SIAM J. NUMER. ANAL.
Vol. 25, No. 6, December 1988

CRITERIA FOR COMBINING INVERSE AND RAYLEIGH QUOTIENT ITERATION*

DANIEL B. SZYLD†

Abstract. A method is presented to find selected eigenvalue-eigenvector pairs of the generalized problem $Ax = \lambda Bx$. Here A and B are real symmetric matrices and B is positive definite. It is further assumed that the matrices are large and sparse and that factorization of either of them is impractical. Our method finds an eigenvalue in a given interval $J = (\gamma - \eta, \gamma + \eta)$ or determines that J is free of eigenvalues while computing an approximation to the eigenvalue closest to this interval. The corresponding eigenvector is also computed. The method consists of inverse and Rayleigh quotient iteration steps. The convergence is studied and it is shown how an inclusion theorem gives one of the criteria for switching from inverse to Rayleigh quotient iteration. The existence of an eigenvalue in the desired interval is guaranteed when this criterion is fulfilled. Some numerical experiments are reported which suggest that the method is optimal in the sense that when the mesh is refined the computational effort grows only linearly with the number of mesh points.

Key words. eigenvalues, eigenvectors, sparse matrices

AMS(MOS) subject classification. 65

1. Introduction.

The generalized eigenvalue problem

$$Ax = \lambda Bx, \tag{1}$$

where A and B are $n \times n$ real symmetric matrices and B is positive definite, appears in many applications, most notably in structural analysis. Eigenvalue calculations are important to determine the natural or resonance frequencies of structures, thus insuring that they are not close to frequencies likely to affect them, i.e., the vibration produced by the propellers, in the case of a ship; the natural wave frequencies that affect an offshore oil platform; or the earthquake band for a building (see [2], [1, p. 146], [14, p. xii]).

It is well known that the eigenvalues of the problem considered are real. We are primarily concerned with finding an eigenvalue in a given interval $J = (\gamma - \eta, \gamma + \eta)$, not necessarily close to the end of the spectrum, and the corresponding eigenvector or to discover that the interval in question is free of eigenvalues. We are particularly interested in the case where the matrices A and B are prohibitively expensive to factor into triangular forms for which few other alternatives exist.

We propose a combination of inverse and Rayleigh quotient iterations of the following form:

$$(A - \mu B)y_{s+1} = Bx_s, \tag{2}$$

$$\omega_{s+1} = (y_{s+1}^T B y_{s+1})^{-1/2}, \tag{3}$$

$$x_{s+1} = \omega_{s+1} y_{s+1} \tag{4}$$

where $\mu = \gamma$ in the case of inverse iteration and $\mu = \mu_s = x_s^T A x_s$ for the Rayleigh quotient case. Inverse iteration converges linearly to the eigenvalue closest to γ, while

* Received by the editors September 15, 1986; accepted for publication (in revised form) October 5, 1987. This work is based in part on the author's Ph.D. thesis, which was submitted to the Department of Mathematics at New York University and was supported in part by the Department of Energy under contract DE-AC02-76-ER03077-V at the Courant Mathematics and Computing Laboratory.

† Computer Science Department, Duke University, Durham, North Carolina 27706-2591.

Rayleigh quotient iteration exhibits cubic convergence. However, the latter method has the drawback that it can converge to different eigenvalues depending on the initial vector x_0 and often to a λ_i far from γ (see, e.g., [14, p. 71]). The choice of $\mu = \gamma$ or $\mu = \mu_s$ is based on criteria that guarantee superlinear convergence of the algorithm to an eigenvalue in J. These criteria are developed below (also see [27]). For the solution of the indefinite system of (2) at each step, we have used in our experiments a stable conjugate gradient-type algorithm, SYMMLQ, developed by Paige and Saunders [13]. Widlund [26] provided a theory for the understanding of the convergence of SYMMLQ (see also Stoer and Freund [21], Stoer [20], Freund [5], and Szyld [23]). The convergence of SYMMLQ is accelerated by the use of preconditioning (see Szyld and Widlund [25], Szyld [23]). See also Morgan and Scott [9] where preconditioning is used in conjunction with Davidson's method but only for the standard case $B = I$.

When factorization of B is acceptable, several methods exist for the solution of (1) (see Stewart [19] for a survey and extensive bibliography). For the case $B = I$ the Lanczos method [8], [14], has become increasingly popular when a few of the eigenvalues at either end of the spectrum are sought. This method has been extensively studied (see, e.g., [11], [12], [17]). However, the Lanczos method cannot be used directly for the solution of the generalized problem (1). Scott [18] has attempted to use the Lanczos method for the solution of (1) by reducing it to a series of standard eigenvalue problems, i.e., with $B = I$, while Kalamboukis [7] presented a Lanczos-type method requiring the B-orthogonalization of the whole basis of the Krylov subspace at each step. Another idea is the combination of the Lanczos method with spectral transformation ([4], [3]).

The idea of combining inverse and Rayleigh quotient iteration (RQI) is of course quite natural. Our contribution consists of the use of an easily computable criterion to switch from one method to the other and the choice of preconditioned SYMMLQ for the solution of the—often almost singular—indefinite systems. The criterion used guarantees that if an eigenvalue exists in the interval J, the method will converge to that eigenvalue and not to any other. Of course, if there is more than one eigenvalue in J, only one of them will be obtained with the method presented. When J is free of eigenvalues, the inverse iteration method will converge linearly to the eigenvalue closest to J. Since the switching criteria is not satisfied, an inclusion theorem guarantees that the method determines the fact that there are no eigenvalues in J.

2. Convergence of inverse and Rayleigh quotient iteration. Let $\{\lambda_i, v_i\}$ be the eigenpairs of the pencil (A, B), i.e., solutions of problem (1). For convenience of presentation, we assume that the eigenvalues are numbered according to their proximity to γ, i.e., $0 < |\lambda_1 - \gamma| \leqq |\lambda_2 - \gamma| \leqq \cdots \leqq |\lambda_n - \gamma|$. We further assume that $|\lambda_1 - \gamma| < |\lambda_2 - \gamma|$ and we choose the eigenvectors so that $v_i^T B v_j = \delta_{ij}$.

At each step of the inverse iteration method (II), there are two residuals that can be computed, $q_s = Ax_s - \gamma Bx_s$, with respect to the shift γ, and $r_s = Ax_s - \mu_s Bx_s$ with respect to the Rayleigh quotient. They are related by the equation $r_s = q_s + (\gamma - \mu_s)Bx_s$. Since $x_s \to v_1$ and $\mu_s \to \lambda_1$, $r_s \to 0$ as $s \to \infty$. We use the B^{-1}-norm to study the convergence of q_s and obtain

$$\|r_s\|_{B^{-1}}^2 = \|q_s\|_{B^{-1}}^2 - (\gamma - \mu_s)^2. \tag{5}$$

This equality is important, since it shows that

$$\|r_s\|_{B^{-1}} \leqq \|q_s\|_{B^{-1}}, \tag{6}$$

$$|\gamma - \mu_s| \leqq \|q_s\|_{B^{-1}} \tag{7}$$

and that

$$\|q_s\|_{B^{-1}} \to |\gamma - \lambda_1| \quad \text{as } s \to \infty. \tag{8}$$

It turns out that the convergence sequence is actually monotonically decreasing; thus we have the following lemma.

LEMMA. $\|q_{s+1}\|_{B^{-1}} \leqq \|q_s\|_{B^{-1}}$.

Proof. Let $C = (A - \gamma B)^{-1}B$. From (2)-(4) we have $x_s = \Omega_s C^s x_0$, where $\Omega_s = \omega_1 \omega_2 \cdots \omega_s$. Then $q_s = \Omega_s B C^{s-1} x_0$, and using the fact that $\|x_s\|_B^2 = 1$ it follows that

$$\|q_s\|_{B^{-1}}^2 = q_s^T B^{-1} q_s = \frac{x_0^T (C^T)^{s-1} B C^{s-1} x_0}{x_0^T (C^T)^s B C^s x_0}. \tag{9}$$

By applying the Cauchy-Schwarz inequality to

$$\|B^{1/2} C^s x_0\|_2^2 = x_0^T (C^T)^s B C^s x_0 = (B^{1/2} C^{s-1} x_0,\ B^{1/2} C^{s+1} x_0),$$

the lemma follows. □

Given x_s and $\mu = \mu_s$, RQI consists of computing x_{s+1} from equations (2)-(4). The convergence of RQI is cubic in all practical cases (see [15], [14] and the references given therein for a full discussion). Wilkinson [27, pp. 364-368], Ruhe [16, pp. 168-170], and Szyld [23] proved the following results:

$$\omega_s \leqq \|r_s\|_{B^{-1}},$$
$$|\mu_s - \mu_{s+1}| \leqq \|r_s\|_{B^{-1}}, \tag{10}$$

$$\|r_{s+1}\|_{B^{-1}} \leqq \|r_s\|_{B^{-1}}, \tag{11}$$

$$|\mu_s - \lambda_i| \leqq \|r_s\|_{B^{-1}}, \tag{12}$$

where λ_i is the eigenvalue closest to μ_s, $|\mu_s - \mu_{s+1}| \to 0$, and in all practical cases, $\|r_s\|_{B^{-1}} \to 0$.

3. The eigenvalue algorithm. We can now turn to some analysis which motivates our algorithm. First, if

$$\varepsilon = \min_{\lambda_i \neq \lambda_j} |\lambda_i - \lambda_j|/4, \tag{13}$$

then equations (10)-(12) guarantee that if $\|r_s\|_{B^{-1}} < 2\varepsilon$, and thus $|\mu_s - \mu_{s+1}| < 2\varepsilon$, then μ_s and μ_{s+1} lie in $(\lambda_i - 2\varepsilon, \lambda_i + 2\varepsilon)$, and thus $\mu_s \to \lambda_i$ and not to any other eigenvalue (cf. [10], [27]). Note that if λ_1 lies in $J = (\gamma - \eta, \gamma + \eta)$ and

$$\eta \leqq \varepsilon, \tag{14}$$

then, by (13) λ_1 is the only eigenvalue in J and furthermore if

$$\|r_s\|_{B^{-1}} < \eta \tag{15}$$

and

$$|\mu_s - \gamma| \leqq \eta, \tag{16}$$

then $|\mu_{s+1} - \lambda_1| < 2\eta$ and $\mu_s \to \lambda_1$, i.e., RQI converges to λ_1, the only eigenvalue in J, and not to any other eigenvalue outside J.

Second, from (8) and the lemma we obtain the inclusion thereom for II

$$|\gamma - \lambda_1| \leqq \|q_s\|_{B^{-1}}. \tag{17}$$

This inequality can also be obtained directly without the lemma (see, e.g., Parlett [14, p. 318]). But the lemma guarantees that the bound (17) never deteriorates as the iteration continues; (also, cf. Jensen [6]). From this it follows that when an eigenvalue is sought in the interval $J=(\gamma-\eta, \gamma+\eta)$ by II using the shift γ, then λ_1 can be guaranteed to lie in J as soon as

$$\|q_s\|_{B^{-1}} < \eta \tag{18}$$

and $\|q_{s'}\|_{B^{-1}} < \eta$ and by (6) $\|r_{s'}\|_{B^{-1}} < \eta$ for $s' \geqq s$.

If the starting vector for RQI is such that its Rayleigh quotient is close enough to λ_1, RQI converges to λ_1. It is then natural to apply a certain number of steps of II, to produce a "good" starting vector to RQI, and then proceed with RQI. The issue is, then, which criterion to use to determine when the switch from II to RQI should take place to guarantee that RQI converges to λ_1.

In our algorithm, there are two criteria for switching. One criterion is (18) and takes care of the case where $\lambda_1 \in J$. From (2)-(4) we point out that $q_s = \omega_s B x_{s-1}$ and thus

$$\|q_s\|_{B^{-1}} = \omega_s. \tag{19}$$

Since ω_s is computed in every iteration anyway, we obtain $\|q_s\|_{B^{-1}}$ (and thus $\|r_s\|_{B^{-1}}$) without further cost and even without computing q_s (or r_s), and (18) can thus be rewritten as

$$\omega_s < \eta. \tag{20}$$

If s is the first value for which (20) holds, then λ_1 is guaranteed to lie in J. We also know, because of (7), that μ_s lies in J. At this stage the method switches to RQI. If (14) holds (although this fact may not be known by the user), then since (6) implies (15), $\mu_s \to \lambda_1$ cubically. If (14) does not hold, from (10) and (6), we only know that $|\mu_{s+1}-\gamma| < 2\eta$, although it is reasonable to expect that the steps of inverse iteration would have diminished the component of x_0 in the directions other than v_1 in such a way that $|\mu_{s+1}-\gamma| \leqq \eta$. This is not taken for granted but rather the Rayleigh quotient μ_s is monitored at every iteration and if at some point it falls outside J, the method reverts to II with shift γ using the latest iterate of RQI as initial vector. In our experience, we have observed the need to return to II only on very rare occasions; and when it occurs, the new initial guesses are so good that only one or two inverse iteration steps are needed to drive the Rayleigh quotient back into J, and then RQI converges to a λ_i in J.

The other criterion for switching to RQI covers the case where there are no eigenvalues in J, i.e., λ_1, the eigenvalue closest to γ, is outside J. In this case, (20) is never satisfied. At the same time that μ_s approaches the eigenvalue λ_1 outside J, the Rayleigh quotient μ_s and the norm of the residuals cease to change rapidly. We recall that μ_s is related to the residuals by (5). We monitor the Rayleigh quotient and when its relative change

$$|\mu_s - \mu_{s-1}|/|\mu_s| < \delta, \tag{21}$$

where δ is a prescribed tolerance, we say that the process has become stationary, and we switch to RQI solely to accelerate the convergence to λ_1 outside J. There exists a danger that condition (21) is satisfied too soon, i.e., before the process actually becomes stationary. This would be the case if the starting vector x_0 happens to be close to an eigenvector v_i, corresponding to an eigenvalue outside J. The inverse iteration method then needs a couple of iterations to diminish the large components in that direction. For that reason, we only switch after a minimum number m of inverse

iterations, which is set by the user. A value of $m = 2$ is recommended. Alternatively, the method could be restarted with a different x_0. The algorithm stops when $\|r_s\|_{B^{-1}}$ falls below a prescribed tolerance τ.

The system in (2) can be solved with any method for indefinite systems. Several methods were compared by Szyld, Sánchez Sarmiento, and Tsujui de Santos [24]. Since we are interested in the case when the system is large and factorizations are impractical, we prefer a preconditioned version of SYMMLQ, a conjugate gradient-type method developed by Paige and Saunders [13] (see also Szyld and Widlund [25], Szyld [23]). We use positive definite preconditioners, and thus we stay in the framework of real eigenvalues. If problem (1) were derived as the discretization of an elliptic operator with variables coefficient, a good choice for a preconditioner would be the discretization of an elliptic operator with constant coefficients or a fast Poisson solver as in [25].

We note that in our experience the algorithm is self-correcting in the sense that if the solution of the indefinite system (2) is solved less accurately, more iterations are needed, but the method ultimately converges to the appropriate eigenpair. Moreover, it should be noted that although the system (2) becomes more singular when μ approaches an eigenvalue, the normalized solution x_s approaches an eigenvector of a pencil close to (A, B) (see [27, pp. 370-373]). This concludes the presentation of the algorithm TLIME, a two-level iterative method for eigenvalue calculations.

4. Numerical experiments. We report tests of the algorithm TLIME with matrices obtained by a finite element discretization of Sturm-Liouville problems (see e.g., [22, Chaps. 1, 6]). Consider the eigenvalue problem

$$-\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)u = \lambda u, \tag{22}$$

where $u = u(x)$, $0 < x < \pi$ and $u(0) = 0$, $u'(\pi) = 0$, and $p(x) > 0$.

We have used TLIME to obtain eigenvalues of the discretization of the problem (22) with $p(x) = 2 + \sin x$ and $q = 1.5$ for several values of n, the size of the problem, and in different parts of the spectrum. To precondition this problem we have used the same discretization of the Sturm-Liouville operator with constant $p = 2$ and $q = 1.5$. We have chosen this model problem for convenience in experimentation since it permits the generation, by a suitable change of parameters, of problems with many levels of difficulty. We realize that this tridiagonal problem can be solved with methods which factor the matrices. The results of some of these experiments are reported in Tables 1 and 2. Table 1 shows timings for different values of n, when TLIME was called with $\gamma = 6$, $\eta = 3$ and in Table 2, the case $\gamma = 200$, $\eta = 30$ is reported. The timings shown were obtained on the Cyber 170/720 of New York University's Academic Computing

TABLE 1

n	Approximate eigenvalue obtained	Log_{10} of residual norm	Number of outer iterations	Total number of SYMMLQ iterations	Time in CPU seconds
250	7.382540	−7	5	24	2.887
500	7.382404	−7	5	24	5.419
1000	7.382370	−6	5	24	10.324
2000	7.382361	−6	5	24	20.177
5000	7.382360	−6	5	24	49.887
7500	7.382595	−6	5	24	74.602

TABLE 2

n	Approximate eigenvalue obtained	Log_{10} of residual norm	Number of outer iterations	Total number of SYMMLQ iterations	Time in CPU seconds
250	190.1242	−11	5	109	8.469
500	189.9881	−10	5	115	17.368
1000	189.9541	−9	5	106	32.255
2000	189.9456	−9	5	102	60.620
5000	189.9432	−8	5	104	152.642
7500	189.9429	−8	5	104	227.594

Facility, and they clearly show how effective the method is. TLIME with the use of preconditioning appears to be optimal in the sense that it takes a fixed amount of computational effort per mesh point.

Acknowledgment. The author is indebted to Olof B. Widlund for enlightening discussions and comments on earlier versions of the manuscript.

REFERENCES

[1] L. ÅASLAND AND P. BJØRSTAD, *The generalized eigenvalue problem in ship design and offshore industry—a comparison of traditional methods with the Lanczos process*, in Matrix Pencils, B. Kågström and A. Ruhe, eds., Proc. Conference Held at Pite Havsbad, Sweden, March 22-24, 1982; Lecture Notes in Mathematics 973, Springer-Verlag, Berlin, New York, 1983, pp. 146-155.

[2] D. CATLEY, *Simultaneous iteration applied to the solution of finite element models in ship vibration*, in Sparse Matrices and Their Uses, I. S. Duff, ed., Academic Press, London, 1981, pp. 325-333.

[3] T. ERICSSON, *Algorithms for large sparse symmetric generalized eignevalue problems*, Ph.D. thesis, Report UMINF-108.83, Umea Universitet, Sweden, May 25, 1983.

[4] T. ERICSSON AND A. RUHE, *The spectral transformation Lanczos method for the numerical solution of large sparse generalized symmetric eigenvalue problems*, Math. Comp., 35 (1980), pp. 1251-1268.

[5] R. FREUND, *Über Einige CG-Ähnliche Verfahren zur Lösung Linearer Gleichungssysteme*, Ph.D. thesis, Universität Würzburg, Federal Republic of Germany, May 1983.

[6] P. S. JENSEN, *An inclusion theorem related to inverse iteration*, Linear Algebra Appl., 6 (1973), pp. 209-215.

[7] T. Z. KALAMBOUKIS, *A Lanczos-type algorithm for the generalized eigenvalue problem* $Ax = \lambda Bx$, J. Comput. Phys., 53 (1984), pp. 82-89.

[8] C. LANCZOS, *An iteration method for the solution of the eigenvalue problem of linear differential operators*, J. Res. Nat. Bur. Standards, 45 (1950), pp. 255-282.

[9] R. B. MORGAN AND D. S. SCOTT, *Generalizations of Davison's method for computing eigenvalues of sparse matrices*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 817-825.

[10] A. M. OSTROWSKI, *On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors*, I, Arch. Rational Mech. Anal., 1 (1957-8), pp. 233-241.

[11] C. C. PAIGE, *Practical use of the symmetric Lanczos process with reorthogonalization*, BIT, 10 (1970) 183-195.

[12] ———, *Computational variants of the Lanczos method for the eigenproblem*, J. Inst. Math. Applics., 10 (1972), pp. 373-381.

[13] C. C. PAIGE AND M. A. SAUNDERS, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617-629.

[14] B. N. PARLETT, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[15] B. N. PARLETT AND W. KAHAN, *On the convergence of a practical QR algorithm*, in Information Processing 68, Vol. 1, A. J. H. Morrell, ed., North-Holland, Amsterdam, 1969, pp. 114-118.

[16] A. RUHE, *Computation of eigenvalues and eigenvectors*, in Sparse Matrix Techniques, V. A. Baker, ed., Copenhagen, 1976; Lecture Notes in Mathematics 572, Springer-Verlag, Berlin, New York, 1977, pp. 130-184.

[17] D. S. SCOTT, *Analysis of the symmetric Lanczos algorithm*, Ph.D. thesis, Department of Mathematics, University of California, Berkeley, CA, 1978.

[18] D. S. SCOTT, *Solving sparse symmetric generalized eigenvalue problems without factorization*, SIAM J. Numer. Anal., 18 (1981), pp. 102–110.

[19] G. W. STEWART, *A bibliographical tour of the large, sparse generalized eigenvalue problem*, in Sparse Matrix Computations, J. R. Bunch and D. J. Rose, eds., Academic Press, New York, 1976, pp. 113–130.

[20] J. STOER, *Solution of large linear systems of equations by conjugate gradient type methods*, in Mathematical Programming The State of the Art—Bonn 1982, A. Bachem, M. Gröstchel, and B. Korte, eds., Springer-Verlag, Berlin, 1983, pp. 540–565.

[21] J. STOER AND R. FREUND, *On the solution of large indefinite systems of linear equations by conjugate gradient algorithms*, in Computer Methods in Applied Sciences and Engineering V, R. Glowinski and J. L. Lions, eds., North-Holland, Amsterdam, 1982, pp. 35–53.

[22] G. STRANG AND G. J. FIX, *An Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs, NJ, 1973.

[23] D. B. SZYLD, *A two-level iterative method for large sparse generalized eigenvalue calculations*, Ph.D. thesis, Department of Mathematics, New York University, New York, October 1983.

[24] D. B. SZYLD, G. SÁNCHEZ SARMIENTO, AND S. TSUJUI DE SANTOS, *Análisis de diversos metodos iterativos para la resolución del problema* $Ax = \lambda Bx$, Revista SIGMA, 6 (1980), pp. 1–15.

[25] D. B. SZYLD AND O. B. WIDLUND, *Applications of conjugate gradient type methods to eigenvalue calculations*, in Advances in Computer Methods for Partial Differential Equations—III, R. Vichnevetsky and R. S. Stepleman, eds., IMACS, 1979, pp. 167–173.

[26] O. B. WIDLUND, Presentation at the Gatlinburg VII Conference in Asilomar, CA, December 16, 1977.

[27] J. H. WILKINSON, *Inverse iteration in theory and practice*, Sympos. Math., 10 (1972), pp. 361–379.