CONF-810251--1

MASTER

ALGORITHMS FOR SPARSE, SYMMETRIC, DEFINITE QUADRATIC λ-MATRIX EIGENPROBLEMS*

David S. Scott and Robert C. Ward

Mathematics and Statistics Research Department
Computer Sciences Division
Union Carbide Corporation - Nuclear Division
Oak Ridge, Tennessee 37830

*Research sponsored by the Applied Mathematical Sciences Research Program, Office of Energy Research, U.S. Department of Energy under contract W-7405-eng-26 with the Union Carbide Corporation.

ALGORITHMS FOR SPARSE, SYMMETRIC, DEFINITE QUADRATIC λ-MATRIX EIGENPROBLEMS

David S. Scott and Robert C. Ward
Computer Sciences Division
Union Carbide Corporation - Nuclear Division
Oak Ridge, Tennessee 37830

ABSTRACT. Methods are presented for computing eigenpairs of the quadratic λ-matrix, $M\lambda^2 + C\lambda + K$, where M, C, and K are large and sparse, and have special symmetry-type properties. These properties are sufficient to insure that all the eigenvalues are real and that theory analogous to the standard symmetric eigenproblem exists. The methods employ some standard techniques such as partial tri-diagonalization via the Lanczos Method and subsequent eigenpair calculation, shift-and- invert strategy and subspace iteration. The methods also employ some new techniques such as Rayleigh-Ritz quadratic roots and the inertia of symmetric, definite, quadratic λ-matrices.

1. INTRODUCTION. Quadratic λ-matrix problems consist of determining scalars λ, called eigenvalues, and corresponding n x 1 nonzero vectors x, called eigenvectors, such that the equation

$$(M\lambda^2 + C\lambda + K)x = 0 \qquad (1)$$

is satisfied, where M, C, and K are given nxn matrices. In addition, we assume that M, C, and K are symmetric or Hermitian, M is definite (either positive or negative definite), and the eigenvalues of (1) are real and can be divided into two disjoint sets $\mathcal{P}$ and $\mathcal{S}$ with the following properties:

P1) If $\lambda_i \in \mathcal{P}$ and $\lambda_j \in \mathcal{S}$, then $\lambda_i > \lambda_j$.

P2) If $\lambda_i \in \mathcal{P}$ ($\mathcal{S}$) and x_i is its associated eigenvector, then λ_i is the larger (smaller) root of the quadratic equation

$$(x_i^* M x_i)\,\lambda^2 + (x_i^* C x_i)\,\lambda + (x_i^* K x_i) = 0.$$

The eigenvalues in $\mathcal{P}$ will be called primary eigenvalues, and those in $\mathcal{S}$ will be called secondary. Their eigenvectors will be referenced similarly.

Problems of this nature occur in several application areas; we will briefly discuss two of them. Lancaster [2] states that the determination of sinusoidal solutions to the equations of motion for vibrating systems which are heavily damped results in such a quadratic λ-matrix problem. In these overdamped systems M, C, and K are symmetric, M and C are positive definite, K is non-negative definite, and the overdamping condition

$$(y^*Cy)^2 - 4(y^*My)(y^*Ky) > 0$$

is satisfied for all vectors $y \neq 0$. Proof that the eigenvalues for overdamped systems are all real and obey properties P1 and P2 above can be found in Lancaster [2]. Problem (1) also arises in the dynamic analysis of rotating structures where the gyroscopic effects cannot be ignored. (See Wildheim [8] and Lancaster [2].) In gyroscopic systems M, C, and K are symmetric (Hermitian), M is negative definite, and K is positive definite. One can determine (Scott and Ward [7]) that all the eigenvalues are real, that P and S are the positive and negative eigenvalues, respectively, and that properties P1 and P2 are satisfied. In both overdamped and gyroscopic systems, the M matrix is usually called the mass matrix and K the stiffness matrix. Thus, we have chosen the notation given in (1) rather than the more standard mathematical notation using A, B, and C for the matrices.

In this paper we present various methods for computing eigenpairs of these quadratic λ-matrices when M, C, and K are also large and sparse. Due to the simplicity of the properties of gyroscopic systems, our model problem for presentation of the methods will be from this application area. That is, we will discuss algorithms for computing eigenpairs of equation (1) where M, C, and K are large, sparse, and symmetric, M is negative definite, and K is positive definite.

In Section 2 we discuss the approach of transforming the quadratic problem into a linear one. Some methods based on the factorization of a nxn matrix are presented in Section 3 with methods not requiring any factorization presented in Section 4. We close the paper by summarizing our results.

2. LINEARIZATION. It may be immediately verified that the eigenpair (λ, x) satisfies the quadratic problem (1) if and only if it also satisfies the 2n x 2n linear problem

$$\left(\begin{bmatrix} 0 & K \\ K & C \end{bmatrix} - \lambda \begin{bmatrix} K & 0 \\ 0 & -M \end{bmatrix} \right) \begin{bmatrix} x \\ \lambda x \end{bmatrix} = 0 , \qquad (2)$$

which we denote as $(A-\lambda B)z = 0$. By the hypotheses on M, C, and K, A and B are symmetric and B is positive definite. Thus from well known linear theory, there are 2n real eigenvalues. Applying the Cauchy interlace theorem to the n x n zero block of A leads to the conclusion that exactly n of the eigenvalues are positive and n are negative. Finally, the eigenvectors of the linear problem are B orthogonal so that if (λ_1, x_1) and (λ_2, x_2) are different eigenpairs, then

$$x_1^* K x_2 - \lambda_1 \lambda_2 x_1^* M x_2 = 0. \qquad (3)$$

Unfortunately, equation (3) involves both λ_1 and λ_2 and does not lead to a useful deflation technique.

Sparse linear eigenvalue problems have been studied in some detail and good solution techniques exist. However, a general linear solver may not be the best choice for solving a quadratic problem in that the linear problem has dimension 2n even though the original problem has dimension n and no advantage will be taken of the special structure of A and B. Also, $A-\sigma B$ is not banded even if M, C, and K are so that factoring $A-\sigma B$, which is an integral part of most linear solvers, will require special care to preserve sparsity.

For these reasons we will investigate solution techniques which take advantage of the underlying quadratic problem.

3. FACTORIZATION TECHNIQUES. In this section we show that the linear problem (2) can be solved using well-known techniques by factoring an nxn matrix only. The Lanczos algorithm and subspace iteration appear to require the factorization of the 2n x 2n matrix $A-\sigma B$. However what is actually needed is the ability to multiply vectors by $(A-\sigma B)^{-1}B$. The special structure of the A and B matrices allows this operator to be realized by factoring only the nxn matrix $W(\sigma) = M\sigma^2 + C\sigma + K$.

Theorem 1. Let A and B be as in equation (2) and let $W(\sigma) = M\sigma^2 + C\sigma + K$. Then

1) The number of negative eigenvalues of W equals the number of eigenvalues of $A-\lambda B$ between σ and 0.

2) $$(A-\sigma B)^{-1}B \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} W(\sigma)^{-1} (-Cx - \sigma Mx - My) \\ W(\sigma)^{-1} (Kx - \sigma My) \end{bmatrix}$$

The proof is given in Scott [5]. Once the operator $(A-\sigma B)^{-1}B$ has been realized then it is straight forward to implement subspace iteration or the Lanczos algorithm (as described in Scott [4]) to find the eigenvalues of $A-\lambda B$ near σ. The number of negative eigenvalues of W can be easily determined as a byproduct of the factorization and so the index of the computed eigenvalues can be found.

If many eigenvalues are desired then a sequence of shifts σ can be used. The eigenvalue count then gives the number of eigenvalues between two consecutive shifts so that no eigenvalue can be knowingly missed.

4. NONFACTORIZATION TECHNIQUES. In this section we assume that the factorization of M, C, K, or any linear combination of them is either impossible or undesirable. Thus, we are basically limited to algorithms similar to the Lanczos Rayleigh Quotient algorithm presented by Scott [6] for the linear pencil eigenproblem which uses only matrix-vector multiplications.

We have developed an algorithm based on techniques for determining the "best" approximation to an eigenvalue given an approximate eigenvector and the "best" approximation to an eigenvector given an approximate eigenvalue. The algorithm alternates between these approximations until convergence, as the following outline illustrates:

I. Set the vector x_0 to random numbers.

II. For i = 1, 2, ... until convergence, do a and b.

a. Determine "best" σ_i from x_{i-1}.

b. Determine "best" x_i from σ_i.

Step II.a uses a generalization of the Rayleigh quotient different from that of Lancaster's [2] and specifically designed for the quadratic problem. Given any nonzero vector x, potential eigenvectors of the linear pencil (A,B) given by (2) would be linear combinations of the vectors $[x^*, 0]^*$ and $[0, x^*]^*$. Using the Rayleigh-Ritz procedure, the "best" approximations to eigenvectors in this space and corresponding eigenvalues can be determined. Best in this context means minimizing the Frobenius norm of the 2 x 2 scaled residual matrix (see Parlett [3]). The characteristic equation of the reduced linear pencil in the Rayleigh-Ritz procedure is equivalent to the quadratic equation

$$(x^*Mx)\,\theta^2 + (x^*Cx)\,\theta + (x^*Kx) = 0. \qquad (4)$$

Thus, the approximations to two eigenvalues of the quadratic λ-matrix are given by its roots, $\theta^+(x)$ and $\theta^-(x)$, which can be easily determined by the quadratic formula. If we are trying to converge to a positive (primary) eigenvalue, then the larger root $\theta^+(x)$ is chosen for σ_i; conversely, the smaller root $\theta^-(x)$ is chosen when trying to converge to a negative (secondary) eigenvalue. The roots of (4) are identical to the primary and secondary functionals discussed by Duffin [1]. However, Duffin does not present a theoretical basis for how and why these roots along with x most closely approximates an eigenpair of the quadratic λ-matrix. A more thorough discussion of Rayleigh quotient generalizations can be found in Scott and Ward [7].

Step II.b is based on the observation that if σ is an eigenvalue of the quadratic λ-matrix with x as its eigenvector, the matrix $W(\sigma)$ defined in Theorem 1 has the eigenpair (0, x). Theorem 1 relates the eigenvalues of the symmetric matrix $W(\sigma)$ to the primary and secondary eigenvalues of the quadratic λ-matrix. Thus, to which eigenvalue we are converging can be controlled by the selection of the appropriate eigenvector of $W(\sigma)$ to be used in Step II.b. For example, the following algorithm is used to converge to the m smallest positive eigenvalues:

I. Set the vector x_0 to random numbers.

II. For $k = 1, 2, \ldots m$, do 1 and 2.

   1. For $i = 1, 2, \ldots$ until convergence, do a and b.

      a. Set $\sigma_i = \theta^+(x_{i-1})$.

      b. Set $x_i = y_k$, where (μ_j, y_j) are eigenpairs of $W(\sigma_i)$ with $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_n$ and y_j unit-length.

   2. Set x_0 to the y_{k+1} computed in step 1.b above.

From Scott and Ward [7], we know that the sequence $\{\sigma_i\}$ for $k = 1$ converges monotonically downward to the smallest positive eigenvalue, and the convergence is asymptotically quadratic. Also, the algorithm is expected to quadratically converge to the other m-1 eigenvalues, but convergence is *not guaranteed.*

A minor modification can be made to the algorithm to guarantee quadratic convergence to interior primary or secondary eigenvalues. This modification requires the solution to a 2k x 2k dense linear pencil eigenproblem in step II.1.a. and the computation of k eigenvectors in step II.1.b. The following algorithm is guaranteed to quadratically convergence to the m smallest positive eigenvalues:

I. Set the vector y_1 to random numbers.

II. For $k = 1, 2, \ldots, m$, do 1 and 2.

   1. Set the r^{th} column of the nxk matrix X to y_r from step I if k = 1 or from step II.2.b otherwise.

   2. For $i = 1, 2, \ldots$ until convergence, do a and b.

      a. Set $\sigma_i = \theta_k$ where $\theta_{-k} \leq \theta_{-k+1} \leq \ldots \leq \theta_{-1} < 0 < \theta_1 \leq \ldots \leq \theta_k$ are the eigenvalues of

$$\begin{bmatrix} X_{i-1} & 0 \\ 0 & X_{i-1} \end{bmatrix}^* \left(\begin{bmatrix} 0 & K \\ K & C \end{bmatrix} - \theta \begin{bmatrix} K & 0 \\ 0 & -M \end{bmatrix} \right) \begin{bmatrix} X_{i-1} & 0 \\ 0 & X_{i-1} \end{bmatrix}.$$

b. Set the r^{th} column of X_i to y_r, where (μ_j, y_j) are the eigenpairs of $W(\sigma_i)$ with $\mu_1 < \mu_2 < \ldots < \mu_n$ and y_j are unit-length.

Similar algorithms can be developed for computing the m largest positive eigenvalues and the m largest and smallest negative eigenvalues.

5. CONCLUSIONS. In this paper we have presented several techniques for solving symmetric, definite, quadratic λ-matrix problems. These techniques are more efficientm, in general, than applying linear techniques to the equivalent $2n \times 2n$ linear problem. The convergence rates of the methods based on factoring $W(\sigma)$ are superior to the convergence rates of the nonfactorization methods presented in Section 4, and so the factorization methods should always be used if the factorization is possible. If the nonfactorization methods must be used, then it is still possible to use preconditioning techniques as in Scott [6] to improve the convergence, if desired. Portable software implementing these algorithms should be available in the near future.

REFERENCES

1. R. J. Duffin, "A Minimax Theory for Overdamped Networks," J. Rational Mech. Anal. 4 (1955), 221-233.

2. Peter Lancaster, Lambda-Matrices and Vibrating Systems, Pergamon Press, New York, 1966.

3. Beresford N. Parlett, The Symmetric Eigenvalue Problem, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1980.

4. D. S. Scott, The Advantages of Inverted Operators in Rayleigh-Ritz Approximations, Technical Report ORNL/CSD-68, Union Carbide Corporation, Nuclear Division, Oak Ridge, Tennessee, 1980.

5. D. S. Scott, Solving Sparse Quadratic λ-Matrix Problems, Technical Report ORNL/CSD-69, Union Carbide Corporation, Nuclear Division, Oak Ridge, Tennessee, 1980.

6. D. S. Scott, "Solving Sparse Symmetric Generalized Eigenvalue Problems Without Factorization," SIAM J. Numer. Anal. 18 (1981), 102-110.

7. D. S. Scott and R. C. Ward, Solving Quadratic λ-Matrix Problems Without Factorization, Technical Report ORNL/CSD-76, Union Carbide Corporation, Nuclear Division, Oak Ridge, Tennessee, 1981.

8. J. Wildheim, "Vibrations of Rotating Circumferentially Periodic Structures," Q. J. Mech. Appl. Math., to appear.