underlying eigenvalue problem is *well-conditioned*. The computation of the roots of the characteristic equation in finite precision may be highly unstable since small perturbations in the coefficients may lead to large perturbations in roots. The numerical computation of the associated eigenvectors from these eigenvalues is even more delicate, particularly when these eigenvectors make small angles with each other. For

Received November 25, 2010

this reason, this method is rarely used. Hence no algorithm can exactly produce all the roots of an arbitrary polynomial of degree n in a finite number of steps and thus eigenvalues for any matrix with a general structure (i.e. not a diagonal or a triangular matrix) must be computed *iteratively*.

Standard eigenvalue problems (of the form $Ax = \lambda x$) continue to be an important area of research in the field of numerical linear algebra and it comes as no surprise that research in this field started as early as 1846, when Jacobi first computed the eigenvalues of symmetric matrices by rotating the matrix to strongly diagonally dominant matrix. Jacobi's method is still of great relevance and has formed the basis of many powerful and popular algorithms such as the QR algorithms. Since computing eigenvalues and vectors is essentially more complicated than solving linear systems, it is not surprising that highly significant developments in this area started with the introduction of electronic computers around 1950, when eigenvalue problems other than the standard ones for instance the generalized eigenvalue problems (of the form $Ax + \lambda Bx + \lambda^2 Cx = 0$) required solution. In the early decades of 20th century, however, important theoretical developments had been made from which computational techniques could grow. Matrix transforming techniques like the ones developed by Jacobi and Householder were found to be inconsistent with larger eigenvalue problems and so proper iterative techniques needed to be devised. Power iteration method wasn't of great significance due to its lack of robustness in terms of its inability to determine eigenvalues other than the dominant ones, although the method is still in use with some modifications, most frequently as a part of very efficient algorithms such as QR-method, Krylov's method etc. Apparently "Google's" *Page-Rank Algorithm* is also based on the Power iteration method [Langville, 2006]. Lanczos and Arnold started their research on improved iteration techniques in the and by the middle of the 19^{th} century [Arnoldi, 1951] developed an eigenvalue algorithm for general (possibly non-Hermitian) matrices and [Lanczos, 1950] came up with an analogous method for Hermitian matrices.

Davidson a chemist came up with an algorithm just around that time that is commonly seen as an extension to Lanczos's method but as [Saad, 1992] points out it's related to the Arnoldi from the view of implementation. The problem of determining reliably full digital information in the subspace spanned by eigenvectors corresponding to coincident or pathologically close eigenvalues has never been satisfactorily solved [Wilkinson, 1965]. Solution of eigenvalue problems and hence the choice of the algorithm most essentially depends on the properties of the matrix, whether the matrix is real or complex, symmetric or nonsymmetric, sparse or dense, hermitian or skew hermitian, unitary etc. Besides the standard eigenproblem, there are a variety of more complicated eigenproblems, for instance $Ax = \lambda Bx$, and more generalized eigenproblems like $Ax + \lambda Bx + \lambda^2 Cx = 0$, higher-order polynomial problems, and nonlinear eigenproblems. All these problems are considerably more complicated than the standard eigenproblem, depending on the operators involved and so problems related to these will be dealt with towards the end sections.

2. SYMMETRIC EIGENVALUE PROBLEM

The symmetric eigenvalue problem with its rich mathematical structure is one of the most aesthetically pleasing problems in numerical linear algebra. Jacobi started his work on real symmetric matrices firstly in 1846 applying plane rotations to the matrix to transform it to a diagonally dominant matrix, elements of whose diagonal were the eigenvalues. As we have already seen that in general the roots of the characteristic polynomial cannot be given in closed form shows that any method must proceed by successive approximations. Although one cannot expect to produce the required eigenvalues exactly in a finite no of steps, there

exist rapidly convergent iterative methods for computing the eigenvalues and eigenvectors numerically for symmetric matrices.

Jacobi's Method

In the classical method of [Jacobi, 1846] plane rotations were used to diagonalize a real symmetric matrix (wherein real plane rotations were used). The number of plane rotations necessary to produce the diagonal form is infinite, which is obvious considering the fact that we cannot, in general, solve a polynomial equation ($n \geq 4$) in a finite number of steps. In practice the process is terminated when the off-diagonal elements are negligible to working accuracy. Jacobi, however, did not use the method to achieve full convergence (which was accomplished later in the 20^{th} century) but combined it with an iterative process (Gauss-Jacobi Iteration) for the missing component of the desired eigenvector (for which he took a unit vector as an initial guess). [Bodewig(pp. 280-287), 1959] pointed out that Jacobi's techniques were kind of a preconditioning form of the Gauss-Jacobi iteration that he also used to solve linear least-squares systems. The preconditioning part of the method, as an iterative technique to diagonalize the matrix was reinvented in 1949 by Goldstine et al and published in a manuscript. After Ostrowski had pointed out that this was actually a rediscovery of Jacobi's method, the adapted manuscript was published in [Goldstein, Murray and Neumann, 1959]. According to [Wilkinson(page 343), 1965], Jacobi's method was already being used independently on desk computers at the National Physical Laboratory in 1947. Since 1949 the method has received extensive coverage in the literature papers by Goldstine, Murray and von Neumann. See papers by [Gregory, 1953], [Henrici, 1958], [Pope and Tompkins, 1957], [Schonhage, 1961] and [Wilkinson, 1962].

Jacobi methods seek to diagonalize a given matrix A by a sequence of Jacobi rotations. Zero elements created at one step will be filled in later and any diagonalizing sequence must be, in principle, infinite. At each step this technique seeks and destroys a maximal off-diagonal element. This rate of convergence sets in after a number of sweeps (i.e. $(n-1)n/2$ elementary rotations), but there is no rigorous bound on the number of sweeps required to achieve the accuracy[Parlett(page 181), 1980]. During the 1960s, considerable effort went into the improvement of this task. In [Corbato, 1963] it is shown that one need only search an array of length n. However, the price for this improvement is twofold: (i) an auxiliary array is needed for the maximal elements in each column, and (ii) the attractive simplicity of the program has been compromised (the number of instructions is more than doubled).

It should be emphasized that by taking advantage of symmetry and working with either the lower or the upper triangle of A, a traditional Jacobi rotation can be made with 4n multiplications and two square roots.

Cyclic Jacobi

The simplest strategy is to annihilate elements, regardless of size, in the order and then begin again for another sweep through the matrix. However, the more vexing problem was the possibility that cyclic Jacobi might not always lead to convergence to a diagonal matrix. [Henrici, 1958] showed that if the angles are suitably restricted the cyclic Jacobi method does indeed converge. In principle the off-diagonal elements could be annihilated in any order at each sweep, provided that no element was missed. Cyclic Jacobi methods waste time annihilating small elements in the early sweeps. Quadratic convergence for the cyclic Jacobi algorithm was proven, under various assumptions, by [Henrici, 1958], [Schonhage, 1961], [Wilkinson, 1962], and [Kempen, 1966]. However, since it does not require off-diagonal search, it is considerably faster than Jacobi's original algorithm.

Improvisations to Jacobi's Method (Givens rotation)

The Jacobi rotation is distinguished by causing the greatest decrease in the sum of squares of the off-diagonal elements. It is locally optimal but not necessarily best for the total computation. A modification to the Jacobi rotation was the Givens rotation which was not introduced until the 1950s. Givens rotation in the (i, j) plane chooses θ to annihilate some element other than a_{ij}. In many ways Givens rotations are more useful than Jacobi rotations. We can, first of all, use Givens original method to reduce A to tridiagonal form T. These rotations more importantly preserve the zeros that have been so carefully created. Next we can do a Jacobi rotation $R_{n-1,n}$. This fills in elements (n, n - 2) and (n - 2, n) and creates a bulge in the tridiagonal form. This bulge can be chased up the matrix and off the top by a sequence of Givens rotations. At this point we have executed one tridiagonal QL transformation with zero shifts. This sequence can be repeated until the off-diagonal elements are negligible. If we relax the habit of zeroing a matrix element at every step, we can do a plane rotation in the (n - 1, n) plane through an angle other than the Jacobi angle. This permits us to incorporate shifts into the QL algorithm and thus accelerate convergence dramatically.

The general superiority of the method of Givens to that of Jacobi for the determination of the eigenvalues gained recognition comparatively slowly in spite of the fact that [Givens, 1954] gave a rigorous a priori error analysis which is one of the landmarks in the history of the subject. Probably the fact that the methods proposed for computing the eigenvectors were found to be unreliable gave rise to unjustified suspicions concerning the accuracy of the eigenvalues. In practice Jacobi programs are only about three times slower than the tridiagonal QL methods, but the trademark of Jacobi's method has been simplicity rather than efficiency. This may be very important in special circumstances (hand-held calculators or computations in space vehicles).

Another Modification to the Jacobi Algorithm

The suggestion, that methods analogous to that of Jacobi should be used for the reduction of a general matrix to triangular form was made by [von Neumann, 1947], [Causey, 1958], [Greenstadt, 1955] and [Lotkin, 1956] in their respective papers. At the matrix conference at Gatlinburg in 1961, Greenstadt gave a summary of the progress which had been made up to that time and concluded that no satisfactory procedure of this kind had yet been developed. [Eberlein, 1962] described a modification of this method based on the observation that for any matrix A there exists a similarity transform $B = P^{-1}AP$ which is arbitrarily close to a normal matrix. In Eberlein's algorithm P is constructed as the product of a sequence of matrices which are generalizations of plane rotations but are no longer unitary. Iteration is continued until B is normal to working accuracy and a feature of the method is that in general the limiting B is the direct sum of a number of 1x1 and 2x2 matrices, so that the eigenvalues are available. Eberlein has also considered the reduction of a general matrix to normal form using a combination of plane rotations and diagonal similarity transformations, and similar ideas have been developed independently by Rutishauser. Generally the guiding strategy at each stage is the reduction of the Henrici departure from normality [Wilkinson(Chapter 3, Sec. 50), 1962]. Development on these lines may yet give rise to methods which are far superior and would seem to be one of the most promising lines of research. Methods of this class are not covered by any of the general error analyses however, one would expect them to be stable since the successive reduced matrices are tending to a normal matrix and the latter has a perfectly conditioned eigenvalue problem.

Advantages and state-of-the algorithm.

A valuable feature of Jacobi is that it never rotates through larger angles than necessary and consequently, when the matrix permits; small eigenvalues can be computed to high relative accuracy. Wilkinson analysed this and showed that the relative error in the eigenvalue approximations is eventually reduced to the order of the condition number of the matrix A times machine precision. This was perfected in 1992, by [Demmel and Veselic, 1992] and [Slapricar, 1992], who showed that for symmetric positive-definite matrices, the condition number of A could be replaced by that of the matrix symmetrically scaled by the diagonal.

The arrival of parallel computers renewed interest in Jacobi methods because n/2 Jacobi rotations can be performed simultaneously if that many processors are available. Hence this technique still is of current interests because in cases where fast storage is tight, it is amenable to parallel computation and because under certain circumstances it has superioir accuracy. Papers by [Modi, 1988], Veselic and [Hari, 1990, 1991] hold much importance to the present state of the art of the Algorithm and its application to parallel computing.

Gerschgorin Theorems

Gerschgorin's theorem has been widely recommended for the location of eigenvalues, [Taussky, 1949] had been a particularly ardent advocate of this theorem and its extensions. The use of diagonal similarity transforms to improve the location of the eigenvalues was discussed in Gerschgorin's original paper [Greschgorin, 1931] but the extension of this device to give the results of classical perturbation theory together with rigorous error bounds appears to be new.

An important solution framework for the symmetric eigenvalue problem involves the production of a sequence of orthogonal transformations $\{Q_k\}$ with the property that the matrices $Q^T_k A Q_k$ are progressively "more diagonal." [Greschgorin theorem, 1931] gives bounds under which all eigenvalues of a matrix lie. The theorem essentially states that if the off-diagonal entries of a square matrix over the complex numbers have small norms, the eigenvalues of the matrix cannot be "far from" the diagonal entries of the matrix. Therefore, by reducing the norms of off-diagonal entries one can attempt to approximate the eigenvalues of the matrix. Of course, diagonal entries may change in the process of minimizing off-diagonal entries.[Wilkinson(Chapter 4), 1962]. Interestingly Greschgorin theorem can be very useful for computation of some eigenvalues without involving softwares or numerical iterative techniques[T. D. Roopamala and S. K. Katti, 2010].

Schur's Factorisation

The Schur decomposition of a square nxn matrix A with complex entries is $A = QUQ^{-1}$, where Q is a unitary matrix (so that its inverse Q^{-1} is also the conjugate transpose Q* of Q), and U is an upper triangular matrix, which is called a Schur form of A. Since U is similar to A, it has the same multiset of eigenvalues, and since it is triangular, those eigenvalues are the diagonal entries of U.

If the triangular matrix U is written as U = D + N, where D is diagonal and N is strictly upper triangular (and thus a nilpotent matrix). The diagonal matrix D contains the eigenvalues of A in arbitrary order (hence its Frobenius norm, squared, is the sum of the squared moduli of the eigenvalues of A, while the Frobenius norm of A, squared, is the sum of the squared singular values of A). The nilpotent part N is generally not unique either, but its Frobenius norm is uniquely determined by A (just because the Frobenius norm of A is equal to the Frobenius norm of U = D + N) [Saad, 1989].

It is clear that if A is a normal matrix, then U from its Schur decomposition must be a diagonal matrix and the column vectors of Q are the eigenvectors of A. Therefore, the Schur

decomposition extends the spectral decomposition. In particular, if A is positive definite, the Schur decomposition of A, its spectral decomposition and its singular value decomposition coincide.

Given square matrices A and B, the generalized Schur decomposition factorizes both matrices as A = QSZ * and B = QTZ *, where Q and Z are unitary, and S and T are upper triangular. The generalized Schur decomposition is also sometimes called the QZ decomposition [Golub and van Loan, 1996].

Householder's Algorithm

In linear algebra, a Householder transformation (also known as Householder reflection or elementary reflector) is a linear transformation that describes a reflection about a plane or hyper plane containing the origin. Householder transformations are widely used in numerical linear algebra, to perform QR decompositions and in the first step of the QR algorithm. The Householder transformation was introduced in 1958 by Alston Scott Householder. The algorithm Householder's tri-diagonalization works for arbitrary symmetric matrices, possibly with multiple eigenvalues. In the latter case, the resulting tridiagonal matrix would decompose if the computation were exact. The algorithms backtransformation derives the eigenvector system of A from an eigenvector system of the tri-diagonal matrix. Householder tri-diagonalization, is however just the first step in the determination of the eigenvalues and eigenvectors of the matrix A, it is followed by a determination of the eigenvalues by the procedure of tri-di-bisection. The determination of corresponding eigenvectors is made possible by applying the procedure of tridiinverse iteration to the resulting tri-diagonal matrix from the Householder Transformation. In the last step the procedure backtransformation is to be applied.

Householder tri-diagonalization may also form the preparation for the application of the QD or QR algorithm.

Householder first suggested the use of the elementary Hermitian matrices for the reduction to tri-diagonal form in a lecture given at Urbana in 1958 and referred again to it briefly in a joint paper with [Bauer, 1959]. Its general superiority to Givens' method both in speed and accuracy was first recognized by [Wilkinson, 1960a]. According to Wilkinson, there are (n - 2) steps in this reduction, in the r^{th} of which the zeros are introduced in the r^{th} row and r^{th} column without destroying the zeros introduced in the previous steps.

Error analyses were given for the fixed-point reduction by [Wilkinson, 1960b] and for the floating-point reduction by [Ortega, 1963]. The calculation of the eigenvalues of a tri-diagonal matrix using the Sturm sequence property was described and an error analysis given by Givens (loc. cit.). This analysis applied to fixed-point computation with ad hoc scaling was the first in which explicit reference was made to what is now known as a 'backward error analysis'(though this idea was implicit in the papers by [von Neumann and Goldstine, 1947] and [Turing, 1948]). The problem of computing eigenvectors of a tri-diagonal matrix when accurate eigenvalues are known has been considered by [Brooker and Sumner, 1966], [Forsythe, 1958], [Wilkinson, 1958], and by Lanczos in the paper by [Rosser et al., 1951].

Power Iteration Method

For the careful analytic treatment of asymptotic properties of powers of matrices refer to the papers by Ostrowski and by Gautschi. [Muntz, 1913a, 1913b] gives an early, possibly the earliest, treatment of simple iteration as a practical computational device. The idea of iterating on a matrix to obtain more than one root at a time is perhaps fairly natural. It was suggested by [Horst, 1937], but the paper was read only by a limited group. [Perron, 1908] provides the

basis for a rigorous justification. More modern treatments are listed under the names of Rutishauser, Bauer, and Francis.

The Power method, for general square matrices, is the simplest of all the methods for solving eigenvalues and eigenvectors. The basic idea is to multiply the matrix A repeatedly by a well-chosen starting vector, so that the component of that vector in the direction of the eigenvector with largest eigenvalue in absolute value is magnified relative to the other components. Householder called this Simple Iteration, and attributed the first treatment of it to [Muntz, 1913a, 1913b]. [Bodewig(p. 250), 1959] attributes the power method to [von Mises, 1929], and acknowledges Muntz for computing approximate eigenvalues from quotients of minors of the explicitly computed matrix A_k , for increasing values of k. The speed of convergence of the Power iteration depends on the ratio of the second largest eigenvalue (in absolute value) to the largest eigenvalue (in absolute value). In many applications this ratio can be close to 1. This has motivated research to improve the effciency of the Power method. If the eigenvalue is multiple, but semi-simple, then the algorithm provides only one eigenvalue and a corresponding eigenvector. A more serious difficulty is that the algorithm will not converge if the dominant eigenvalue is complex and the original matrix as well as the initial vector is real. This is because for real matrices the complex eigenvalues come in complex pairs and as result there will be (at least) two distinct eigenvalues that will have the largest modulus in the spectrum. Then the theorem will not guarantee convergence. There are remedies to all these difficulties like the shift-and-invert techniques.

If the power method is used not on the original matrix but on the shifted matrix A + I maybe on a matrix of the form $B(\sigma) = A + \sigma I$ for any positive σ(and the choice $\sigma = 1$ is a rather arbitrary choice), there are better choices of the shift σ, but more generally, when the eigenvalues are real it is not too difficult to find the optimal value of σ, i.e., the shift that maximizes the asymptotic convergence rate. The scalars σ are called shifts of origin. The important property that is used is that shifting does not alter the eigenvectors and that it does change the eigenvalues in a simple known way, it shifts them by σ. From the above observations, one can think of changing the shift σ occasionally into a value that is known to be a better approximation of λ_1 than the previous σ. For example, one can replace occasionally σ by the estimated eigenvalue of A that is derived from the information that α_k converges to $1/(\lambda_1-\sigma)$, i.e., we can take $\sigma_{new} = \sigma_{old} + 1/\alpha_k$. Strategies of this sort are often referred to as shift-and-invert techniques. Another possibility, which may be very efficient in the Hermitian case, is to take the new shift to be the Rayleigh quotient of the latest approximate eigenvector v_k. However since the LU factorization is expensive so it is desirable to keep such shift changes to a minimum.

At one extreme where the shift is never changed, we obtain the simple Inverse Power method. At the other extreme, one can also change the shift at every step. The algorithm corresponding to this case is called Rayleigh Quotient Iteration (RQI) and has been extensively studied for Hermitian matrices.

Inverse Iteration

It is interesting that the most effective variant of the Power method is the Inverse Power method, in which one works with the matrix $(A - \lambda I)^{-1}$. This variant was proposed as late as 1944 by Wielandt (Wielandt's fractional iteration). Wielandt also proposed continuing the process after the largest eigenvalue has converged, by working with the defected matrix $A-\lambda vv^*$, for which λ, v is the computed eigenpair (with $\|v\|_2 = 1$), associated with the largest eigenvalue in magnitude. (The deflation procedure outlined here is for symmetric matrices. For nonsymmetric matrices it is necessary to work with at least two vectors; the choice of one of the vectors may not be unique.) This is called implicit deflation; another possibility is to

keep the iteration vectors orthogonal to the computed eigenvector(s): explicit deflation. A compact description and analysis of these deflation techniques was given by [Parlett, 1980].

There are two ways in which we can implement the inverse iteration process. One obvious possibility would be to use the original matrix $A \epsilon R^{nxn}$. An alternative is to replace A in this equation by the tridiagonal matrix $T \epsilon R^{nxn}$ supplied by [Householder's method, 1964]. The calculation is then very much quicker, but produces the eigenvector of T. To obtain the corresponding eigenvector of A we must then apply to this vector the sequence of Householder transformations which were used in the original reduction to tridiagonal form.

In general, the nearer σ is to the eigenvalue λ_k the faster the power iteration $(A - \sigma I)^{-1}$ must converge to the eigenvector x_k. However, as σ tends to λ_k the matrix $(A - \sigma I)$ becomes more and more ill-conditioned until when $\sigma = \lambda_k$ it is singular.

The Power method and the Inverse Power method, in their pure form are no longer competitive methods even for the computation of a few eigenpairs, but they are still of interest since they are explicitly or implicitly part of most modern methods such as the QR method, and the methods of Lanczos and Arnoldi.

Rayleigh Quotient Iteration

Another possible and very useful variant to the Power Method is working with properly updated shifts σ in the inverse process and in particular, if one takes the Rayleigh quotient with the most recent vector as a shift, then one obtains the Rayleigh quotient iteration. According to [Parlett(p. 71), 1980], Lord Rayleigh used in the 1870s a less powerful technique, he did a single shift-and-invert step with a Rayleigh quotient for an eigenvector approximation, but with a unit vector as the right-hand side. (This saves the refactoring of the matrix $(A - \sigma I)$, at each iteration.) The modern RQI, in which one takes the most current eigenvector approximation as the right-hand side, leads to very fast convergence. Ostrowski, in a series of six papers [Ostrowski, 1958-1959], studied the convergence properties for variance of RQI for the symmetric and nonsymmetric case. He was able to establish cubic convergence in both cases under various circumstances (in the nonsymmetric case for a properly generalized Rayleigh quotient). The use of iterative methods in which a number of vectors are used simultaneously has been described in papers by [Bauer, 1957 and 1958]. The method of [Wilkinson(Sec. 38), 1962] has been discussed more recently by [Voyevodin, 1962]. A detailed discussion on this topic is given in[Householder's book(Chapter 7), 1964].

Treppen-iteration, the orthogonalization technique and bi-iteration may be performed using the inverse matrix. We have typically in the case of treppen-iteration

$$AX_{s+1} = L_s,\ X_{s+1} = L_{s+1}R_{s+1}$$

and it is obvious that A^{-1} need not be determined explicitly. We need only carry out a triangular decomposition of A, and this is particularly important if A is of Hessenberg form. We may incorporate shifts of origin, but a triangular decomposition of $(A - \sigma I)$ is required for each different value of σ. Again we have the weakness referred to in [Wilkinson(Sec. 34), 1962] that σ cannot be chosen freely after a value of λ_1 has been determined.

QR-Algorithm

The QL, or QR, algorithm is the currently the most preferred way to compute the eigenvalues of matrices and can be seen invariably as the more sophisticated version of the Power Algorithm. The basic idea being to perform a QR decomposition, writing the matrix as a product of an orthogonal matrix and an upper triangular matrix, multiply the factors in the other order, and iterate.

According to Parlett, the key idea came from Rutishauser with his construction of a related algorithm called LR way back in 1958. Wilkinson however asserts Rutishauser's first account

of the LR algorithm to be given in 1955, with shifts of origin and interchanges applied to Hessenberg matrices in 1959, mainly on matrices with complex elements. However much of the credit for the development of QR algorithm is due to [Francis(England), 1961, 1962] who made the three observations required to make the method successful to wit (1) the basic QR transformation, (2) the invariance of the Hessenberg form, and (3) the use of origin shifts to hasten convergence. According to Parlett, Francis received some assistance from Strachey and Wilkinson. [Kublanovskaya, 1963] working independently during the same period discovered (1) and (3), but without the improvements induced by (2) the algorithm is very slow. The method can be seen as a natural analogue of the LR algorithm, however Francis' papers not only describe just the mere detail of the QR algorithm, but also the technique for combining complex conjugate shifts of origin and methods for dealing with consecutive small sub-diagonal elements [Wilkinson(Sec. 38), 1962] which contribute heavily to the effectiveness of the QR algorithm. For the proofs of convergence of the QR algorithm we refer to [Wilkinson(Sec. 29-32), 1962]. However, proofs based on the more sophisticated determinantal theory have been given by [Kublanovskaya, 1963] and [Householder, 1964].

The intimate connection between the RQI and QR was noticed by both Kahan and Wilkinson. In fact, the observed monotonic decay of the last off- diagonal element in the tridiagonal QR algorithm led Kahan to the global convergence analysis of RQI [Parlett(Chapter 4), 1980]. In 1968 Wilkinson proved that the tridiagonal QL algorithm can never fail when his shift is used. His proof is based on the monotonic decline of the product $|\beta_1\beta_2|$. However, the analysis is simpler when $\beta_1^2\,\beta_2$ is used in place of $\beta_1\beta_2$ [Parlett(Sec. 8.9, 8.10), 1980]. The rapid and global convergence of QL with Wilkinson's shift is very attractive. Apparently in 1985, [Erxiong and Zhenye, 1985] came up with a better strategy.

The PWK version of the root-free QL algorithm [Parlett(Sec. 8.15), 1980] has not appeared in the open literature. Backward stability of PWK was proved by [Feng, 1991]. The algorithm was incorporated in the LAPACK library under the name SSTERF using QR rather than QL format. Gates and Gragg in 1996 demonstrated how to get rid of one multiplication in the inner loop while preserving stability in the algorithm. In PWK (QL shifted version) the operations

$$\gamma_i \leftarrow c_i^2(\alpha_i-\sigma) - s_i^2\gamma_{i+1},$$

$$\tilde{\alpha}_i \leftarrow \gamma_{i+1} + \alpha_i - \gamma_{i.}$$

were replaced by Gates and Gragg by introducing a new variable u_i and rewrite these operations (in QL format) as

$$u_i \leftarrow \alpha_i + \gamma_{i+1},$$

$$\gamma_i \leftarrow c_i^2(u_i-\sigma) - \gamma_{i+1},$$

$$\tilde{\alpha}_i \leftarrow u_i - \gamma_{i.}$$

The variable u_i is a "temporary" variable and the quantity $\alpha_i + \gamma_{i+1}$ was never sent back to the memory. In addition Gates and Gragg show how the sign of c_k may be recovered from c_k^2 (assuming $s_k > 0$). Observe that $\text{sign}(c_k) = \text{sign}(\pi_{k+1})$, by definition of c_k, and, in addition $\gamma_k = \pi_k c_k$ so that $\text{sign}(\gamma_k) = \text{sign}(\pi_k)\text{sign}(c_k)$. Hence, sequentially,

$$\text{sign}(\pi_k) = \begin{cases} \text{sign}(\gamma_k)\text{sign}(\pi_k), & \pi^2_{k+1} \neq 0, \\ -\text{sign}(\pi_{k+2}) & \text{otherwise.} \end{cases}$$

The QR method for computing all of the eigenvalues and associated eigenvectors of a dense symmetric matrix is really powerful and robust and has essential enhancements that make the method extend its application to the solution of the nonsymmetric eigenvalue problem with the general GR algorithm.

Lanczos Method

The [Lanczos algorithm, 1950] is an iterative algorithm invented by Cornelius Lanczos in 1950 that is an adaptation of power methods to find eigenvalues and eigenvectors of a square matrix or the singular value decomposition of a rectangular matrix. Lanczos algorithm emerged from the pioneering Ph.D. thesis by [Paige, 1971]. Some important facets of that work were published in [Paige, 1972, 1976]. Algorithm holds special interests for finding decompositions of very large sparse matrices like for instance in Latent Semantic Indexing where matrices relating millions of documents to hundreds of thousands of terms must be reduced to singular-value form. Variations on the Lanczos algorithm exist where the vectors involved are tall, narrow matrices instead of vectors and the normalizing constants are small square matrices. The idea of using the algorithm iteratively, i.e., restarting periodically, goes back to some of the early attempts to use Lanczos[Parlett, 1980] on the computers available in the 1950s. These are called "block" Lanczos algorithms, first developed by [Golub and Underwood, 1977]. These algorithms can be much faster on computers with large numbers of registers and long memory-fetch times. The error bounds on the Hermitian Lanczos algorithm are described [Saad, 2000]. Bounds of a different type have been proposed by [Kaniel, 1966] (however there were a few errors for the case i > l in Kaniel's original paper and some of these errors were later corrected by [Paige, 1971]). [Saad, 1980] also developed similar bounds for Block Lanczos algorithm.

Selective orthogonalization was introduced by [Parlett and Scott, 1979]. Scott also advocated that SO be applied only at the wanted end of the spectrum. [Cullum and Willoughby, 1985] showed how modifications in the Lanczos algorithm can be used to compute the whole spectrum of large matrices. We mention two recent developments related to the Lanczos algorithm. Ruhe takes the idea of shift-and-invert Lanczos, i.e., use $(A - \sigma I)^{-1}$ as the operator, and extends it to present us with rational Krylov subspace methods [Ruhe, 1984].

The second idea was developed for the nonsymmetric eigenvalue problem but extends readily to the symmetric Lanczos algorithm. Saad developed the idea of explicitly restarting a Lanczos run when storage is limited. However, it seems preferable to restart the Lanczos algorithm implicitly by explicitly post multiplying the matrix Q_j of Lanczos vectors by cleverly chosen orthogonal matrices.

Many implementations of the Lanczos algorithm restart after a certain number of iterations. One of the most influential restarted variations is the implicitly restarted Lanczos method [Calvetti, Reichel, and Sorensen, 1994], which is implemented in ARPACK. This has led into a number of other restarted variations such as restarted Lanczos bi-diagonalization [Kokiopoulou, Bekas and Gallopoulos, 2004]. Another successful restarted variation is the Thick-Restart Lanczos method [Wu and Simon, 2000], which has been implemented in a software package called TRLan. Implementations on massively parallel machines have recently been discussed by [Petiton, 1991] on the CM-2 and by [Scott, 1989] on the iPSC/2. Cullum and Willoughby offer a FORTRAN code for the Hermitian case in their book [Cullum and Willoughby, 1985] based on the Lanczos algorithm without any form of re-orthogonalization. A similar (research) code was also developed by [Parlett and Reid, 1981]. Recently, [Freund, Gutknecht, and Nachtigal, 1990] published a FORTRAN implementation of their Look-Ahead Lanczos algorithm. We know of no other codes based on the Lanczos algorithm with or without re-orthogonalization. There have been a few implementations of the

Hermitian Lanczos and the Block Lanczos algorithm with some form of re-orthogonalization. Parlett came up with a survey concerning software availability in 1984 [Parlett, 1984].

The Lanczos algorithm has had a diversified history since it was first introduced in 1950. Although Lanczos pointed out that his method could be used to find a few eigenvectors of a symmetric matrix it was heralded at that time as a way to reduce the whole matrix to tridiagonal form. In this capacity it flopped unless very expensive modifications were incorporated. Twenty years later, Paige showed that despite its sensitivity to roundoff, the simple Lanczos algorithm is nevertheless an effective tool for computing some outer eigenvalues and their eigenvectors [Parlett, 1980]. Lanczos procedure has been modified with success for large, difficult problems by incorporating a variety of checks on the progress of the computation by [Whitehead, 1972], [Davidson, 1975] and [van Kats and van der Vorst, 1976].

3. NON-SYMMETRIC EIGENVALUE PROBLEM

For nonsymmetric matrices the picture is less rosy as compared to the symmetric problem. Unfortunately, it is not always possible to diagonalize a nonsymmetric matrix, and even if it is known that all eigenvalues are distinct, then it may be numerically undesirable to do this. The most stable methods seem to be based on the Schur factorization, that is for each nxn matrix A, there exists an orthogonal Q, so that $Q^TAQ = R$, where R is upper triangular. Apart from permutations and signs, the matrix Q is unique. The Schur factorization reveals much of the eigenstructure of A, its diagonal elements are the eigenvalues of A and the off-diagonal elements of R indicate how small the angles between eigenvectors may be. Other methods like Power method, Inverse and Rayleigh quotient iteration can also be applied to nonsymmetric matrices but several research papers have shown that there exist a set of matrices for which the algorithms fail.

For matrices not too large, QR is the method of choice, but for larger matrices the picture is less clear. Modern variants of the Arnoldi method seem to be the first choice at the moment, and, if approximations are available, the Jacobi-Davidson method may be attractive.

Jacobi Davidson Method

Jacobi and Davidson originally presented their methods for symmetric matrices, but as is well known and as we will see in the present discussion, both methods can easily be formulated for nonsymmetric matrices. The Jacobi-Davidson (JD) subspace iteration method can be applied to nonsymmetric and complex matrices A and can also be extended to solve generalized eigenvalue problems [Fokkema, Sleijpen, and Vorst, 1996]. In this approach a search subspace is generated onto which the given eigenvalue problem is projected. For the construction of effective subspaces we observe that for a given approximate Ritz pair (θ, s), the residual is given by

$$r = As - \theta s$$

Following the historical technique of [Jacobi, 1846](for strongly diagonally dominant matrices), it was suggested in [Sleijpen and Vorst, 1996] to compute a correction Δs for s in the subspace orthogonal to s, such that the residual vanishes in that subspace.

That is, we want to solve

$$(I - ss^*)(A- \theta I) (I - ss^*) \Delta s = -r,$$

For $\Delta s \perp s$

It can be shown that for $\theta = \lambda$ (an eigenvalue of A), this correction Δs leads immediately to the corresponding eigenvector $y = s + \Delta s$; $Ay= \lambda y$.

For the expansion of the subspace we solve above equation for a given Ritz value θ, and we expand the subspace with Δs. We compute a new Ritz pair with respect to the expanded subspace and we repeat the above procedure. This is the basis for the Jacobi-Davidson Method. The sketched procedure leads to quadratic convergence of the Ritz value to an eigenvalue if A is nonsymmetric, and to cubic convergence if A is symmetric.

A major advantage of Jacobi-Davidson is that the correction equations allow approximate solutions, hence they can be solved by existing iterative methods; preconditioners may be incorporated. In contrast, shift-invert Arnoldi-type methods generally require rather accurate system solves, which presents difficulty for preconditioned iterative linear solvers. One main reason for the different accuracy requirement is that, in an Arnoldi-type method, the residual vector at each iteration is computed as a by-product of the Arnoldi decomposition, it needs high accuracy in order to keep the Arnoldi structure; while a JD method preforms the Rayleigh-Ritz procedure and computes a new residual vector ($r = Ax - \mu x$) at each iteration step. The extra work frees a Davidson-type method from the need to solve its correction equations accurately.

Schur and Jordan decomposition

Most of the algorithms involve transforming the matrix A into simpler or canonical forms, from which it is easy to compute its eigenvalues and eigenvectors. These transformations are called similarity transformation. The two most common canonical forms are called the Jordan form and Schur form. The Jordan form is useful theoretically but it is very hard to compute in a numerically stable fashion, which is why Schur decomposition is the preferred way to transform a matrix to triangular (or diagonal) matrix.

Schur's Decomposition

The proof of the existence of Schur decompositions is essentially the one given by [Schur, 1909], an early example of the use of partitioned matrices. The decomposition can be computed stably by the QR algorithm [Stewart(Sec. 2-3, Chapter 2), 2001] and it can often replace more sophisticated decompositions-the Jordan canonical form in matrix algorithms.

A block upper triangular matrix with either 1-by-l or 2-by-2 diagonal blocks is upper quasi-triangular. The real Schur decomposition amounts to a real reduction to upper quasi-triangular form.

As we know that a real matrix can have complex eigenvalues since the roots of its characteristic polynomial may be real or complex. Therefore, there is not always a real triangular matrix with the same eigenvalues as a real general matrix, since a real triangular matrix can only have real eigenvalues. Therefore, we must either use complex numbers or look beyond real triangular matrices for our canonical forms for real matrices. It will turn out to be sufficient to consider block-triangular matrices. The canonical forms that are computed presently are mostly block triangular and proceed computationally by breaking up large diagonal blocks into smaller ones. For a complex matrix A, the ultimate canonical form is triangular and for a real matrix A, the ultimate canonical form is quasi-triangular.

Block Diagonalization

Although rounding error prevents matrices that are exactly defective from arising in practice, one frequently encounters matrices with very ill-conditioned systems of eigenvectors. Block diagonalization would seem to be an attractive way out of this problem. The idea is to generate clusters of eigenvalues associated with nearly dependent eigenvectors, find bases for their eigenspaces, and reduce the matrix to block diagonal form. If the blocks are small, one can afford to lavish computation on their subsequent analysis.

Unfortunately, this sensible-sounding procedure is difficult to implement. Leaving aside the possibility that the smallest cluster may consist of all the eigenvalues, the problem is to find the clusters. Also it has been shown that eigenvalues that should be clustered do not have to lie very close to one another. Consequently, schemes for automatically block diagonalizing a matrix by a well-conditioned similarity transformation have many ad hoc features. For detailed description see [Bavely and Stewart, 1979], [Demmel and Kagstorm, 1987], [Demmel, 1983], [Demmel, 1986], [Gu., 1994].

Jordan Decomposition

The Jordan canonical form was established in 1870 by Camile Jordan [Jordan, 1870]. The characteristic polynomials of the Jordan blocks of a matrix A are called the elementary divisors of A. Thus to say that A has linear elementary divisors is to say that A is non-defective or that A is diagonalizable. The issues regarding the computing of block diagonal forms of a matrix apply to the Jordan form as well. For details refer [Golub and Wilkinson, 1976], [Kagstorm and Ruhe, 1980], [Ruhe, 1970]

The Jordan form tells everything that we might want to know about a matrix and its eigenvalues, eigenvectors, and invariant subspaces. There are also explicit formulas based on the Jordan form to comput e^A or any other function of matrix. But it is bad to compute the Jordan form for two numerical reasons:

1. It is a discontinuous function of A, so any rounding error can change it completely.
2. It cannot be computed stably in general.

Note that the Schur form is not unique, because the eigenvalues may appear on the diagonal of T in any order. This introduces complex numbers even when A is real, therefore a canonical form is preferred that uses only real numbers, because it will be cheaper to compute. This also means that we will have to sacrifice a triangular canonical form and settle for a block-triangular canonical form.

Power Method, Inverse and the RQI

As an explicit method for computing a dominant eigenpair, the power method goes back at least to [Miintz, 1913]. Hotelling in his famous paper on variance components [Hotelling, 1933] used it to compute his numerical results. [Aitken, 1937] also gave an extensive analysis of the method.

Power method is used most often and sometimes unknowingly, for example, a simulation of a discrete-time discrete-state Markov chain amounts to applying the power method to the matrix of transition probabilities. For a time, before the advent of the QR algorithm, the power method was one of the standard methods for finding eigenvalues and eigenvectors on a digital computer. Wilkinson, working at the National Physical Laboratory in England, honed this tool to as fine an edge as its limitations allow. A particularly important problem was to deflate eigenvalues as they are found so that they do not get in the way of finding smaller eigenvalues.

The Inverse power method is due to [Wielandt, 1944]. [Ipsen, 1997] gives a survey of the method and its properties, along with an extensive list of references. The method is chiefly used to compute eigenvectors of matrices whose eigenvalues can be approximated by other means [Stewart(Algorithm 2.6, Chapter 3), 2001].

The fact that A - kI is ill conditioned when k is near an eigenvalue of A makes certain that the vector in the inverse power method will be inaccurately calculated. The analysis given here which shows that the inaccuracies do no harm is due to [Wilkinson, page 619-622, 1965].

The Rayleigh quotient is due to Lord Rayleigh (J. W. Strutt), who defined it for symmetric matrices. Some of its many generalizations have appeared frequently in the form of many research papers. The fact that the Rayleigh quotient gives an optimal residual is due to [Wilkinson, p. 172, 1965]. Wilkinson used shifting to good effect at the National Physical Laboratory. The shift was actually entered by machine operators in binary by switches at the console before each iteration [Wilkinson, p. 577, 1965]. It is important to use the Rayleigh quotient to compute the optimal residual for the obvious reason that if we use the constant shift k; the residual can never converge to zero. A more subtle reason however, is the behaviour of the residual when the matrix is strongly non-normal—that is, when the off-diagonal part of its Schur form dominates the diagonal part [Henrici, 1962]. In that case, the first residual will generally be extremely small and then will increase in size before decreasing again. The use of the Rayleigh quotient to compute the optimal residual mitigates this problem. For details see [Golub and Wilkinson, 1976] and [Ipsen, 1997]. The Rayleigh quotient iteration has been analysed extensively in a series of papers by [Ostrowski, 1958a, 1958b, 1959a, 1959b, and 1959c].

As we have already seen that if the matrix is symmetric, then Rayleigh quotient iteration succeeds [Parlett, 1980], [Batterson and Smillie]. Parlett has also questioned the algorithms (or some of its modification) success when applied to the more general case of nonsymmetric matrices. But unfortunately there exists an open set of nonsymmetric real matrices for which the algorithm fails. Since the characteristic polynomials can have both real and complex roots, invariant subspaces to which the algorithm converges are the reasonable targets for an eigenvector algorithm.

1R. an eigenvector,

2R. the one-dimensional span of an eigenvector,

3R. an eigenspace, or

4R. a generalized eigenspace.

For a symmetric matrix, Rayleigh quotient iteration succeeds with the solution having the form 2R. If the matrix is nonsymmetric and the characteristic polynomial has complex roots, then it is perhaps unreasonable to expect an eigenvalue algorithm to converge to one of the above types of solution. For each of the four types of solutions there is a complex analogue (1C-4C, where the space referred to in 2C is two-dimensional). An algorithm is said to weakly succeed for a matrix A provided that the set of initial points for which the sequence converges to either 4R or 4C is a set of full measure. An algorithm which does not weakly succeed is said to strongly fail which is the case for RQI for nonsymmetric matrices [Batterson and Smillie, 1990].

Givens and Householder's transformation

Rotations used by [Jacobi, 1846] in his celebrated algorithm for the symmetric eigenvalue problem, are usually distinguished from plane rotations because Jacobi chose his angle to diagonalize a 2x2 symmetric matrix. [Givens, 1954] was the first to use them to introduce a zero at a critical point in a matrix; hence they are often called Givens rotations. Givens gave the method for the inversion of a nonsymmetric matrix, which has been in use at Oak Ridge National Laboratory and has proved to be highly stable numerically but requires a rather large number of arithmetic operations, including a total of n(n - 1)/2 square roots. Strictly, the method achieves the triangularization of the matrix, after which any standard method may be employed for inverting the triangle. The triangular form is brought about by means of a sequence of n(n - 1)/2 plane rotations, whose product is an orthogonal matrix. Each rotation requires the extraction of a square root. The advantage in using the method lies in the fact that an orthogonal matrix is perfectly conditioned. Hence the condition of the matrix cannot

deteriorate through successive transformations. In fact, if one defines the condition number of a matrix A to be [Householder, 1958a]

$$\gamma(A) = \|A\| \, \|A^{-1}\|,$$

where the norm is the spectral norm, then for any orthogonal matrix W, $\gamma(W) = 1$ and the condition of any matrix is preserved under multiplication by an orthogonal matrix:

$$\gamma(WA) = \gamma(A).$$

To look at the matter another way, if $WA = R$, where R is an upper triangle, then

$$A^TA = A^TW^TWA = R^TR,$$

so that R is precisely the triangle one would obtain from the application of the Choleski square-root method to the positive definite matrix A^TA. It is, in fact, the matrix to which [von Neumann and Goldstine, 1947] are led in their study of Gaussian elimination as applied to a positive definite matrix. To obtain the precise triangle that would result from Gaussian elimination with A^TA, one has only to remove as a factor the diagonal of R:

$$R = DU,$$

where U has a unit diagonal.[Householder, 1958b]

For error analyses of plane rotations see [Wilkinson, pp. 131-143], [Gentleman, 1975], and especially [Higham Sec 18.5, 1996]. The superior performance of plane rotations on graded matrices is part of the folklore. In special cases, however, it may be possible to show something. For example, [Demmel and Veselic, 1992] have shown that Jacobi's method applied to a positive definite matrices is superior to Householder tridiagonalization followed by the QR algorithm. Mention should also be made of the analysis of [Anda and Park, 1996].

Householder transformations seem first to have appeared in a text by [Turnbull and Aitken, 1932], where they were used to establish Schur's result [Schur, 1909] that any square matrix can be triangularized by a unitary similarity transformation. They also appear as a special case of a class of transformations in [Feller and Forsythe, 1951]. [Householder, 1958b], who discovered the transformations independently, was the first to realize their computational significance. Householder called his transformations elementary Hermitian matrices in his Theory of Matrices in Numerical Analysis [Householder, 1964], a usage which has gone out fashion. Since the Householder transformation $I - uu^T$ reflects the vector u through its orthogonal complement (which remains invariant), these transformations have also been called elementary reflectors. Householder seems to have missed the fact that there are two transformations that will reduce a vector to a multiple of e_1 and that the natural construction of one of them is unstable. This oversight was corrected by [Wilkinson, 1960]. [Parlett, 1971, 1998]] has shown how to generate an alternative transformation in a stable manner. Although Householder derived his triangularization algorithm for a square matrix, he pointed out that it could be applied to rectangular matrices as well.

QR-Algorithm

The situation for real, nonsymmetric matrices is much more complicated than for the symmetric matrices . In this case, the given matrix has real elements, but its eigenvalues may well be complex. Real matrices are used throughout, with a double shift strategy that can handle two real eigenvalues, or a complex conjugate pair. Even thirty years ago, counter examples to the basic iteration were known and Wilkinson introduced an “ad-hoc” shift to handle them. But no one has been able to prove a complete convergence theorem, infact counter examples exist that will cause the real, nonsymmetric QR algorithm to fail, even with Wilkinson’s ad hoc shift, but only on certain computers.

The real Schur decomposition amounts to a real reduction to upper quasi-triangular form. The most prominent shortcoming associated with Schur’s decomposition is that each step requires a full QR factorization costing $O(n^3)$ flops. Fortunately, the amount of work per iteration can

be reduced by an order of magnitude if the orthogonal matrix U_0 is judiciously chosen. In particular, if $U^T_0 AU_0 = H_0 = (h_{ij})$ is upper Hessenberg ($h_{ij} = 0$, $i > j + 1$), then each subsequent H_k requires only $O(n^2)$ flops to calculate. To see this we look at the computations $H = QR$ and $H_+ = RQ$ when H is upper Hessenberg. As we can always upper triangularize H with a sequence of n - 1 Givens rotations:

$Q^TH = G^T_{n-1}.....G^T_1H = R$. Here, $G_i = G(i, i + 1,\theta_i)$. this being known as the Hessenberg QR step. Hence reducing A to Hessenberg form using 'Hessenberg QR step' and then iterating with QR algorithm using the deflation technique to produce the real Schur form is the standard means by which the dense nonsymmetric eigenproblem is solved.

Although the reduction to Hessenberg form which enables us to carry out each step of the QR algorithm with order n^2 rather than n^3, multiplications and additions, for some matrices this is not sufficient to insure rapid convergence of the algorithm and so shifts of origin are necessary [Francis(p271), 1961] and [Parlett(p118),1967]

Another reason for introducing origin shifts is to get around the problem of convergence when distinct eigenvalues have the same modulus. Such eigenvalues lie on circle in the complex plane (with centre at the origin) and shifting the origin by appropriately can change the fact that their moduli are equal.

It should be pointed out that when the algorithm converges [Nasier(p. 40), 1967] it is observed that the convergence of the sub-diagonal elements depends on the ration of the eigenvalues $|\lambda_i/ \lambda_{i-1}|$, and unless this ratio is small the rate of convergence can be very unsatisfactory.

Therefore by means of incorporating shifts (viz. Single shift, Double shift) in the QR algorithm convergence rate to upper quasi-triangular form can be accelerated.

Explicitly Shifted QR

The lineage of the QR algorithm can be traced back to a theorem of [Konig, 1884], which states that if the function defined by the power series

$$a_0 + a_1 + a_2z^2 + \ldots.$$

has one simple pole r at its radius of convergence then $r = \lim_{k\to\infty} a_k/a_{k+1}$. Later [Hadamard, 1892] gave detrimental expressions in terms of the coefficients of the above equation for poles beyond the radius of convergence. [Aitken, 1926] independently derived these expressions and gave recurrences for their computation. [Rutishauser, 1954a, 1954b] consolidated and extended the recurrences in his QD-algorithm. In 1955, [Rutishauser, 1955] then showed that if the numbers from one stage of the QD algorithm were arranged in a tridiagonal matrix T then the numbers for the next stage could be found by computing the LU factorization of T and multiplying the factors in reverse order. He generalized the process to a full matrix and showed that under certain conditions the iterates converged to an upper triangular matrix whose diagonals were the eigenvalues of the original matrix. The result was what is known as the LR algorithm. Later [Rutishauser, 1958] introduced a shift to speed convergence. Except for special cases, like positive definite matrices, the LU algorithm is numerically unstable. [Kublanovskaya, 1961] and [Francis, 1961] independently proposed substituting the stable QR decomposition for the LU decomposition. However, Francis went beyond a simple substitution and produced the algorithm that we use today, complete with a preliminary reduction to Hessenberg form and the implicit double shift for real matrices.

The name of the QR decomposition comes from the fact that in Francis's notation the basic step was to decompose the matrix A into the product QR of an orthogonal matrix and an upper triangular matrix. Thus the algorithm preceded the decomposition. It is also interesting

to note that Q was originally O for "orthogonal" and was changed to Q to avoid confusion with zero. For more see Parlett's survey [Parlett, 1964] and the paper by [Fernando and Parlett, 1994]. The relation of the QR algorithm to the inverse power method was known to J. H. Wilkinson and W. Kahan [Parlett, p. 173, 1980], although the first explicit statement of the relation is in a textbook by [Stewart p. 352, 1973]. The relation to the power method was also known from first, as was the convergence which was originally established using determinants.[Wilkinson, 1965b], [Wilkinson(Sec. 29-30 Ch. 8), 1965a]. The iteration converges quadratically to a non-defective multiple eigenvalue. For the special case of the Rayleigh-quotient shift, this result is implied by work of Ostrowski on the Rayleigh-quotient method [Ostrowski, 1959], however, no general analysis seems to have appeared in the literature. [Watkins and Eisner, 1991] have given a very general convergence theory that holds for both the QR and LR algorithms.

The Hessenberg QR algorithm

The reduction of general matrix to Hessenberg form is due to [Householder, 1958] and [Wilkinson, 1960], who worked out the numerical details. The stability analysis is due to Wilkinson. His general approach to backward error analysis of orthogonal transformations is given in his Algebraic Eigenvalue Problem (1965) and is the basis for most of the rounding-error results [Wilkinson(Sec. 6.8-19), 1965]. It is possible to use non-orthogonal transformations to reduce a symmetric matrix (as already seen) to tridiagonal form, a process that bears the same relation to Givens' and Householder's method as Gaussian elimination bears to orthogonal triangularization. The case for non-orthogonal reduction, however, is not as compelling as the case for Gaussian elimination, and the method is not used in the major packages.

The implementation of the QR algorithm is a large task, and most of the algorithms do not serve as a complete implementation. For further details, there is little to do but go to the better packages and look at the code. The Handbook ALGOL codes [Wilkinson, 1971] are well documented but show their age, as do their EISPACK translations into FORTRAN [Garbow, Boyle, Dongarra, and Moler, 1977]. The LAPACK [Andersan et. al, 1995] codes are state of the art, at least for large matrices but they are undocumented, sparsely commented, and difficult to read. The observation that the QR step preserves Hessenberg form is due to [Francis, 1961].

The Wilkinson shift, Deflation and Preprocessing

The Wilkinson shift was proposed by [Wilkinson, 1968] for symmetric tridiagonal matrices, where it insures global convergence of the QR algorithm. Its use in the general QR algorithm is traditional.

A personal anecdote may illustrate the difficulties in finding satisfactory criteria for declaring a sub-diagonal element zero. The idea of deflating when there are two consecutive small sub-diagonal elements is due to [Francis, 1961]. The computational advantages of this procedure are obvious. In addition, many people felt that such small elements could retard the convergence of the implicitly shifted QR algorithm (to be discussed in the next section). However, [Watkins, 1995] has shown that the shift is accurately transmitted through the small elements. A personal anecdote may illustrate the difficulties in finding satisfactory criteria for declaring a sub-diagonal element zero.

The balancing algorithm [Stewart(p. 107), 2001] is due to [Osborne, 1960] and was used, along with the strategy for the initial deflation of eigenvalues, in the Handbook [Parlett and Reinsch, 1969] and in subsequent packages.

Implicitly Shifted QR

The ideas in this section, from the real Schur form to the implicit double shift, are due to [Francis, 1961], Although focus has been more on the double shift version, there is a single shift variant, which we still in use to compute eigenvalues of a symmetric tridiagonal matrix. There is also a multishift variant (see below), which is implemented in the LAPACK routine xHSEQR. The idea of an implicit shift has consequences extending beyond the computation of a real Schur form.

When a 2 x 2 block has converged, it should be split if it has real eigenvalues. The LAPACK routine XLANV2 does this. It also transforms a block with complex eigenvalues into one in which the diagonal elements are equal. The implicit QR shifts can be also be done by incorporating multiple shifts, there is no need to confine ourselves to two shifts. For example, when we treat the symmetric eigenvalue problem, we will use single shifts. We can also use more than two shifts. Specifically, suppose that we have chosen m shifts k_1, k_2, k_3,...k_m. We compute the first column c of $(A - k_mI)$..... $(A - k_1I)$, which will have m + 1 leading nonzero components. We then determine an (m+l) x (m+l) Householder transformation Q_0 that reduces c to e_1 and then apply it to A. The matrix Q_0AQ_0 has an (m+l) x (m+l) bulge below its diagonal, which is chased down out of the matrix in a generalization of Algorithm. The shifts can be computed as the eigenvalues of the trailing m x m sub-matrix.

This algorithm, proposed by [Bai and Demmel, 1989] and implemented in the LAPACK routine HSEQR, can take advantage of level-3 BLAS, the larger m the greater the advantage. Unfortunately, [Watkins, 1996] has shown that if the number of shifts is great they are not transmitted accurately through the bulge chasing process. An alternative exists if the bulge from a 2 x 2 shift is chased two steps down the diagonal then another double shift can be started after it. If this process is continued, the result is a sequence of small, adjacent bulges being chased down the matrix. This algorithm and a new deflation procedure is treated by [Braman, Byers, and Mathias, 1999]. The paper also contains an extensive bibliography on the subject.

There also exist variants to the QR algorithm. There can be four factorizations of a matrix into a unitary and a triangular matrix, which are represented symbolically by QR, QL, RQ, and LQ, where R and L stand for upper and lower triangular. The theory of the QR algorithm, applies to the other three variants. The shifted variants of the QR and QL algorithms exhibit fast convergence in the last and first rows; those of the RQ and the LQ, in the first and last columns. See [Stewart Sec. 1.3, Chapter 3, 2001] for details on QL algorithm.

Non-Symmetric Lanczos Method and Arnoldi Iteration

For nonsymmetric matrices it took longer for the Krylov subspace methods to gain popularity. An influential paper, that helped to promote Arnoldi's method as a useful tool, was published by [Saad, 1980]. The Arnoldi method, for orthogonal reduction to upper Hessenberg form, was not only too expensive if one wanted to know only a few eigenpairs, it also suffered from poor convergence for specific eigenvalue distributions. Well-known is the [Saad-Schultz, 1986] example, which is a permuted identity matrix. The method leads to trivial approximations after the first n − 1 steps and after n steps all eigenpairs suddenly appear. This however, is at a much higher cost than for Householder's reduction. Lanczos is seen as the extension of Arnoldi for the non-hermitian matrices; however, the non-hermitian Lanczos algorithm is an oblique projection technique and is quite different in concept from Arnoldi's method. From the practical point of view, a big difference between the non-Hermitian Lanczos procedure and the Arnoldi methods is that we now only need to save a few vectors in memory to execute the algorithm if no re-orthogonalization is performed. More precisely, we

need 6 vectors of length n plus some storage for the tri-diagonal matrix, no matter how large m is. This is clearly a significant advantage.

Lanczos Method for non-Hermitian matrices

Non-Symmetric Lanczos process, also referred to as the two-sided Lanczos method, received quite a lot of attention. Initially, the method was notorious for its break-down possibilities, its behaviour in finite precision arithmetic and the fact that the reduction operators to tridiagonal form are non-orthogonal. During the 1980s, much work was devoted to exploiting the basic non-Hermitian Lanczos algorithm by [Parlett, Taylor and Liu, 1985] and by [Cullum and Willoughby, 1984, 1985] and [Cullum, Kerner and Willoughby, 1989]. The first successful application of the code in a real life problem seems to be in the work by [Carnoy and Geradin, 1982], who used a version of the algorithm in a finite element model. [Cullum and Willoughby, 1986], presented a code based on the two-sided Lanczos method, in which they solved a number of practical problems; this included a clever trick for identifying the spurious eigenvalues due to rounding errors. The code gained some popularity, for instance for plasma-physics eigenvalue computations [Parlett, 1986]. [Parlett and co-workers, 1985] introduced the concept of "look-ahead", mainly in order to improve the numerical stability of the process. The look-ahead strategy, introduced in order to prevent breakdown, was further perfected by [Freund and Nachtigal, 1996]. They published a code based on quasi-minimization of residuals, and included look-ahead strategies, in which most of the original Lanczos problems were repaired (but the non-orthogonal reductions were still there). [Gutknecht, 1992] published a thorough theoretical overview of the two-sided Lanczos algorithm and exploited its relation to Pade approximations. This gave a better understanding of look-ahead strategies and the convergence behaviour of the method (in the context of solving linear systems). Block variants of the two-sided Lanczos process [Demmel et. al., 1999] were discussed in Day's Ph.D. thesis in 1993.

Arnoldi Iteration

Arnoldi Iteration technique finds the eigenvalues of general (possibly non-Hermitian) matrices (an analogous method for Hermitian matrices is the Lanczos iteration). The Arnoldi iteration was invented by W. E. Arnoldi in 1951. We mention firstly polynomial preconditioning, discussed extensively in Saad's book [Saad, 1992], which damps unwanted parts of the spectrum, and secondly, sophisticated restarting strategies. The method becomes effective for matrices for which shift-and-invert operations can be applied for given vectors. But the many (increasingly expensive) iterations for relevant problems were a bottleneck. A real breakthrough for the Arnoldi method was realized by [Sorensen, 1991] with the so-called Implicit Restart Technique. This is a clever technique by which unwanted information can be filtered away from the process. This leads to a reduced subspace with a basis, for which the matrix still has a Hessenberg form, so that Arnoldi's process can be continued with a subspace (rather than with a single vector as with the more classical restart techniques).

The equivalent Block Arnoldi algorithm, has not been given much attention, except in control problems where it is closely associated with the notion of controllability for the multiple-input case [Boley and Golub, 1987]. In fact Arnoldi's method (single input case) and its block analogue (multiple input case) are useful in many areas in control; see for example [Saad, 1990a, 1990b].

The term iterative method, used to describe Arnoldi, can perhaps be somewhat confusing. Note that all general eigenvalue algorithms must be iterative. This is not what is referred to when we say Arnoldi is an iterative method. Rather, Arnoldi belongs to a class of linear algebra algorithms (based on the idea of Krylov subspaces) that give a partial result after a

relatively small number of iterations. This is in contrast to so-called direct methods, which must complete to give any useful results.[Wikipedia-6]

Arnoldi iteration is a typical large sparse matrix algorithm: It does not access the elements of the matrix directly, but rather makes the matrix map vectors and makes its conclusions from their images. This is the motivation for building the Krylov subspace. There have been several papers published on Arnoldi's method and its variants for solving eigenproblems. The original paper by [Arnoldi, 1951] came out about one year after Lanczos' breakthrough paper [Lanczos, 1950] and is quite different in nature. The author hints that his method can be viewed as a projection method and that it might be used to approximate eigenvalues of large matrices. Note that the primary goal of the method is to reduce an arbitrary (dense) matrix to Hessenberg form. At the time, the QR algorithm was not yet invented, so the Hessenberg form was desired only because it leads to a simple recurrence for the characteristic polynomial. The 1980 paper by [Saad, 1980] showed that the method could indeed be quite useful as a projection method to solve large eigenproblems, and gave a few variations of it. Later, sophisticated versions have been developed and used in realistic applications, see [Natarajan, 1992], [Natarajan and Acrivos, 1992], [Papanastasiou, 1990], [Parlett and Saad, 1987].

4. THE GENERALIZED EIGENVALUE PROBLEM

The generalized eigenvalue problem (regarded as a pair of bilinear forms) for finding nontrivial solutions of the equations

$$Ax = \lambda Bx,\ Ax + \lambda Bx + \lambda^2 Cx = 0$$

goes back at least to [Weierstrass,1 868], who established the equivalent of the Jordan form for regular matrix pencils. In [Gantmacher, 1959] the matrix $A - \lambda B$ is called a matrix pencil. The rather strange use of the word "pencil" comes from optics and geometry: an aggregate of (light) rays converging to a point does suggest the sharp end of a pencil and, by a natural extension, the term came to be used for any one parameter family of curves, spaces, matrices, or other mathematical objects. [Jordan, 1874] later gave a new proof for singular pencils. The problem reduces to the ordinary eigenvalue problem when $B = I$, which is why it is called a generalized eigenvalue problem. Although the generalization results from the trivial replacement of an identity matrix by an arbitrary matrix B, the problem has many features not shared with the ordinary eigenvalue problem. For example, a generalized eigenvalue problem can have infinite eigenvalues. In spite of the differences between the two problems, the generalized eigenvalue problem has the equivalents of a Hessenberg form and a Schur form. Moreover, the QR algorithm can be adapted to compute the latter; the resulting algorithm is widely known as the QZ algorithm. The power method and RQI are techniques for solving the homogeneous system of equations $(A - \lambda B)X = O$ (or the matrix pencils). Of course A is unknown and so the problem is not linear. Nevertheless almost every known technique for solving linear systems yields an analogous iteration for the Non-Linear eigenvalue problem. For example, there is a successive overrelaxation (SOR) method which can be very effective for special problems when triangular factorization is not possible. See [Ruhe, 1975 and 1977] for a full treatment of these ideas. [Kronecker, 1890] extended the result to rectangular pencils. For modern treatments refer to some of the advanced texts: [Wilkinson, 1979], [Dooren, 1979], [Gantmacher, 1959], [Kagstrom and Wiberg, 1999] and [Stewart, 1990]. The generalized Schur form is due to [Stewart, 1972].

Linear Eigenvalue Problem

In linear eigenvalue problems $Ax = \lambda x$ or $Ax = \lambda Bx$, well established methods are available, that include error and condition estimates. These methods are able to deal with most of the small or large scale problems in practice and specific methods have been developed to handle extra structures [Anderson et. al., 1995], [Benner et. al., 2002], [Benner and Faßbender, 1998], [Mehrmann, 1991], [Mehrmann and Watkins, 2001 and 2002].

Generalized Schur form

Just as we can reduce a matrix to simpler forms by similarity transformations, so can we simplify a pencil by equivalence transformations. For example, there is an analogue of the Jordan form for regular pencils called the Weierstrass form. Unfortunately, this form, like the Jordan form, is not numerically well determined, and the transformations that produce it can be ill conditioned [Stewart(Sec. 1.6, Chapter 1), 2000]. Fortunately, if we restrict ourselves to unitary equivalences, which are well conditioned, we can reduce the matrices of a pencil to triangular form.

The Schur form has enabled us to solve the mathematical problem of defining the multiplicity of generalized eigenvalues—finite or infinite. It also suggests how to put finite and infinite eigenvalues on the same notational footing.

Perturbation Theory

For a general treatment of the perturbation theory of the generalized eigenvalue problem along with a bibliography see [Stewart, 1990]. The use of the chordal metric in generalized eigenvalue problems is due to [Stewart, 1975], who is happy to acknowledge a broad hint from W. Kahan. For a chordless perturbation theory and backward error bounds see [Higham and Higham, 1998].

QZ Algorithm

Like the QR algorithm, the doubly shifted QZ algorithm is an iterative reduction of a real Hessenberg-triangular pencil to real generalized Schur form. The QZ algorithm, including the preliminary reduction to Hessenberg-triangular form, is due to [Moler and Stewart, 1973]. The double shift strategy, of course, works when the eigenvalues of the trailing pencil are real. However, [Ward, 1975] has observed that the algorithm performs better if real shifts are treated by a single shift strategy. This option has been incorporated into the EISPACK and LAPACK codes. The balancing algorithm is due to [Ward, 1981], who builds on earlier work of [Curtis and Reid, 1972]. [Ward, 1975] and [Watkins, 2000] consider the problem of infinite eigenvalues. In particular, both the reduction to Hessenberg-triangular form and the steps of the QZ algorithm move zero elements on the diagonal of B toward the top. Moreover, Watkins shows that the presence of small or zero elements on the diagonal of B does not interfere with the transmission of the shift. Thus an alternative to deflating infinite eigenvalues from the middle of the pencil is to wait until they reach the top, where they can be deflated by a single rotation. For the treatment of 2x2 blocks see the EISPACK code qzval or the LAPACK codeSLAG2.

There does not seem to be any exhaustive coverage of the generalized eigenvalue problems, theory and algorithms. In addition, there seems to be a dichotomy between the need of users, mostly in finite elements modeling, and the numerical methods that numerical analysts develop. One of the first papers on the numerical solution of quadratic eigenvalue problems is [Borri and Mantegazza, 1977]. Quadratic eigenvalue problems are rarely solved in structural engineering. The models are simplified first by neglecting damping and the leading eigenvalues of the resulting generalized eigenproblem are computed. Then the eigenvalues of the whole problem are approximated by performing a projection process onto the computed

invariant subspace of the approximate problem [Jennings, 1981]. This may very well change in the future, as models are improving and computer power is making rapid gains.

Reduction to Hessenberg-triangular form

The first step in the solution of the ordinary eigenvalue problem by the QR algorithm is a reduction of the matrix in question to Hessenberg form, followed by a reduction of a pencil (A, B) to a form in which A is upper Hessenberg and B is triangular. Note that if A is upper Hessenberg and B is lower triangular, then AB^{-1} is upper Hessenberg. Thus the reduction to Hessenberg-triangular form corresponds to the reduction to Hessenberg form of the matrix AB^{-1}. The process begins by determining an orthogonal matrix Q such that Q^TB is upper triangular. The matrix Q can be determined as a product of Householder transformations [Refer any book] .The transformation Q^T is also applied to A (considering a 5 x 5 matrix). Now plane rotations are used to reduce A to Hessenberg form while reserving the upper triangularity of B. The reduction proceeds by columns. Zeros are introduced in A beginning at the bottom of the first column. The elimination of the (5, l)-element of A by pre-multiplication by the rotation Q_{45} in the (4, 5)-plane introduces a nonzero into the (5,4)-element of B. This nonzero element is then annihilated by post-multiplication by a rotation Z_{54} in the (5,4)-plane. The annihilation of the (4, 1) and (3, l) elements of A proceeds similarly.

$$\left[\begin{pmatrix} a & a & a & a & a \\ a & a & a & a & a \\ a & a & a & a & a \\ a & a & a & a & a \\ \widehat{a} & a & a & a & a \end{pmatrix} \begin{pmatrix} b & b & b & b & b \\ 0 & b & b & b & b \\ 0 & 0 & b & b & b \\ 0 & 0 & 0 & b & b \\ 0 & 0 & 0 & 0 & b \end{pmatrix}\right] \xrightarrow{Q_{45}(A,B)}$$

$$\left[\begin{pmatrix} a & a & a & a & a \\ a & a & a & a & a \\ a & a & a & a & a \\ a & a & a & a & a \\ 0 & a & a & a & a \end{pmatrix} \begin{pmatrix} b & b & b & b & b \\ 0 & b & b & b & b \\ 0 & 0 & b & b & b \\ 0 & 0 & 0 & b & b \\ 0 & 0 & 0 & \widehat{b} & b \end{pmatrix}\right] \xrightarrow{(A,B)Z_{54}}$$

$$\left[\begin{pmatrix} a & a & a & a & a \\ a & a & a & a & a \\ a & a & a & a & a \\ \widehat{a} & a & a & a & a \\ 0 & a & a & a & a \end{pmatrix} \begin{pmatrix} b & b & b & b & b \\ 0 & b & b & b & b \\ 0 & 0 & b & b & b \\ 0 & 0 & 0 & b & b \\ 0 & 0 & 0 & 0 & b \end{pmatrix}\right] \xrightarrow{Q_{34}(A,B)}$$

$$\left[\begin{pmatrix} a & a & a & a & a \\ a & a & a & a & a \\ a & a & a & a & a \\ 0 & a & a & a & a \\ 0 & a & a & a & a \end{pmatrix} \begin{pmatrix} b & b & b & b & b \\ 0 & b & b & b & b \\ 0 & 0 & b & b & b \\ 0 & 0 & \widehat{b} & b & b \\ 0 & 0 & 0 & 0 & b \end{pmatrix}\right] \xrightarrow{(A,B)Z_{43}}$$

$$\left[\begin{pmatrix} a & a & a & a & a \\ a & a & a & a & a \\ \widehat{a} & a & a & a & a \\ 0 & a & a & a & a \\ 0 & a & a & a & a \end{pmatrix} \begin{pmatrix} b & b & b & b & b \\ 0 & b & b & b & b \\ 0 & 0 & b & b & b \\ 0 & 0 & 0 & b & b \\ 0 & 0 & 0 & 0 & b \end{pmatrix}\right] \xrightarrow{Q_{23}(A,B)}$$

$$\left[\begin{pmatrix} a & a & a & a & a \\ a & a & a & a & a \\ 0 & a & a & a & a \\ 0 & a & a & a & a \\ 0 & a & a & a & a \end{pmatrix} \begin{pmatrix} b & b & b & b & b \\ 0 & b & b & b & b \\ 0 & \widehat{b} & b & b & b \\ 0 & 0 & 0 & b & b \\ 0 & 0 & 0 & 0 & b \end{pmatrix}\right] \xrightarrow{(A,B)Z_{32}}$$

$$\left[\begin{pmatrix} a & a & a & a & a \\ a & a & a & a & a \\ 0 & a & a & a & a \\ 0 & a & a & a & a \\ 0 & a & a & a & a \end{pmatrix} \begin{pmatrix} b & b & b & b & b \\ 0 & b & b & b & b \\ 0 & 0 & b & b & b \\ 0 & 0 & 0 & b & b \\ 0 & 0 & 0 & 0 & b \end{pmatrix}\right]$$

\

This process is repeated on columns 2 through n-2 of the pencil, to obtain the Hessenberg-triangular form.

Iterative projection methods for linear problems

For sparse linear eigenvalue problems $Ax = \lambda Bx$, iterative projection methods like the Lanczos, Arnoldi, rational Krylov or Jacobi–Davidson method are well established. The basic idea of all these methods is the construction of a search space (typically a Krylov subspace) followed by projection of the problem into this subspace. This leads to a small dense problem that is then handled by a dense solver and the eigenvalues of the projected problem are used as approximations to the eigenvalues of the large sparse problem. The main feature of all these methods is that matrix factorizations are avoided as much as possible (except in the context

of preconditioning) and the generation of the search space is usually done via an iterative procedure that is based on matrix vector products that can be cheaply obtained.

Two basic types of iterative projection methods are in use: The first type consists of methods which expand the subspaces independently of the eigenpair of the projected problem and which use Krylov subspaces of A or $(A - \sigma I)^{-1}$ for some shift σ. These methods include the Arnoldi, Lanczos or rational Krylov method. The Arnoldi method together with its variants is today a standard solver for sparse linear eigenproblems. It is implemented in the package ARPACK [Davidson, 1975] and the MATLAB command eigs, see also [Baglama, Calvetti and Reichel, 1996]. The method typically converges to the extreme eigenvalues first.

The other type of methods like the Jacobi-Davidson, aim at a particular eigenpair and choose the expansion q such that it has a high approximation potential for a desired eigenvalue/eigenvector or invariant subspace.

Non Linear Eigenvalue Problem

For nonlinear eigenvalue problems, there are essentially no analogous packages that reach the standard of those for linear problems. However, there is a vast literature on numerical methods for nonlinear eigenvalue problems. In general, one has to distinguish between dense and large sparse problems. For dense problems, the size of the problems that can be treated with numerical methods is limited to a few thousand depending on the available storage capacities. Methods for small dense problems, however, are needed in most of the iterative projection methods for large sparse problems.

The theoretical analysis and the numerical methods for polynomial eigenvalue problems usually proceed via linearization, i.e. via the embedding of the nonlinear eigenvalue problem into a larger linear eigenvalue problem [Gohberg, Lancaster and Rodman, 1982].

Newton type methods and inverse Iteration

For general nonlinear eigenvalue problems, the classical approach is to formulate the eigenvalue problem as a system of nonlinear equations and to use variations of Newton's method or the inverse iteration method. An improved version of this method was suggested in [Jain, Singhal and Huseyin, 1983], [Jain and Singhal, 1983] and also quadratic convergence was shown. A similar approach was presented in [Yang, 1983], via a representation of Newton's method using the LU-factorization. Other variations of this method can be found in [Zurm and Falk, 1984, 1986]. However, this relatively simple idea is not efficient, since it computes eigenvalues one at a time and needs several $O(n^3)$ factorizations per eigenvalue. It is, however, useful in the context of iterative refinement of computed eigenvalues and eigenvectors.

Another method that also serves the purpose of iterative refinement is the nonlinear version of inverse iteration. The disadvantage of inverse iteration with respect to efficiency is the large number of factorizations that are needed for each of the eigenvalues. The obvious idea then is to use a version of a simplified Newton method, where the shift σ is kept fixed during the iteration.

Iterative projection methods for nonlinear problems

Arnoldi type methods are quite efficient in solving sparse nonlinear eigenproblems if an accurate

preconditioner is at hand. If this is not the case, then the convergence deteriorates considerably. In this situation Jacobi–Davidson type methods offer an appealing alternative. As in the linear case the correction equation does not have to be solved exactly to maintain fast convergence, but usually a few steps of a Krylov solver with an appropriate preconditioner suffice to obtain a good expansion direction of the search space.

A different approach was developed in [Ruhe, 1996a, 1996b, 2000, 2004], which generalizes the rational Krylov method for linear eigenproblems [Ruhe, 1998] to sparse nonlinear eigenvalue problems by nesting the linearization of problem (by regula falsi) and the solution of the resulting linear eigenproblem (by Arnoldi's method). Hence, in this approach the two numerical subtasks, i.e. reducing the large dimension to a much smaller one and solving a nonlinear eigenproblem are attacked simultaneously. This method was applied in [Hager and Wiberg, 2000], [Hager, 2001] to the rational eigenvalue problem governing damped vibrations of a structure.

Iterative projection methods for nonlinear eigenvalue problems have been studied extensively in recent years, but still have not reached a stage of maturity as have the methods for linear problems. More research is needed to improve the current methods, to analyse their stability and convergence behavior, in particular in finite precision, and to generate implementations that can be used in a simple and robust way by non-experts.

5. MODERN APPROACHES

In the last few years, rapid improvement in computer architecture and substantial advances in algorithmic research have enabled application scientists to tackle eigenvalue problems with tens of millions degrees of freedom. In this section we briefly discuss the recent developments in the solution of large scale eigenvalue problems giving the state of art of the algorithms that are in practical use for solving the problems and providing proper references to places with more extensive coverage on the subject. The detailed discussion on this topic can be found in [Yang, 2005].

Krylov Subspace Method

Krylov subspace methods (KSM) remain the most reliable and effective tools for solving large-scale eigenvalue problems. In a KSM, approximations to the desired eigenpairs of an n by n matrix A are extracted from a k-dimensional Krylov subspace

$$K(v_0,A; k) = \text{span}\{v_0,Av_0, \cdots ,A^{(k-1)}v_0\},$$

where v_0 is often a randomly chosen starting vector and $k \ll n$. It is well known that dominant eigenvalues well separated from the rest of the spectrum of A converge rapidly in a standard KSM (e.g., the Lanczos or Arnoldi iteration). Furthermore, if the starting vector v_0 contains a linear combination of a few eigenvectors corresponding to the desired eigenvalues, then $K(v_0,A; k)$ becomes invariant for a small value of k, and the eigenvalues of the projected matrix H_k are indeed the eigenvalues of A.

Two types of strategies have been developed to accelerate the convergence of KSM. The implicitly restarting technique originally developed in [Sorensen, 1992] and implemented in ARPACK [Lehoucq, 1998]is designed to gradually filter out the unwanted spectral components from v_0 by applying the implicitly shifted QR algorithm to the projected matrix H_k. The filter applied in ARPACK is a polynomial filter with roots set near the unwanted eigenvalues. This technique can be extended to include rational filters with poles set near the eigenvalues of interest [Sorensen and Yang, 1998], [Yang, 1998]. The use of implicit restart enables one to extract desired eigenpairs from a Krylov subspace of small dimension, thereby keeping the cost of orthogonalization low.

The second type of techniques for enhancing the convergence of KSM involves transforming the original eigenvalue problem into one that has a more favourable eigenvalue distribution. This technique is particularly helpful when the eigenvalues of interest are near a target shift σ in the interior of the spectrum. In this case, computing the largest eigenvalues of the shifted inverse $(A-\sigma I)^{-1}$ is often more effective than computing those of A directly. Although rational transformation is the most commonly used spectral transformation, one may also use polynomial transformations [Sorensen and Yang, 1997] when it is prohibitively expensive to factor $A-\sigma I$ into a product of triangular matrices.

Alternative Methods

The difficulty of introducing a preconditioner into a KSM in a straightforward fashion is fundamentally related to the fact that the eigenvectors of a preconditioned matrix $P^{-1}A$ is generally different from the eigenvectors of A. Thus, building a Krylov subspace in terms of $P^{-1}A$ does not readily provide approximation to the eigenvectors of A directly.

Two Alternatives have been pursued in the last few years to overcome this problem.

Solving the Eigenvalue Problem as an Optimization Problem

For problems in which A or K is symmetric (or Hermitian) and the eigenvalues of interest are the smallest or the largest, one may solve the eigenvalue problem as a constrained optimization problem.

$$\min_{x^T x=1} \rho(x) = x^T A x$$

The use of a preconditioner in this formulation simply amounts to a change of variable. The constrained optimization approach has been effective for solving large-scale eigenvalue problems arising from the self-consistent field iteration used in electronic structure calculation.

Solving the Eigenvalue Problem as a System of Non-linear Equations

An eigenvalue problem can also be viewed as a set of nonlinear equations

$$Ax = (x^T A x)x,\ x^T x = 1.$$

By treating an eigenvalue problem as a nonlinear system, one can exploit the possibility of using Newton's method to compute the desired eigenvalues and eigenvectors. Although significant progress has been made in large-scale eigenvalue calculation research, a number of challenges still remain. This is particularly true in the area of non-linear eigenvalue problems.

This approach leads to the Jacobi-Davidson (JD) [Sleijpen and Vorst, 1996] algorithm, which is an extension of the [Davidson, 1975] and [Olsen, Jorgensen and Simons, 1990] algorithms. The method can be described as an inner-outer iteration. Approximations to the desired eigenvalues and eigenvectors are computed in the outer iteration by projecting A into a subspace V.

Successive Eigenvalue Relaxation

[Ovtchinnikov and Xanthis, 2000] presented a new subspace iteration method for the efficient computation of several smallest eigenvalues of the generalized eigenvalue problem $Ax = \lambda Bx$ for symmetric positive definite operators A and B. The SER method is the homoechon of the classical successive over-relaxation, or SOR method for linear systems. In particular, there are two significant features of SER which render it computationally attractive: (i) it can effectively deal with preconditioned large-scale eigenvalue problems, and (ii) its practical implementation does not require any information about the preconditioner used: it can routinely accommodate sophisticated preconditioners designed to meet more exacting requirements (e.g. three-dimensional elasticity problems with small thickness parameters). SER is endowed with theoretical convergence estimates which cover the case of multiple eigenvalues and eigenvalue clusters and identify the important parameters essentially affecting the convergence.

Given an approximation I^n to the invariant subspace I corresponding to the first m eigenvalues of

$Ax = \lambda Bx$, we choose a number k_n, $0 < k_n < m$, and define the new approximation as

$$I^{n+1} = \text{span}\{u_i(I^n)\}_{i=0,\, m-1}, \quad I^n = I^n + \text{span}\{r_{kn}(I^n)\}.$$

In other words, we add the residual on the Ritz vector $u_{kn}(I^n)$ to the set of Ritz vectors in I^n, then we apply the Rayleigh-Ritz projection onto the subspace I^n spanned by these m + 1 vectors and we define the new approximation I^{n+1} to the invariant subspace I as the span of the first m Ritz eigenvectors in In (the last one is discarded so that the dimension of I^{n+1} is the same as I^n). The above iterative procedure is known as the successive eigenvalue relaxation, or SER method. The convergence rate of SER iterations depends on the choice of k_n. The purpose of introducing the auxiliary subspaces is to ensure that on each step we 'relax' the Ritz eigenvalues corresponding to Ritz eigenvectors which approximate different eigenvectors of $Ax = \lambda Bx$. For full computational details see the paper by [Ovtchinnikov and Xanthis, 2000].

SER radically overcomes difficulties highlighted above, thus eloquently justifying its raison d 'etre and unique role amongst other eigensolvers (cf. Prolegomena and Introduction). Although SER is particularly attractive for pre- conditioned large-scale problems, it also performs reasonably well for other classes of problems unrelated to preconditioning. The full computational capabilities of SER will emerge when one undertakes extensive numerical (large-scale) computations to demonstrate its performance and relative merits compared with other methods (eigensolvers).

BIBLIOGRAPHY

- C. Aitken, On Bernoulli's numerical solution of algebraic equations. Proceedings of the Royal Society of Edinburgh, 46:289-305,1926.
- C. Aitken, Studies in practical mathematics, II: The evaluation of the latent roots and latent vectors of a matrix. Proceedings of the Royal Society of Edinburgh, 57:269-304,1937.
- A. Anda and H. Park, Self-scaling fast rotations for stiff least squares problems. Linear Algebra and Its Applications, 234:137-162,1996.
- Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen. LAPACK Users' Guide. SIAM, Philadelphia, second edition, 1995.

- W. E. Arnoldi, The principle of minimized iterations in the solution of the matrix eigenvalue problem, Quarterly of Applied Mathematics, Volume 9, pages 17–29, 1951.
- E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen. LAPACK Users' Guide. SIAM Publications, Philadelphia, second edition, 1995.
- Z. Bai and J. Demmel, On a block implementation of Hessenberg QR iteration. International Journal of High Speed Computing, 1:97-112, 1989.
- Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, H. van der Vorst, Templates for the Solution of Algebraic Eigenvalue Problems: a Practical Guide, SIAM, Philadelphia, PA, 1999.
- J. Baglama, D. Calvetti, and L. Reichel, Iterative methods for the computation of a few eigenvalues of a large symmetric matrix. BIT, 37:400-440,1996.
- S. Batterson and J. Smillie, The dynamics of Rayleigh quotient iteration, SIAM J. Numer.Anal. 26 (1989), 624-636.
- Steve Batterson and John Smillie, Rayleigh Quotient Iteration for nonsymmetric matrices, Mathematics of Computation, Vol 5, pages 169-171, 1990.
- F. L Bauer, Das Verfahren der Treppeniteration und verwandte Ver- fahren zur Losung algebraischer Eigenwertprobleme, Z. angew. Math. Phys. 8, 214-235, 1957.
- F. L Bauer, On modern matrix iteration processes of Bernoulli and Graeffe type, J. Ass. corp. Mach. 5, pages 246-257, 1958.
- F. L. Bauer, Sequential reduction to tridiagonal form. J. Soc. industr. appl. Math. 7, 107-113, 1959.
- C. Bavely and G. W. Stewart, An algorithm for computing reducing subspaces by block diagonalization. SIAM Journal on Numerical Analysis, 16:359-367, 1979.
- Boley and G. H. Golub, The Lanczos-Arnoldi algorithm and controllability. Systems and Control Letters, 4:317-324, 1987.
- M. Borri and P. Mantegazza, Efficient solution of quadratic eigenproblems arising in dynamic analysis of structures. Comp. Meth. Appl. Mech. and Engng, 12:19-31, 1977.
- E. Bodewig, Matrix Calculus (2nd edition), Interscience Publishers, New York, North-Holland Publishing Company, Amsterdam, 1959.
- K. Braman, R. Byers, and R. Mathias, The multi-shift QR-algorithm: Aggressive deflation, maintaining well focused shifts, and level 3 performance. Technical report, Department of Mathematics, University of Kansas, Lawrence, 1999.
- P. Benner and H. Faßbender, The symplectic eigenvalue problem, the butterfly form, the SR algorithm, and the Lanczos method. Linear Algebra Appl., 275/276:19–47, 1998.
- P. Benner, R. Byers, V. Mehrmann, and H. Xu, Numerical computation of deflating subspaces of skew Hamiltonian/Hamiltonian pencils. SIAM J. Matrix Anal. Appl., 24:165–190, 2002.
- R. A. Brooker and F. H. Sumker, The method of Lanczos for calculating the characteristic roots and vectors of a real symmetric matrix, Proc. Instn elect. Engrs B.103 Suppl., pages 114-119, 1956
- Calvetti, L. Reichel, and D.C. Sorensen , An Implicitly Restarted Lanczos Method for Large Symmetric Eigenvalue Problems, 1994.
- Carnoy and M. Geradin. On the practical use of the Lanczos algorithm in finite element applications to vibration and bifurcation problems. In Axel Ruhe, editor,

Proceedings of the Conference on Matrix Pencils, Lulea, Sweden, March 1982, pages 156-176, New York, 1982. University of Umea, Springer Verlag.

- Robert L. Causey , Computing eigenvalues of non-Hermitian matrices by methodes of Jacobi type, J. Soc. Indust. Appl. Math. 6, pages 172-181, 1958.
- M. Clint and A. Jennings, The evaluation of eigenvalues and eigenvectors of real symmetric matrices by simultaneous iteration method. Journal of the Institute of Mathematics and its Applications, 8:111-121, 1971.
- F. J. Corbato, On the coding of Jacobi's method for computing eigenvalues and eigenvectors of real symmetric matrices, J. Assoc. Comput. Mach. 10:123-125, 1963.
- J. Cullum and W. E. Donath, A Block Generalization of the Symmetric S-step Lanczos Algorithm, Report #RC 4845 (#21570), Yorktown Heights, IBM Thomas J. Watson Research Center, New York, 1974.
- J. Cullum and R. Willoughby, A Lanczos procedure for the modal analysis of very large nonsymmetric matrices. In Proceedings of the 23rd Conference on Decision and Control, Las Vegas, 1984.
- J. Cullum and R. Willoughby, Lanczos Algorithms for Large Symmetric Eigenvalue Computations, Birkhauser, Basel, 1985.
- J. Cullum, R.A. Willoughby, A practical procedure for computing eigenvalues of large sparse nonsymmetric matrices, in: J. Cullum, R.A. Willoughby (Eds.), Large Scale Eigenvalue Problems, Amsterdam, North-Holland, pp. 193-240, 1986.
- J. Cullum, W. Kerner, and R. Willoughby. A generalized non- symmetric Lanczos procedure. Computer Physics Communications, 53, 1989.
- A. R. Curtis and J. K. Reid, On the automatic scaling of matrices for Gaussian elimination. Journal of the Institute for Mathematics and Applications, 10:118-124,1972.
- D. Davis, Explicit Functional Calculus, Lin. Alg. and Its Applic. 6, 193-99, 1973.
- E. R. Davidson, The iterative calculation of a few of the lowest eigenvalues of corresponding eigenvectors of large real symmetric matrices, J. Comput. Phys. Vol 17, pages 87-94, 1975.
- J. W. Demmel, The condition number of equivalence transformations that block diagonalize matrix pencils. SIAM Journal on Numerical Analysis, 20:599-610, 1983.
- J. W. Demmel, Computing stable eigendecompositions of matrices. Linear Algebra and Its Applications, 79:163-193,1986.
- J. Demmel and B. Kagstrom, Computing stable eigendecompositions of matrix pencils. Linear Algebra and Its Applications, 88/89:139-186,1987.
- J.W. Demmel, K. Veselic, Jacobi's method is more accurate than QR 1204-1246, SIAM J. Matrix Anal. Appl. 13, 1992.
- J.W. Demmel, Applied Numerical Linear Algebra, SIAM, Philadelphia, PA, 1996.
- J. Descloux, Bounds for the Spectral Norm of Functions of Matrices, Numer. Math. 5, 185-90, 1963.
- Diederik R. Fokkema, Gerard L.G. Sleijpen, and Henk A, Van der Vorst, Jacobi-Davidson style QR and QZ algorithms for the partial reduction of matrix pencils , Preprint 941, Department of Mathematics, University Utrecht (January, 1996), SIAM J. Sc. Comput., 20:1 pp. 94-125, 1998.
- P. Van Dooren, The computation of Kronecker's canonical form of a singular pencil, Linear Algebra and Its Applications, 27:103-140,1979.

- P. J. Eberlein, A Jaoobi-like method for the automatic computation of eigenvalues and eigenvectors of an arbitrary matrix, J. Soc. industr. appl. Math. 10, pages 74-88, 1962.
- J. Erxiong and Z. Zhenye, A new shift of the QL algorithm for irreducible symmetric tridiagonal matrices, Linear Algebra AppL Vol 65, pages 261- 272, 1985.
- W. Feller and G. E. Forsythe, New matrix transformations for obtaining characteristic vectors. Quarterly of Applied Mathematics, 8:325-331,1951.
- Y. Feng, Backward Stability of Root Free QR Algorithms, Ph.D. Thesis. Berkeley, University of California, 1991.
- K. V. Fernando and B. N. Parlett. Accurate singular values and differential QD algorithms. Numerische Mathematik, 67:191-229,1994.
- J.G.F. Francis, The QR Transformation: A Unitary Analogue to the LR Transformation, Parts I and II Corny. J. 4, 265-72, 332-45, 1961.
- R. W. Freund, M. H. Gutknecht, and N. M. Nachtigal, An implementation of the Look-Ahead Lanczos algorithm for non-Hermitian matrices, Part I. Technical Report 90-11, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1990.
- R.W. Freund, N.M. Nachtigal, QMRPACK: a package of QMR algorithms, ACM Trans. Math. Software 22, 46-77, 1996.
- G. E. Forsythe, Singularity and near singularity in numerical analysis, Amer. math. Mon. Vol 65, pages 229-240, 1958.
- J.G.F. Francis, The QR Transformation, I, The Computer Journal, Vol. 4, No. 3, pages 265-271, 1961.
- J.G.F. Francis, The QR Transformation, II, The Computer Journal, Vol. 4, No. 4, pages 332-345 1962.
- S. Garbow, J. M. Boyle, J. J. Dongarra, and C. B. Moler. Matrix eigensystem routines-EISPACK guide extension. In Lecture Notes in Computer Science. Springer-Verlag, New York, 1977.
- F. R. Gantmacher, The Theory of Matrices, Vols. 1,11. Chelsea Publishing Company, New York, 1959.
- W. M. Gentleman, Error analysis of QR decompositions by Givens transformations. Linear Algebra and Its Applications, 10:189-197,1975.
- G. H. Golub and J. H. Wilkinson, Ill-conditioned eigensystems and the computation of the Jordan canonical form. SIAM Review, 18:578-619,1976.
- G. H. Golub and R. Underwood, The block Lanczos method for computing eigenvalue, In J. R. Rice, editor, Mathematical Software III, pages 361-377, Academic press, New York, 1977.
- G.H. Golub, C.F. Van Loan, Matrix Computations, The Johns Hopkins University Press, Baltimore, 1996.
- G.H. Golub, C.F. Van Loan, Matrix Computations, The Johns Hopkins University Press, Baltimore, 1996.
- Gene H. Golub, Henk A. van der Vorst, 150 years old and still alive: Eigenproblems SCCM, Stanford University, Stanford, USA, Department of Mathematics, Utrecht University, Utrecht, The Netherlands, 2000.
- Wallace Givens, Numerical Computation of the Characteristic Values of a Real Symmetric Matrix. Oak Ridge National Laboratory, Report ORNL-157 4, 1954.
- S. Gerschgorin, On bounding the eigenvalues of a matrix (in german) Vol 1 749-754,. Izv. Akad. Nauk. SSSR Otd Mat. Estest., 1931.

- I. Gohberg, P. Lancaster, and L. Rodman. Matrix Polynomials. Academic Press, New York, 1982.
- Goldstein, F. J. Murray and John Von Neumann, The Jacobi method for real symmetric matrices, J. Assoc. Comput. Mach. 6, pages 59-96, 1959.
- R. T. Gregory, Computing eigenvalues and eigenvectors of a symmetric matrix on the ILLIAC, Math. Tab., Wash. 7, pages 215-220, 1953.
- John Greenstadt, A method for finding roots of arbitrary matrices, Math. Tables Aids Comput. 9, pages 47-52, 1955.
- M. Gu, Finding well-conditioned similarities to block-diagonalized nonsymmetric matrices is NP-hard. Technical Report LBL-35100, Physics Division, Lawrence Berkeley Laboratory, Berkeley, CA, 1994.
- M.H. Gutknecht, A completed theory of the unsymmetric Lanczos process and related algorithms, Part I, SIAM J. Matrix Anal. Appl. 13 (1992) 594-639.
- M. J. Hadamard, Essai sur 1'etude des fonctions donnees par leur developpement de taylor. Journ. deMath. (4e serie), 8:101-186,1892.
- P. Hager and N.E. Wiberg, The rational Krylov algorithm for nonlinear eigenvalue problems. In B.H.V. Topping, editor, Computational Mechanics for the Twenty-First Century, pages 379 – 402. Saxe–Coburg Publications, Edinburgh, 2000.
- P. Hager, Eigenfrequency Analysis, FE-Adaptivity and a Nonlinear Eigenvalue Problem. PhD thesis, Chalmers University of Technology, G¨oteborg, 2001.
- V. Hari, On sharp quadratic convergence bounds for the serial Jacobi methods, Numer. Math. 60, pages 375-406,1991.
- P. Henrici, On the speed of convergence of cyclic and quasicyclic Jacobi methods for computing eigenvalues of Hermitian matrices, J. Soc. industr. appl. Math. 6, pages 144-162, 1958.
- P. Henrici, Bounds for iterates, inverses, spectral variation and fields of values of nonnormal matrices. Numerische Mathematik, 4:24-39,1962.
- N. J. Higham, Accuracy and Stability of Numerical Algorithms. SLAM, Philadelphia, 1996.
- D. J. Higham and N. J. Higham, Structured backward error and condition of generalized eigenvalue problems. SIAM Journal on Matrix Analysis and Applications, 20:493-512,1998.
- Paul Horst, A method of factor analysis by means of which all coordinates of the factor matrix are given simultaneousl, Psychometrika 2, 225- 236, 1937.
- H. Hotelling, Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24:417'-441 and 498-520,1933.
- A. S. Householder and F. L. Bauer, On certain methods for expanding the characteristic polynomial. Numerische Mathematik, 1:29-37,1958.
- A. S. Householder, A class of methods for inverting matrices. J. Soc. Ind. Appl. Math. 6, 189-195, 1958a.
- A. S. Housholder, Unitary Triangularization of a Nonsymmetric Matrix, J. Asoc. Comp., Mach. 5, pages 339-342, 1958b.
- Alston S. Housesholder, The Theory of Matrices in Numerical Analysis 178-199, Dover Publications, Inc., New York, 1964. Originally published by Blaisdell, New York.
- C. F. Ipsen, Computing an eigenvector with inverse iteration. SIAM Review, 39:254-91,1997.

- C.G.J. Jacobi, Ueber ein leichtes Verfahren, die in der Theorie der Sscularstorungen vorkommended Gleichungen numerisch aufzulosen, Journal for die reine and angewandte, Mathematik, pages 51-94, 1846.
- N. K. Jain and K. Singhal, On Kublanovskaya's approach to the solution of the generalized latent value problem for functional λ–matrices. SIAM J. Numer. Anal., 20:1062–1070, 1983.
- N. K. Jain, K. Singhal, and K. Huseyin, On roots of functional lambda matrices. Comput. Meth. Appl. Mech. Eng., 40:277–292, 1983.
- A. Jennings, Eigenvalue methods and the analysis of structural vibrations. In I. S. Duff, editor, Sparse Matrices and their Uses, pages 109-138. Academic Press, New York, 1981.
- C. Jordan, Traite des Substitutions et des Equations Algebriques. Paris, 1870.
- C. Jordan, Memoire sur les formes bilineaires, Journal de Mathematiques Pures et Appliquees, Deuxieme Serie, 19:35-54,1874.
- J. M. van Kats and H. A. van der Vorst, Numerical Results of the Paige-Style Lanczos Method for the Computation of Extreme Eigenvalues of Large Sparse Matrices, Tech. Report 3. Utrecht, Academic Computer Center, Netherlands, 1976.
- S. Kaniel, Estimates for some computational techniques in linear algebra, Mathematics of Computations, Vol 20, pages 369-378, 1966.
- B. Kagstrom and A. Ruhe, Algorithm 560: JNF, an algorithm for numerical computation of the Jordan normal form of a complex matrix. Transactions on Mathematical Software, 6:437-443 and 398-419,1980.
- B. Kagstrom and P Wiberg, Extracting partial canonical structure for large scale eigenvalue problems. Numerical Algorithms, 24:195-237, 1999.
- H.P.M. van Kempen, On quadratic convergence of the special cyclic Jacobi method 19-22, Numer. Math. 9, 1966.
- Kokiopoulou and C. Bekas and E. Gallopoulos, Computing smallest singular triplets with implicitly restarted Lanczos bidiagonalization, 2004.
- Vera N. Kublanovskaya, On some algorithms for the solution of the complete eigenvalue problem, USSR Computational Mathematics and Mathematical Physics, Vol. 1, No. 3, pages 637–657 1963(received Feb 1961). Also published in: Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, Vol.1, No. 4, pages 555–570, 1961.
- D. LaBudde, The reduction of an arbitrary real square matrix to tridiagonal form using similarity transformations. Mathematics of Computation (American Mathematical Society) 17 (84): 433–437, 1963.
- C Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, Journal of Research of the National Bureau of Standards, Vol 45, pages255-282, 1950.
- Langville, Amy N. Meyer, Carl D, Google's PageRank and Beyond: The Science of Search Engine Rankings, Princeton University Press 2006.
- R. B. Lehoucq, D. C. Sorensen, and C. Yang, ARPACK USERS GUIDE: Solution of Large Scale Eigenvalue Problems by Implicitly Restarted Arnoldi Methods. SIAM, Philadelphia, PA, 1998.
- C. Van Loan, A general matrix eigenvalue problem, SIAM J. Numerical Analysis, 12, 819-834, 1975.
- Mark Lotkin, Characteristic values of arbitrary matrices, Quart. Appl. Math. 14, pages 267-275, 1956.

- R. von Mises, H. Pollaczek-Geiringer, Praktische Verfahren der Gleichungsauflosung, Z. Angew. Math. Mech.9, 58-77, 152-164. 1929.
- J. J. Modi, Parallel Algorithms and Matrix Computation, Oxford,Clarendon Press,1988.
- C. B. Moler and G. W. Stewart, An algorithm for generalized matrix eigenvalue problems. SIAM Journal on Numerical Analysis, 10:241-256,1973.
- V. Mehrmann, The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution. Number 163 in Lecture Notes in Control and Information Sciences. Springer Verlag, Heidelberg, July 1991.
- V. Mehrmann and D. Watkins, Structure-preserving methods for computing eigenpairs of large sparse skew-hamiltonian/hamiltonian pencils. SIAM J. Sci. Comput., 22:1905–1925, 2001.
- V. Mehrmann and D. Watkins, Polynomial eigenvalue problems with hamiltonian structure. Electr. Trans. Num. Anal., 13:106–113, 2002.
- L. Miintz, Solution direct de 1'equation seculaire et de quelques problemes analogues transcendents. Comptes Rendus de I'Academie des Sciences, Paris, 156:43-6,1913.
- D. D. Morrison, Remarks on the Unitary Triangularization of a Nonsymmetric Matrix, Journal of the ACM 7 (2): 185–186, 1960.
- C. H. Muntz, Solution direct de l'equation seculaire et de quelques problemes analogues transcendents, C.R. Acad. Sci. Paris Vol 156, pages 43-46, 1913a.
- C. H. Muntz, Sur la solution des equations seculaires et des equations integrales, C.R. Acad. Sci. Paris Vol 156, pages 860-862, 1913b.
- R. Natarajan, An Arnoldi-based iterative scheme for nonsymmetric matrix pencils arising in finite element stability problems. Journal of Computational Physics, 100:128-142, 1992.
- R. Natarajan and A. Acrivos, The instability of the steady flow past spheres and disks. Technical Report RC 18235, IBM Res. div., T. J. Watson Res. ctr, Yorktown Heights, 1992.
- J. von Neumann and H. H. Goldstine, Numerical inverting of matrices of high order, Bull. Amer. math. Soc. 53, pages 1021-1099, 1947.
- J. Olsen, P. Jorgensen, and J. Simons, Passing the one-billion limit in full configuration-interaction (fci) calculations. Chemical Physics Letters, 169:463–472, 1990.
- J. M. Ortega, An error analysis of Householder's method for the symmetric eigenvalue problem, Numerische Math. 5, pages 211-225, 1963.
- E. Osborne, On pre-conditioning of matrices. Journalof the ACM, 7:338-345,1960.
- A.M. Ostrowski, On the convergence of the Rayleigh quotient iteration for the computation of characteristic roots and vectors I-VI, Arch. Rational. Mesh. Anal. I 1:233-241, II 2:423-428, III 3:325-340, IV 3:341-347, V 3:472-481, VI 4:153-165, 1958-1959.
- Evgueni E. Ovtchinnikov and Leonidas S. Xanthis, Successive eigenvalue relaxation: a new method for the generalized eigenvalue problem and convergence estimates, Centre for Techno-Mathematics and Scientific Computing Laboratory, University of Westminster, London 2000.
- C.C Paige, The Computation of Eigenvalues and Eigenvectors of Very Large Sparse Matrices, Ph.D. thesis, Univ. of London, London, 1971.
- C.C Paige, The Computational variants of the Lanczos method for the eigenproblem, J. Inst. Math. Appl, Vol 10, pages 373-381, 1972.

- C.C Paige, Error analysis of the Lanczos algorithm for tridiagonalizing a symmetric matrix,. J. Inst. Math. Appl. 18:341-349, 1976.
- P. C. Papanastasiou, Numerical analysis of localization phenomena with application to deep boreholes. PhD thesis, University of Minnesota, Dept. Civil and Mineral Engineering, Minneapolis, MN, 1990.
- B. N. Parlett, The origin and development of methods of the LR type. SIAM Review, 6:275-295,1964.
- B.N. Parlett , Canonical Decomposition of Hessenberg Matrices, Math. Comp. 21, 223-27, 1967.
- B. N. Parlett and C. Reinsch, Balancing a matrix for calculation of eigenvalues and eigenvectors. Numerische Mathematik, 13:293-304,1969.
- B. N. Parlett, Analysis of algorithms for reflections in bisectors. SIAM Review, 13:197-208,1971.
- B. N. Parlett and D. S. Scott, The Lanczos algorithm with selective orthogonalization, Math. Comp. Vol 33, pages217-238, 1979.
- B. N. Parlett and J. K. Reid, Tracking the progress of the Lanczos algorithm for large symmetric eigenproblems, IMA J. Num. Anal, Vol 1, pages 135-155, 1981.
- B. N. Parlett, The software scene in the extraction of eigenvalues from sparse matrices, SIAM J. of Sci. Stat. Comput., Vol 5 Edition 3, pages, 590- 604, 1984.
- B. N. Parlett, D. R. Taylor, and Z. S. Liu. A look-ahead Lanczos algorithm for nonsymmetric matrices. Mathematics of Computation, 44:105-124, 1985.
- B. N. Parlett and Y. Saad, Complex shift and invert strategies for real matrices. Linear Algebra and its Applications, 88/89:575-595, 1987.
- B. N. Parlett, The Symmetric Eigenvalue Problem. Prentice-Hall, Englewood Cliffs, NJ, 1980. Reissued with revisions by SIAM, Philadelphia, 1998.
- S. Petiton, Parallel subspace method for non-Hermitian eigenproblems on the connection machine (CM-2), Technical Report YALEU/DCS/RR-859, Yale University, Computer Science dept., New Haven, CT, 1991.
- Oskar Perron, Zur Theorie der Matrizen, Math. Ann. 64, 248- 263, 1908.
- D. A. Pope and C. Tompkins, Maximizing functions of rotations experiments concerning speed of diagonalization of symmetric matrices using Jacobi's method, J. Ass. corp. Mach. 4, pages 459-466, 1957.
- T. D. Roopamala and S. K. Katti, Eigenvalues of Structural Matrices Via Gerschgorin Theorem, International Journal of Computer Science and Network Security, Vol.10 No.1, January 2010.
- J. B. Rosser, C. Lanczos, M. R. Hestenes And W. Karush, The separation of close eigenvalues of a real symmetric matrix, J. Res. not. Bur. Stand Vol 47, pages 291-297, 1951.
- A. Ruhe, Perturbation bounds for means of eigenvalues and invariant subspaces. BIT, 10:343-354,1970.
- A. Ruhe, Computing nonlinear eigenvalues with spectral transformation Arnoldi. ZAMM, 76:S2:17 – 20, 1996a.
- A. Ruhe, Rational Krylov for nonlinear eigenvalues. Talk presented at the Householder
  XIII Conference at Pontresina, Switzerland, 1996b.
- A. Ruhe, Rational Krylov, a practical algorithm for large sparse nonsymmetric matrix pencils.
  SIAM J. Sci. Comp., 19(5):1535–1551, 1998.
-

- A. Ruhe, A rational Krylov algorithm for nonlinear matrix eigenvalue problems. Zapiski Nauchnyh Seminarov POMI, 268:176 – 180, 2000.
- A. Ruhe, Rational Krylov for large nonlinear eigenproblems. Technical report, Royal Institute
  of Technology, Stockholm, Sweden, 2004. To appear in Proceedings of PARA'04, Lyngby, Denmark, 2004.
- H. Rutishauser, Anwendungen Des Quotienten-Differenzen-Algorithmus. Zeitschrift filr angewandte Mathematik und Physik, 5:496-508,1954a.
- H. Rutishauser, Der Quotienten-Differenzen-Algorithmus. Zeitschrift fiir angewandte Mathematik und Physik, 5:233-251,1954b.
- H. Rutishauser, Une methode pour la determination des valeurs propres d'une matrice. Comptes Rendus de VAcademic des Sciences, Paris, 240:34-36,1955.
- H. Rutishauser, Solution of eigenvalue problems with the LR-transformation. National Bureau of Standards Applied Mathematics Series, 49:47-81,1958.
- Y. Saad, On the rates of convergence of the Lanczos and the block Lanczos methods, SIAM J. Numer. Anal, Vol 17, pages 687-706, 1980.
- Y. Saad, Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices, Linear Algebra and its Applications, Vol 34, pages 269-295, 1980.
- Y. Saad, M.H. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Statist. Comput. 7 856-869, 1986.
- Y. Saad, Numerical solution of large Lyapunov equations. In M. A. Kaashoek, J. H. van Schuppen, and A. C. Ran, editors, Signal Processing, Scattering, Operator Theory, and Numerical Methods. Proceedings of the international symposium MTNS-89, vol III, pages 503-511, Boston, 1990a. Birkhauser.
- Y. Saad, An overview of Krylov subspace methods with applications to control problems. In M. A. Kaashoek, J. H. van Schuppen, and A. C. Ran, editors, Signal Processing, Scattering, Operator Theory, and Numerical Methods. Proceedings of the international symposium MTNS-89, vol III, pages 401-410, Boston, 1990b. Birkhauser.
- Y. Saad, Numerical Methods for Large Eigenvalue Problems, Manchester University Press, Manchester, UK, 1992.
- Schonhage, Zur Konvergenz des Jacobi-Verfahrens, Numerische Math. 3, 374-380, 1961.
- S. Scott, Implementing Lanczos-like algorithms on Hypercube architectures. Computer Physics Communications, 53:271-282, 1989.
- J. Schur, Uber die charakteristischen Wiirzeln einer linearen Substitution miteiner Anwendung auf die Theorie der Integralgleichungen. MathematischeAnnalen, 66:448-510,1909.
- Slapricar, Accurate Symmetric eigenreduction by a Jacobi Method, Ph.D. thesis, Hagen, Germany, Ferruniversitat, 1992.
- G. L. G. Sleijpen and H. A. van der Vorst, A Jacobi-Davidson Iteration Method for Linear Eigenvalue Problems, SIAM Journal on Matrix Analysis and Applications, 17 (1996) pp. 401-425. Reprinted in SIAM Review, 42 (2000), pp. 267-293.
- G. W. Stewart, On the sensitivity of the eigenvalue problem $Ax = \lambda Bx$, SIAM Journal on Numerical Analysis, 4:669-686,1972.
- G. W. Stewart, Introduction to Matrix Computations. Academic Press, New York, 1973.

- G. W. Stewart, Gerschgorin theory for the generalized eigenvalue problem Ax = λBx. Mathematics of Computation, 29:600-606,1975.
- G. W. Stewart and J. G. Sun, Matrix Perturbation Theory. Academic Press, New York, 1990.
- G. W. Stewart, Matrix Algorithms, Vol I- Basic Decompositions, SIAM, Philadelphia, 1998.
- G. W. Stewart, Matrix Algorithms, Vol II- Eigensystems, SIAM, Philadelphia, 2001.
- D.C. Sorensen, Implicit application of polynomial filters in a k-step Arnoldi method, SIAM J. Matrix Anal. Appl.13 (1) 357-385, 1992.
- D. C. Sorensen and C. Yang, Accelerating the Lanczos algorithm via polynomial spectral transformations. Technical report TR 97-29, Department of Computational & Applied Mathematics, Rice University, Houston, TX, 1997.
- D. C. Sorensen and C. Yang, A truncated RQ-iteration for large scale eigenvalue calculations. SIAM J. Matrix Anal. Appl., 19(4):1045–1073, 1998.
- O. Taussky, A recurring theorem on determinants. Amer. math. Mon. 56, pages672-676, 1949.
- H. W. Turnbull and A. C. Aitken, An Introduction to the Theory of Canonical Matrices. Blackie and Son, London, 1932.
- M. Turing, Rounding-off errors in matrix processes, Quart. J. Mech. 1, pages 287-308, 1948.
- K. Veselic and V. Hari, A note on a one-sided Jacobi algorithm, Numer. Math, 56 pages 627-633, 1990.
- V. V. Voyevodin, A method for the solution of the complete eigenvalue problem, Zh. vych. Mat. 2, pages 15-24, 1962.
- R. C. Ward, The combination shift QZ algorithm. SIAM Journal on Numerical Analysis, 12:835-853,1975.
- R. C. Ward, Balancing the generalized eigenvalue problem. SIAM Journal on Scientific and Statistical Computing, 2:141-152,1981.
- D. S. Watkins and L. Eisner, Convergence of algorithms of decomposition type for the eigenvalue problem. Linear Algebra and Its Applications, 143:19-47, 1991.
- D. S. Watkins, Forward stability and transmission of shifts in the QR algorithm. SIAM Journal on Matrix Analysis and Applications, 16:469—487,1995.
- D. S. Watkins, On the transmission of shifts and shift blurring in the QR algorithm. Linear Algebra and Its Applications, 241-243:877-896,1996.
- D. S. Watkins, Performance of the QZ algorithm in the presence of infinite eigenvalues. SIAM Journal on Matrix Analysis and Applications, 22:364-375, 2000.
- K. Weierstrass, Zur Theorie der bilinearen und quadratischen Formen. Monatshefte Akadamie Wissenshaften Berlin, pages 310-338, 1868.
- R. R. Whitehead, A numerical approach to nuclear shell-model calculations, Nuclear Phys. A1 82, pages 290-300, 1972.
- J. H. Wilkinson, The calculation of the eigenvectors of codiagonal matrices, Computer J. 1, 90-96, 1958.
- J. H. Wilkinson, Householder's method for the solution of the algebraic eigenvalue problem. Computer J. 3, pages 23-27, 1960a.
- J. H. Wilkinson, Error analysis of floating-point computation, Numerische Math. 2, pages 319-340, 1960b.
- J. H. Wilkinson, Note on the quadratic convergence of the cyclic Jacobi process, Numerische Math. 4, pages 296-300, 1962.
- J.H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, 1965a.

- J. H. Wilkinson, Convergence of the LR, QR, and related algorithms. Computer Journal, 8:77-84,1965b.
- J. H. Wilkinson, Global convergence of tridiagonal QR algorithm with origin shifts. Linear Algebra and Its Applications, 1:409-420,1968.
- J. H. Wilkinson and C. Reinsch, Handbook for Automatic Computation. Vol. H Linear Algebra. Springer-Verlag, New York, 1971.
- J. H. Wilkinson, Kronecker's canonical form and the QZ algorithm, Linear Algebra and Its Applications, 28:285-303,1979.
- H. Wielandt, Beitrage zur mathematischen Behandlung komplexer Eigenwertprobleme V: Bestimmung hoherer Eigenwerte durch gebrochene Iteration. Bericht B 44/J/37, Aerodynamische Versuchsanstalt Gottingen, Germany, 1944.
- Wikipedia-1 http://en.wikipedia.org/wiki/Abel%E2%80%93Ruffini_theorem.
- Wikipedia-2 http://en.wikipedia.org/wiki/Schur_decomposition.
- Wikipedia-3 http://en.wikipedia.org/wiki/Householder_transformation.
- Wikipedia-4 http://en.wikipedia.org/wiki/QR_algorithm.
- Wikipedia-5 http://en.wikipedia.org/wiki/Lanczos_algorithm.
- Wikipedia-6 http://en.wikipedia.org/wiki/Arnoldi_iteration.
- Kesheng Wu and Horst Simon, Thick-Restart Lanczos Method for Large Symmetric Eigenvalue Problems, 2000.
- C. Yang, Convergence analysis of an inexact truncated RQ-iteration. Electronic Transaction on Numerical Analysis, 7(4):40–55, 1998.
- Chao Yang, Solving Large-scale Eigenvalue Problems in SciDAC Applications, Computational Research Division, Lawrence Berkeley National Lab, Berkeley, CA, 94597, 2005.
- W.H. Yang, A method for eigenvalues of sparse -matrices. Internat. J. Numer. Meth. Eng.,
  19:943 – 948, 1983.
- David M. Young and Robert Todd Gregory, A survey of numerical mathematics, Volume 2, Dover Publications, Mineola, N.Y, 1988.
- Yunkai Zhou, Studies on Jacobi-Davidson, Rayleigh Quotient Iteration, Inverse Iteration Generalized Davidson and Newton Updates, Dept. of Computer Science & Engineering, University of Minnesota, Minneapolis, MN 55455, USA, 2005.
- R. Zurm¨uhl and S. Falk, Matrizen und ihre Anwendungen, Bd. I. Springer Verlag, Berlin, fifth edition, 1984.
- R. Zurm¨uhl and S. Falk, Matrizen und ihre Anwendungen, Bd. II. Springer Verlag, Berlin, fifth edition, 1986.