NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl., **5,** 33–55 (1998)

Harmonic Projection Methods for Large Non-symmetric Eigenvalue Problems

Ronald B. Morgan[1] and Min Zeng[2]

[1]*Department of Mathematics, Baylor University, Waco, TX 76798-7328, USA. e-mail: morganr@baylor.edu.*
[2]*Department of Mathematics, The University of Missouri, Columbia, MO 65211, USA*

The problem of finding interior eigenvalues of a large nonsymmetric matrix is examined. A procedure for extracting approximate eigenpairs from a subspace is discussed. It is related to the Rayleigh–Ritz procedure, but is designed for finding interior eigenvalues. Harmonic Ritz values and other approximate eigenvalues are generated. This procedure can be applied to the Arnoldi method, to preconditioning methods, and to other methods for nonsymmetric eigenvalue problems that use the Rayleigh–Ritz procedure. The subject of estimating the boundary of the entire spectrum is briefly discussed, and the importance of preconditioning for interior eigenvalue problems is mentioned. © 1998 John Wiley & Sons, Ltd.

KEY WORDS eigenvalues; Arnoldi; Krylov subspaces; preconditioning; Davidson's method

1. Introduction

We examine the problem of finding a few of the interior eigenvalues of a large nonsymmetric matrix. These eigenvalues are far enough toward the middle of the spectrum that it is impractical to compute all of the eigenvalues exterior to them. It is also assumed that factoring the matrix is impractical. This makes the problem very difficult.

There are an increasing number of applications for which interior eigenvalues are desired. Eigenvalues in the middle of the spectrum are needed for studying tidal motion [7], and for both adaptive polynomial preconditioning [1, 13] and Richardson's iteration [5]

Contract grant sponsor: National Science Foundation; Contract grant number: CCR-9102221.

CCC 1070–5325/98/010033– 23 $17.50
©1998 John Wiley & Sons, Ltd.
Received 15 September 1994
Revised 13 August 1997

for indefinite systems of linear equations. Nonsymmetric examples include some weather forecasting models [51], and stability analysis in computational fluid dynamics [25]. In addition, some iterative methods for non-symmetric indefinite systems of linear equations [12, 23, 53] require estimates of interior eigenvalues. And approximations of the eigenvectors corresponding to interior eigenvalues can be used to improve the convergence of the GMRES method for non-symmetric systems of linear equations [30].

Let B be a large, n by n matrix. The standard way of computing eigenvalues is to use the Rayleigh–Ritz procedure [35, 39] to extract approximate eigenvectors from a subspace of $\mathbf{R^n}$.

1.1. The Rayleigh–Ritz procedure

1. Choose a subspace S of dimension j. Compute P, an n by j orthonormal matrix with columns spanning S.
2. Calculate the j by j matrix $G = P^* B P$.
3. Compute eigenpairs (ρ_i, g_i) of G.
4. The ρ_i's are approximate eigenvalues of B. They are called Ritz values. If desired, compute the corresponding approximate eigenvectors $y_i = P g_i$, called Ritz vectors. Norms of residual vectors $r_i \equiv \frac{B y_i - \rho_i y_i}{||y_i||}$ can be computed to check the accuracy.

The subspace for the Rayleigh–Ritz procedure is often chosen to be a Krylov subspace [35, 39]. For symmetric matrices this leads to the Lanczos algorithm [21, 35]. For non-symmetric problems, the Lanczos algorithm can be generalized as the Arnoldi method [2, 28, 39, 40, 42, 49] or as the non-symmetric Lanczos algorithm [21, 36, 39]. Subspace iteration [18, 35] is a related method that also uses the Rayleigh–Ritz procedure. Another approach is to use preconditioning methods such as Davidson's method and its generalizations [9, 27, 29, 31, 48, 50]. In these methods the Rayleigh–Ritz procedure is applied to a subspace generated with a preconditioned operator.

Unfortunately, the Rayleigh–Ritz procedure usually does a better job of finding estimates for exterior eigenvalues than for interior eigenvalues. For the case of a symmetric matrix, the Rayleigh–Ritz procedure achieves a collective optimality for all of its approximations [35, pp. 215–217]. But also, approximations of the most exterior eigenvalues are optimal in a certain sense. By this we refer to the fact that one of the Ritz values is the largest possible Rayleigh quotient from a vector in the subspace [35, p. 215]. Thus, this Ritz value is the closest possible approximation to the largest eigenvalue. This Ritz value is not accurate if a poor subspace is used, but it is the best possible from the given subspace. With a non-symmetric matrix, no optimality is guaranteed, but the Rayleigh–Ritz procedure still usually does a good job finding exterior eigenvalues (for highly non-normal matrices, the situation can be more complicated; see, for example, [52]). For interior eigenvalues, the Rayleigh–Ritz procedure may give poor approximations even in the symmetric case. We discuss this further in the next paragraph and explain what generally happens.

When a Ritz value is near the exterior of the spectrum, the corresponding Ritz vector usually has some significance. Even if it is not an accurate approximation to a particular eigenvector, it is generally weighted more heavily in directions of eigenvectors corresponding to eigenvalues in that region of the spectrum. Now consider a Ritz value in the interior of the spectrum. The Ritz vector may be mostly a combination of nearby eigenvectors, or it may have large components of eigenvectors from all over the spectrum. The effects of these different components can cancel each other out to produce a Ritz value in the interior. We

© 1998 John Wiley & Sons, Ltd.

refer to this as a spurious Ritz value, because its value is not significant. Suppose now that the subspace contains an accurate approximation of a particular eigenvector. Then if the corresponding eigenvalue is on the exterior of the spectrum, the Rayleigh–Ritz procedure will generally give a Ritz vector of about the desired accuracy. But for an interior eigenvalue, there may or may not be a Ritz vector of that accuracy (see Example 1 in [26]). Loss of accuracy happens when there is a spurious Ritz value that falls near the significant Ritz value. With two Ritz values near each other, the corresponding eigenvectors blend together. There is no way to know when a spurious Ritz value will fall near and decrease the accuracy of the approximate eigenvector. However, as the approximation in the subspace becomes more accurate, it is less likely to be perturbed by a spurious Ritz value. So there are fewer problems as the subspace gets larger.

In conclusion, there are two difficulties in using the Rayleigh–Ritz procedure to compute interior eigenvalues. First, the approximations may be less accurate than expected from the given subspace. Second, even if there is a good approximation, it may be difficult to distinguish it from spurious Ritz values. To tell the difference between the significant and the spurious Ritz values, it is necessary to compute the residual norms.

Because the Rayleigh–Ritz procedure is better for exterior eigenvalues, it would be desirable to convert interior eigenvalues into exterior ones. This paper discusses a way of doing this implicitly. A version of the Rayleigh–Ritz procedure is given that is designed for finding interior eigenvalues. It is called 'Interior' or 'Harmonic' Rayleigh–Ritz. It is an extension of the procedure in [26] for symmetric matrices. See [33] for further discussion and for relationships for solving linear equations.

Some specific cases of the procedure presented in this paper have been considered elsewhere, as is mentioned in the next paragragh. The main contribution of this paper is to give the general procedure and give a full development, including use with different eigenvalue methods. Extending the procedure from the symmetric case in [26] is fairly straightforward, but the application and analysis are somewhat different. We feel there are actually three tools needed for solving realistic interior eigenvalue problems: an effective projection technique, a method for improving the spectrum with preconditioning and good preconditioners for indefinite matrices. This paper concentrates on the first two of these tools. It is demonstrated that the Interior Rayleigh–Ritz procedure is better than standard Rayleigh–Ritz for interior eigenvalues, and it is shown that the preconditioning approach of Davidson type methods can generate much better subspaces. However, the paper does not give preconditioners for general interior eigenvalue problems. It is only indicated that it makes sense to try preconditioners that have been developed for indefinite linear equations.

Section 2 gives the projection method for extracting interior eigenvalues from a given subspace. The next two sections consider applications of this procedure to specific eigenvalue methods. Section 3 deals with the Arnoldi method. An interior eigenvalue version of the Arnoldi method is developed. A special case of this method is considered briefly in [4] and is also related to finding zeros of the GMRES polynomial [14, 15, 22, 23, 32, 44]. There is also a relationship to the Temple–Lehmann bounds; see [19, 20, 33]. Section 4 deals with preconditioning methods for eigenvalue problems. See also [26] for consideration of the symmetric case, and see [48] for related work with the Jacobi–Davidson method.

© 1998 John Wiley & Sons, Ltd.

2. A modified Rayleigh–Ritz procedure for interior eigenvalues

We consider the n by n non-symmetric eigenvalue problem

$$Bz = \lambda z \tag{2.1}$$

Let S be a j-dimensional subspace of $\mathbf{R^n}$. It is from this subspace that we wish to extract approximate eigenvectors. Suppose the eigenvalues nearest to σ are needed. Then ideally, the Rayleigh–Ritz procedure should be used with the shifted and inverted operator $(B-\sigma I)^{-1}$ [11, 45]. The significant eigenvalues are shifted to the exterior of the spectrum of this operator. The shifted and inverted operator not only yields a better Rayleigh–Ritz procedure, but can also be used to develop a very good subspace. However, we avoid this operator because a matrix–vector product with $(B-\sigma I)^{-1}$ requires either a matrix factorization (assumed to be impractical) or an iterative method to solve a system of linear equations. We make use of the this inverted operator in deriving a procedure, but in implementing the actual procedure, no inverted operator is needed; and importantly, the subspace from which approximate eigenvectors are extracted is not generated with the inverted operator. The Rayleigh–Ritz extraction is improved, but the subspace is not improved.

The generalized Rayleigh–Ritz procedure uses the reduced eigenvalue problem

$$Q^*BQg = \theta Q^*Qg \tag{2.2}$$

where Q is an n by j matrix whose columns span the desired subspace S. Unlike for standard Rayleigh–Ritz, Q is not required to have orthonormal columns.

For the shifted and inverted problem, applying generalized Rayleigh–Ritz gives

$$Q^*(B-\sigma I)^{-1}Qg = \frac{1}{\theta-\sigma}\, Q^*Q\, g \tag{2.3}$$

where (θ, Qg) is an approximate eigenpair for B. Ideally, this should be solved with Q representing the subspace S. Because of the difficulty with the inverted operator, instead let P have columns spanning S and let $Q = (B-\sigma I)P$. Then (2.3) becomes

$$P^*(B-\sigma I)^*Pg = \frac{1}{\theta-\sigma}\, P^*(B-\sigma I)^*(B-\sigma I)Pg \tag{2.4}$$

Solving this small j by j generalized eigenvalue problem gives approximate eigenpairs $(\frac{1}{\theta-\sigma}, Qg)$ of $(B-\sigma I)^{-1}$, since equation (2.3) is indirectly being solved. Corresponding to this are approximate eigenpairs (θ, Qg) of B, but a better choice for an approximate eigenpair of B is (ρ, Pg), where ρ is the Rayleigh quotient of Pg with respect to B. The vector Pg is generally a better approximate eigenvector than Qg because $Pg = (B-\sigma I)^{-1}Qg$. So, Pg is the result of applying one step of a shifted inverse iteration to Qg. Since ρ is the Rayleigh quotient of a better vector, it is generally a better approximate eigenvalue than θ. Using the terminology in [33], θ is a harmonic Ritz value and Pg is a harmonic Ritz vector.

We next give the algorithm for this approach, which we call Interior Rayleigh–Ritz. As mentioned in the introduction, it is a generalization of the procedure given for symmetric matrices in [26].

© 1998 John Wiley & Sons, Ltd.

2.1. *'Interior' or 'Harmonic' Rayleigh–Ritz*

1. Choose a subspace of dimension j. Compute P, an n by j matrix with columns spanning the subspace.
2. Calculate $G = P^*(B - \sigma I)P$ and $W = P^*(B - \sigma I)^*(B - \sigma I)P$.
3. Compute the desired number of the largest eigenpairs of $G^* g_i = \alpha_i W g_i$.
4. As desired, compute $\theta_i = \sigma + \frac{1}{\alpha_i}$, $y_i = P g_i$, and $\rho_i = \frac{y_i^* B y_i}{y_i^* y_i}$. Norms of residual vectors $r_i = \frac{B y_i - \rho_i y_i}{||y_i||}$ can be computed to check the accuracy.

We expect that the approximations given by Interior Rayleigh–Ritz will be significant because $(\frac{1}{\theta_i - \sigma}, Q g_i)$ is an approximation of an exterior eigenpair of $(B - \sigma I)^{-1}$. As discussed in Section 1, we expect approximations to exterior eigenvalues to have some significance. In the symmetric case, an indirect optimality can be established; see [26].

In comparing this procedure with standard Rayleigh–Ritz, we do not expect a remarkable improvement. This is because the two methods use the same subspace. As mentioned earlier, for a greater improvement, the shifted-and-inverted operator should be used to generate a better subspace. This situation is somewhat analogous to comparing SYMMLQ with MINRES [33, 34] or comparing FOM with GMRES [43] for solving indefinite linear equations. While MINRES (or GMRES) is expected to be better than SYMMLQ (or FOM), the difference may be marginal because both methods use the same subspace. In the non-symmetric case with GMRES versus FOM, the difference is more important because restarting is frequently used, and it is important to restart with the best approximate solution. Likewise for non-symmetric eigenvalue problems, the Interior Rayleigh–Ritz may be more important than it was for symmetric problems. There is an important exception to this analogy with linear equations methods. The preconditioning methods for eigenvalues that are discussed in Section 4 use an approximate eigenvector at every step to help generate the subspace. So the accuracy of the subspace is dependent upon having good approximate eigenvectors along the way. The Interior Rayleigh–Ritz procedure can then have a quite significant advantage.

Effectiveness of the Interior Rayleigh–Ritz method is reduced when σ is equal or very near to an eigenvalue of B, but the procedure seems to still be useful (see the last part of example 1 in the next section and also the discussion in [26]).

The ρ_i values can be found without forming y_i, if P is orthonormal. For g_i of unit length,

$$\begin{aligned} \rho_i &= \frac{(P g_i)^* B P g_i}{(P g_i)^* P g_i} \\ &= \sigma + \frac{g_i^* P^*(B - \sigma I) P g_i}{g_i^* P^* P g_i} \\ &= \sigma + g_i^* G g_i \end{aligned} \tag{2.5}$$

The following theorem gives a simple formula for the residual norm. This is a little surprising, because there is not a corresponding formula for standard Rayleigh–Ritz. The additional work done in forming and solving the small generalized eigenvalue problem (2.4) allows this result. Note however, that the formula can lose accuracy as the approximate eigenvector converges and the θ and ρ values become too close.

© 1998 John Wiley & Sons, Ltd.

Theorem 2.1. *The residual norm can be calculated with the formula*

$$||r_i||^2 = (\rho_i - \sigma)^*(\theta_i - \rho_i) \tag{2.6}$$

Proof
For simplicity, we drop the subscripts. Let $y = Pg$ be a unit vector. Using first equation (2.4), then the definition of ρ,

$$\begin{aligned} y^*(B - \sigma I)^*(B - \sigma I)y &= (\theta - \sigma)y^*(B - \sigma I)^* y \\ &= (\theta - \sigma)(\rho - \sigma)^* \end{aligned} \tag{2.7}$$

Next

$$\begin{aligned} r^*r &= (By - \rho y)^*(By - \rho y) \\ &= \big((By - \sigma y) + (\sigma - \rho)y\big)^*\big((By - \sigma y) + (\sigma - \rho)y\big) \end{aligned}$$

Multiplying this out and using equation (2.7) and the definition of ρ, we have

$$\begin{aligned} r^*r &= (\rho - \sigma)^*(\theta - \sigma) + (\sigma - \rho)^*(\rho - \sigma) + (\sigma - \rho)(\rho - \sigma)^* + (\sigma - \rho)^*(\sigma - \rho) \\ &= (\rho - \sigma)^*(\theta - \rho), \end{aligned}$$

and the proof is complete. ■

The Rayleigh–Ritz procedure is often used iteratively with subspaces of increasing dimension and improving accuracy. Assume that $r_i^* r_i$ is converging to zero. If σ is not equal to an eigenvalue of B, we can bound ρ_i away from σ. Then from (2.6), $|\theta_i - \rho_i|$ converges to zero. So the θ_i and ρ_i values generally get closer together as the approximation gets more accurate. The following interesting relationship between θ_i and ρ_i can also be derived.

Corollary 2.1. *The values σ, θ_i and ρ_i are on the same line in the complex plane. Furthermore, ρ_i is between σ and θ_i.*

Proof
It is a simple calculation to establish for complex numbers u and v, that if u^*v is a real number, then the vectors in the complex plane corresponding to u and v are parallel. Furthermore, if $u^*v > 0$, then the vectors u and v have the same direction. Note that r^*r is a positive-real number. From equation (2.6) it follows that if the complex numbers $(\rho_i - \sigma)$ and $(\theta_i - \rho_i)$ are viewed as vectors, they have the same direction. The desired result follows. ■

Even though the θ or harmonic values are not expected to be as accurate as the ρ values, they can be useful. It is possible to have ρ values near σ that are spurious. With the θ's this is very unlikely, since θ comes from the shifted and inverted problem. In fact, it is easy to show that there is a region around σ in which no θ can fall.

Theorem 2.2. *Assume σ is not an eigenvalue of B. Then there exists a positive number δ such that $|\theta_i - \sigma| > \delta$.*

Proof
The matrix $(B - \sigma I)^{-1}$ exists, by assumption. Since $\frac{1}{\theta_i - \sigma}$ is the Rayleigh quotient of Qg_i with respect to $(B - \sigma I)^{-1}$, $\frac{1}{\theta_i - \sigma} \in F\big((B - \sigma I)^{-1}\big)$, the field of values of $(B - \sigma)^{-1}$.

© 1998 John Wiley & Sons, Ltd.

$F\left((B-\sigma I)^{-1}\right)$ is bounded, so there is a $\beta > 0$ such that $|\alpha| < \beta$, for all $\alpha \in F\left((B-\sigma I)^{-1}\right)$. Therefore $|\frac{1}{\theta_i - \sigma}| < \beta$. Let $\delta = \frac{1}{\beta}$. Then $|\theta_i - \sigma| > \delta$. ■

The θ values can be used to pick which approximate eigenpairs are significant. The approximate eigenpairs corresponding to the θ values nearest σ are generally the most significant.

The Interior Rayleigh–Ritz procedure can be viewed as an oblique Petrov–Galerkin projection [39]. The matrix is $(B - \sigma I)$, the approximations are from the subspace spanned by the columns of P, and the orthogonalization is enforced against the subspace spanned by the columns of $(B - \sigma I)P$.

An orthogonality condition can be given for the vectors $\tilde{r}_i = By_i - \theta_i y_i$. Note these vectors are not the residual vectors given in the Interior Rayleigh–Ritz algorithm. Here the θ value is used instead of the better ρ value. Again following [33], we call these harmonic residual vectors. This result is also given in [48], for the case of $\sigma = 0$.

Theorem 2.3. *The harmonic residual vectors $\tilde{r}_i \equiv By_i - \theta_i y_i$ are orthogonal to the subspace spanned by the columns of $(B - \sigma I)P$.*

Proof

$$\begin{aligned}
\left((B-\sigma I)P\right)^* \tilde{r}_i &= P^*(B-\sigma I)^*(By_i - \theta_i y_i) \\
&= P^*(B-\sigma I)^*\left((B-\sigma I)y_i - (\theta_i - \sigma)y_i\right) \\
&= P^*(B-\sigma I)^*(B-\sigma I)y_i - (\theta_i - \sigma)P^*(B-\sigma I)^* y_i \\
&= P^*(B-\sigma I)^*(B-\sigma I)Pg_i - (\theta_i - \sigma)P^*(B-\sigma I)^* P g_i \\
&= 0
\end{aligned}$$

The last step is because (θ_i, g_i) is a solution of (2.4). ■

3. Applications to the Arnoldi method

The Arnoldi method [2, 28, 39, 40, 42, 49] is a standard way to find eigenpairs of a large non-symmetric matrix. The Arnoldi iteration for matrix $B - \sigma I$ can be represented by

$$(B - \sigma I)P = PG + \beta_j p_{j+1} e_j^*$$

Here P is an n by j orthonormal matrix with columns spanning a Krylov subspace of dimension j, p_{j+1} will be the next column of P, G is a j by j upper-Hessenberg matrix, and β_j will be the $(j+1, j)$ entry of the next G. Equation (2.4) becomes

$$G^* g = \frac{1}{\theta - \sigma}(G^* G + \beta_j^2 e_j e_j^*) g \tag{3.1}$$

This can be converted to a nicer form with an orthogonal factorization $G = S\bar{R}$, where S is orthogonal (or unitary if σ is complex) and $\bar{R}$ is upper-triangular. Then

$$\begin{aligned}
G^* G + \beta_j^2 e_j e_j^* &= \bar{R}^* \bar{R} + \beta_j^2 e_j e_j^* \\
&= R^* R
\end{aligned}$$

© 1998 John Wiley & Sons, Ltd.

where R is upper triangular and differs from $\bar{R}$ only in the jj element. For that element, $r_{jj}^2 = \bar{r}_{jj}^2 + \beta_j^2$. So equation (3.1) becomes

$$G^* g = \frac{1}{\theta - \sigma} (R^* R) g. \tag{3.2}$$

For experiments in this paper, we used (3.2), but there are ways to rewrite it. See [14, equation (5.22)], and from [33, 17] there is the formula

$$(G + \beta_j^2 f e_j^*) g = (\theta - \sigma) g$$

where $f = G^{-*} e_j$. This follows quickly from (3.1).

3.1. Interior or Harmonic Arnoldi

1. *Initialize:* Choose σ and starting vector p_1 of norm one.
2. *Generation of Krylov subspace:* for $j = 1, 2, ...,$ do:
   $g_{ij} = (p_i, (B - \sigma I) p_j), i = 1, 2, ..., j$
   $\hat{p}_{j+1} = (B - \sigma I) p_j - \sum_{i=1}^{j} g_{ij} p_i$
   $g_{j+1,j} = ||\hat{p}_{j+1}||$
   $\beta_j = g_{j+1,j}$, and
   $p_{j+1} = \hat{p}_{j+1} / \beta_j$
3. *The reduced problem:*
   Set up and solve $G^* g = \frac{1}{\theta - \sigma} (G^* G + \beta_j^2 e_j e_j^*) g$,
   for the eigenpairs with largest magnitude of $\frac{1}{\theta - \sigma}$.
4. *Find approximate eigenvalues and eigenvectors:* Compute ρ_i's using (2.5) and if desired $y_i = P g_i$. Residual norms can be checked for convergence.

The residual norms can be computed using (2.6), or for this Arnoldi case there is a more reliable formula:

$$\| r_i \|^2 = \| (G - \rho_i I) g_i \|^2 + \beta_j^2 g_{mi}^2$$

where g_{mi} is the mth entry of g_i

For the case of $\sigma = 0$, the θ_i values are mentioned in [4] and also have been studied recently as the zeros of the residual polynomials generated by GMRES. Use of GMRES zeros as approximate eigenvalues is discussed in [14, 15, 22, 23, 32, 44].

The reduced problem (3.2) is more complicated to set up and solve than the reduced problem in standard Arnoldi, particularly if σ is complex. However, the factored form of the right-hand side allows for a more efficient solution. Also, in terms of n, the expense of the interior version of Arnoldi is the same as for standard Arnoldi.

The Arnoldi iteration is independent of σ except for shifting the main diagonal of G (this has been observed before, see for example [8, 26]). We can get eigenvalue approximations near another σ by doing another orthogonal factorization (of a differently shifted G). In this way it is possible to get eigenvalue approximations over whatever range is desired. We note that the method is limited by the accuracy of the Krylov subspace. The subspace would not be expected to be effective for interior eigenvalues, unless either the desired eigenvalues are well separated from the rest of the spectrum or the subspace is quite large.

© 1998 John Wiley & Sons, Ltd.

We next characterize the approximate eigenvectors found by the Interior Arnoldi method. The harmonic Ritz vectors y_i can be written as $y_i = \phi(B)p_1$, where p_1 is the first column of P and ϕ is a polynomial of degree $j-1$. This is because y_i is from the Krylov subspace of dimension j. The following theorem shows that the roots of the polynomial ϕ are the harmonic Ritz values (the θ values from equation (3.2)). See [23] for a similar discussion of the GMRES polynomial.

Theorem 3.1. *If y_i is a Ritz vector from the Interior Arnoldi algorithm, then it can be written as $y_i = \phi(B)p_1$, where ϕ is a polynomial such that*

$$\phi(\lambda) = \alpha \prod_{k \neq i} (\lambda - \theta_k) \tag{3.3}$$

The θ values are defined in equations (2.4) and (3.2), and α is a normalizing factor.

Proof
Let $y_i = \phi(B)p_1$. As in Theorem 2.4, let the harmonic residual vector be $\tilde{r}_i \equiv By_i - \theta_i y_i$. Define a new polynomial $\psi(\lambda) = (\lambda - \theta_i)\phi(\lambda)$. Then $\tilde{r}_i = (B - \theta_i I)\phi(B)p_1 = \psi(B)p_1$. We will show that each θ_k, for k from 1 to j, is a root of ψ. From (2.4),

$$\begin{aligned} 0 &= P^*(B-\sigma I)^*(B-\sigma I)Pg_k - (\theta_k - \sigma)P^*(B-\sigma I)^*Pg_k. \\ &= P^*(B-\sigma I)^*(B-\theta_k I)Pg_k \\ &= \big((B-\sigma I)P\big)^*(B-\theta_k I)Pg_k \end{aligned}$$

We have shown that $(B-\theta_k I)Pg_k$ is orthogonal to the subspace spanned by the columns of $(B-\sigma I)P$. By Theorem 2.4, $\tilde{r}_i$ is also orthogonal to that subspace. So we have two vectors, $\tilde{r}_i$ and $(B-\theta_k I)Pg_k$, that are both in the Krylov subspace of dimension $j+1$ and are both orthogonal to a j-dimensional subspace. The two vectors must be multiples of each other. The polynomial corresponding to $(B-\theta_k I)Pg_k$ has a root at θ_k, so ψ has a root at θ_k.

Now, since ψ has roots at each θ_k, we have that ϕ has roots at all θ_k except for θ_i. These are the only roots, because ϕ is degree $j-1$. ∎

The result (3.3) also holds for standard Ritz vectors, except the θ values are then the standard Ritz values instead of harmonic. The proof is similar to the one just given; it uses orthogonality of the residual vectors to the subspace spanned by the columns of P. We just state the result. See also [35, p. 240, 46, 47].

Proposition 3.1. *If y_i is a Ritz vector from the standard Arnoldi algorithm, then it can be written as $y_i = \phi(B)p_1$, where ϕ is a polynomial and*

$$\phi(\lambda) = \alpha \prod_{k \neq i} (\lambda - \theta_k) \tag{3.4}$$

The θ values are the standard Arnoldi Ritz values, and α is a normalizing factor.

We can now return to the discussion in Section 1 about the accuracy of approximate eigenvectors for standard Rayleigh–Ritz. We look at the effect of a spurious Ritz value near a good Ritz value in the standard Arnoldi method (see [33] for a related discussion of the effect on the conjugate gradient method of inaccurate Ritz values that happen to fall near

© 1998 John Wiley & Sons, Ltd.

zero). Equation (3.4) shows why the Ritz vector can loose accuracy. Let λ_p be the desired eigenvalue, and let θ_i be a standard Ritz value approximating λ_p. Suppose another Ritz value, say θ_{i+1}, happens to fall near λ_p. For the Ritz vector y_i to be an accurate eigenvector approximation, the corresponding polynomial ϕ needs to be large at λ_p, compared with its magnitude at the other eigenvalues [35, ch. 12]. But,

$$\phi(\lambda_p) = \alpha \prod_{k \neq i} (\lambda_p - \theta_k)$$

and one of the factors in this product is $(\lambda_p - \theta_{i+1})$. This causes ϕ to be relatively small at λ_p and the Ritz vector to be inaccurate. Now, compare this to Arnoldi with Interior Rayleigh–Ritz. If λ_p is the eigenvalue nearest σ, it is very unlikely that a spurious θ_{i+1} value will hinder the accuracy of y_i. This is because the θ values in (3.3) are from Interior Rayleigh–Ritz. As mentioned in the previous section, they generally do not come near σ unless they correspond to accurate Ritz vectors. Next, we look at some examples for Interior Arnoldi.

Example 1.
As a test matrix, we generate a 998 by 998 non-symmetric, block-diagonal matrix. The eigenvalues of the matrix are distributed around two circles in the complex plane centered at -2 and 2 with radii of 1. On each circle, the real parts of the eigenvalues are equally spaced. The matrix has two 1 by 1 blocks with entries 1.0 and -1.0. There are 249 blocks that are 2 by 2 with diagonal entries both α, superdiagonal entry $\alpha - 3$, and subdiagonal entry $\alpha - 1$, where $\alpha = 1 + \frac{2}{250}, 1 + \frac{4}{250}, \ldots, 3 - \frac{2}{250}$. There are 249 blocks that are 2 by 2 with α on the diagonal, $\alpha + 1$ above, and $\alpha + 3$ below, where $\alpha = -1 - \frac{2}{250}, -1 - \frac{4}{250}, \ldots, -3 + \frac{2}{250}$. The desired eigenvalue is the one at 1. The value of σ is chosen to be 0.9. The starting vector has all entries equal to 1. Both standard Arnoldi and Interior Arnoldi are restarted every 50 iterations. Standard Arnoldi is restarted with the Ritz pair closest to 0.9, while interior Arnoldi uses the θ closest to 0.9 to pick the approximate eigenpair (ρ_i, y_i) for restarting. Figure 1 gives the results for 50 restarted runs. The residual norm of the best approximation to the eigenvalue at 1, at the time of each restart, is given. Interior Arnoldi is better than standard Arnoldi. When the standard Arnoldi method restarts, it does not always have a good approximation. However, neither method is satisfactory.

Next, we implement these methods differently. Additional approximate eigenvectors are included in the new subspace after a restart is done. We call these methods 'Arnoldi with Eigenvectors' and 'Interior Arnoldi with Eigenvectors'.

3.2. Arnoldi with eigenvectors

1. Let $y_1, y_2, \ldots, y_k$ be approximate eigenvectors.
2. Generate a Krylov subspace of dimension j with y_1 as the starting vector, using the Arnoldi iteration.
3. Orthogonalize $y_2, \ldots, y_k$ against the Krylov subspace. Let P be the matrix whose columns span the combined space Span$\{y_1, By_1, \ldots, B^{j-1}y_1, y_2, \ldots, y_k\}$.
4. Apply the Rayleigh–Ritz procedure to the subspace spanned by the columns of P. Choose Ritz vectors as the new approximate eigenvectors, and if needed, restart at Step 1.

© 1998 John Wiley & Sons, Ltd.

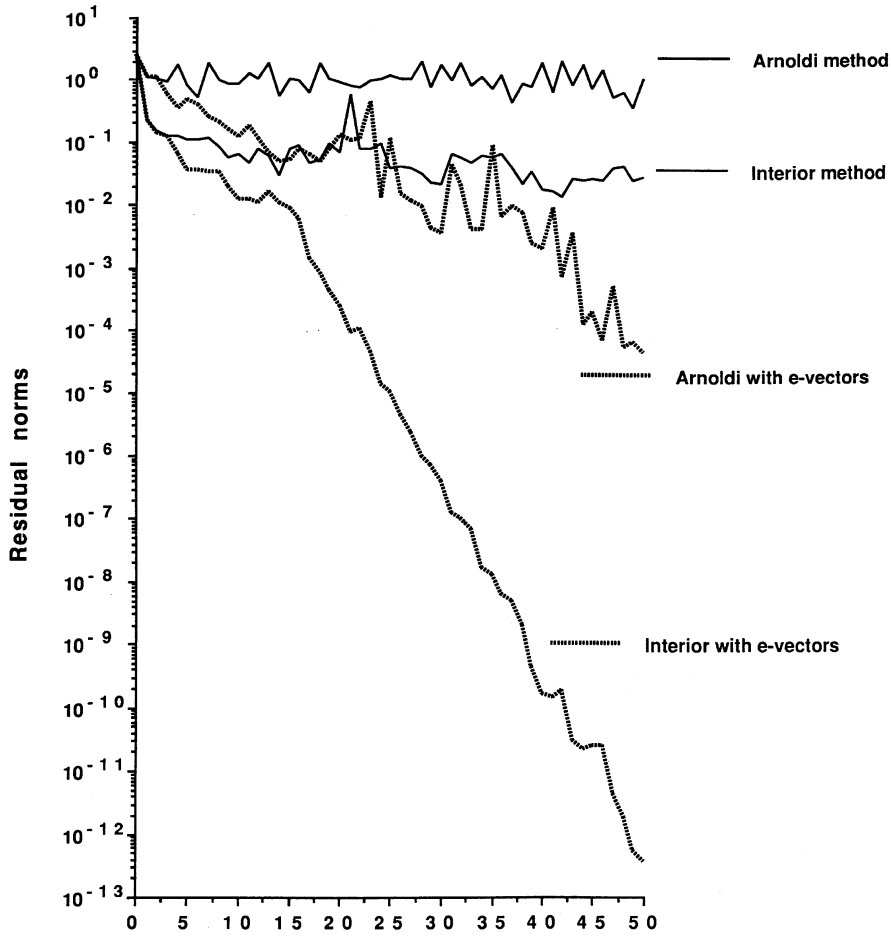

Figure 1. Matrix with eigenvalues on two circles.

© 1998 John Wiley & Sons, Ltd.

See [28] for a discussion of adding approximate eigenvectors to the subspace and for implementation details. It is interesting that the Arnoldi with Eigenvectors method is equivalent to Sorensen's Implicitly Restarted Arnoldi [49], as long as no eigenvector approximations are specified at the beginning of the procedure (before any restarts) [28].

Interior Arnoldi with Eigenvectors is the same except that Interior Rayleigh–Ritz is used. See [30] for an implemention. The brief algorithm is as follows.

3.3. Interior Arnoldi with eigenvectors

1. Let $y_1, y_2, \ldots, y_k$ be approximate eigenvectors.
2. Generate a Krylov subspace of dimension j with y_1 as the starting vector, using the Arnoldi iteration.
3. Apply the Interior Rayleigh–Ritz procedure to the subspace Span$\{y_1, By_1, \ldots, B^{j-1}y_1, y_2, \ldots, y_k\}$.
   Choose interior (harmonic) Ritz vectors as the new approximate eigenvectors, and if needed, go back to 1.

For our example, we use 47 Krylov basis vectors and three additional approximate eigenvectors, so the subspace is again of dimension 50. The results are also shown in Figure 1. Both the standard and interior methods improve with the addition of approximate eigenvectors. Between the two methods, the interior version is considerably better.

Finally, the value of σ is changed to 1.0, so that it equals the eigenvalue. The Interior Arnoldi with Eigenvectors method is tested, and it reaches residual norm of 0.48×10^{-5} after 50 runs. So, the method does not perform as well with $\sigma = \lambda$, but it is still as good as the standard Arnoldi with Eigenvectors method. Multiple copies of the eigenvalue at 1.0 appear, which could perhaps be used as an indication that σ should be changed [26].

Example 2.
Now we use the 1 001 by 1 001 tridiagonal matrix with 1 in each superdiagonal position, -1 for the subdiagonal elements, and with $-510, -509, \ldots, -12, -11, 0, 11, 12, \ldots 510$ on the main diagonal. The desired eigenvalue is at 0.0. The four eigenvalues nearest are $\pm 11.91 \pm 0.71i$, and the next few out from the origin are ± 13.1, ± 14.0, and ± 15.0. The σ value is chosen to be 1.0. Otherwise, the testing is done as in the previous example. The results are given in Figure 2. Again, the interior versions are better. They converge faster and also have smoother convergence. For this problem, using extra approximate eigenvectors is not really necessary with the interior method.

Example 3.
In this example, the goal is to extract information about the spectrum from an Arnoldi run without restarting. We use a matrix considered by Manteuffel and Starke [23]. It is a normal matrix with eigenvalues as shown in Figure 3. The best possible approximation to the region in the complex plane containing the spectrum is needed for solving linear equations with an iterative method [23]. Figure 4 has a plot of the standard Ritz values found during 100 different runs, all with subspaces of dimension 15, but with different random starting vectors (this is similar to a plot in Figure 2.1 in [23]). The Ritz values do not give a full picture of the spectrum. In particular, the location of the part of the spectrum nearest zero is unclear.

Manteuffel and Starke use the harmonic θ values from (2.4) with $\sigma = 0$ to get more information about the spectrum. Figure 5 shows this is helpful for identifying the portion of the spectrum furthest from the origin. However, θ values do not always work so well. If we

© 1998 John Wiley & Sons, Ltd.

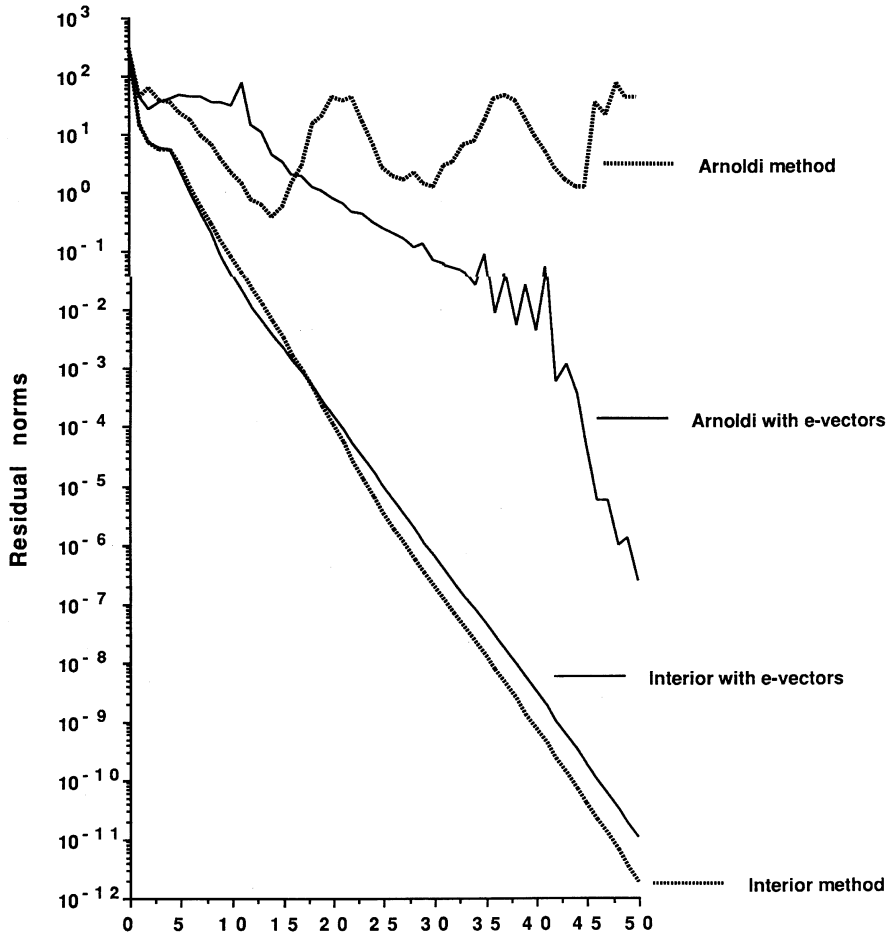

Figure 2. Tridiagonal matrix.

© 1998 John Wiley & Sons, Ltd.

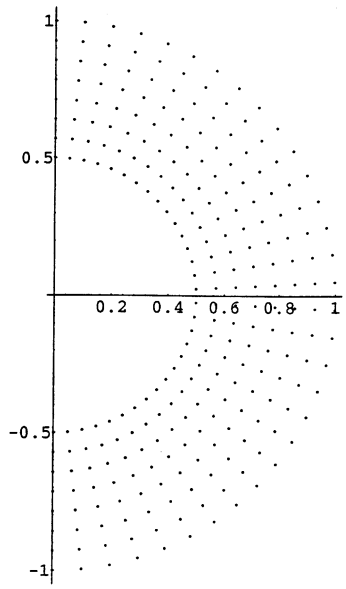

Figure 3. True eigenvalues.

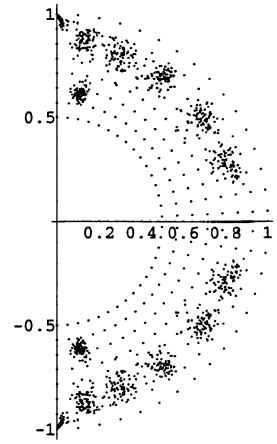

Figure 4. Regular Ritz values.

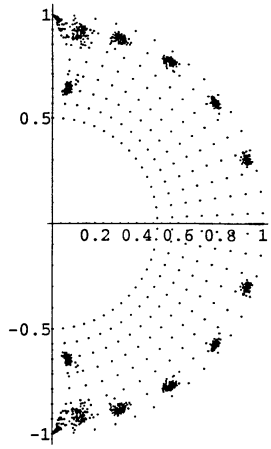

Figure 5. Theta values, sigma = 0.

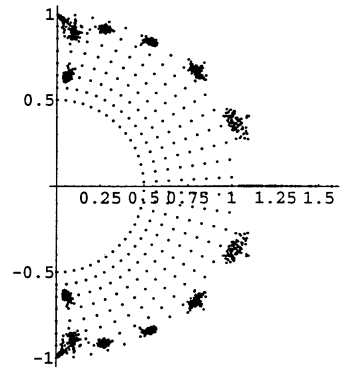

Figure 6. Theta values, sigma = 5.

© 1998 John Wiley & Sons, Ltd.

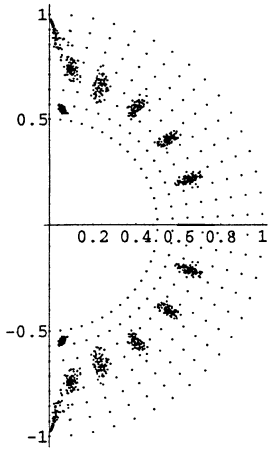

Figure 7. Rho values, sigma = 0.

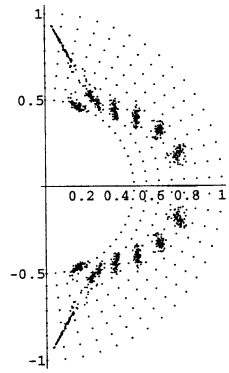

Figure 8. Rho values, sigma = 0.5.

recall where the θ values come from, they are designed for approximating the eigenvalues near σ. In the symmetric case, the θ's provide an outer bound on the eigenvalues near σ (there are eigenvalues between σ and the nearest θ's). But the θ's are not as accurate as ρ values for estimating eigenalues near σ, and for the portion of the spectrum away from σ, little accuracy can be expected. Figure 6 shows the θ values with $\sigma = 0.5$. There is one real θ value larger than 1.5, and many others outside of the spectrum.

Next consider the ρ values with $\sigma = 0$; see Figure 7. They give a better picture of the portion of the spectrum nearest to the origin. Figure 8 shows the ρ's with $\sigma = 0.5$. There is some danger in relying on ρ values if σ is chosen so that to one side of σ there are two separate portions of the spectrum with a gap in between. There is then a chance that ρ values will fall in the gap. This happens in Figure 8, with some of the ρ values creeping slightly into the empty circular region surrounding the origin.

It is possible to take the ρ values with several different σ's and combine them in one picture. Figure 9 has the standard Ritz values and the ρ values with $\sigma = -0.5, 0.0, 0.25, 0.5, 0.75, 1.0$ and 2.0 (we have used only real σ's, but complex values are also possible). By combining these different values, a good picture of the spectrum is obtained. Note that with a different procedure, Manteuffel and Starke also obtain a better picture of the spectrum than that given by just the Ritz and θ values [23].

We next look at the approximation of the spectrum that can be obtained if just one run of Arnoldi is available. Figure 10 has a plot with the Ritz values and the various ρ values as in Figure 9, but using only one run with a random starting vector. The picture of the spectrum is about as good as with 100 runs.

4. Applications to preconditioning methods

The Davidson and generalized Davidson (GD) methods [9, 27, 29, 31, 48, 50] use a subspace generated by a preconditioned operator instead of a Krylov subspace. However, like Arnoldi, they use the Rayleigh–Ritz procedure to extract approximate solutions from the subspace. The preconditioning can improve convergence, particularly for interior eigenvalues. We

© 1998 John Wiley & Sons, Ltd.

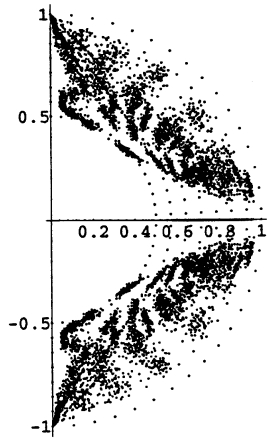

Figure 9. Ritz and rho values.

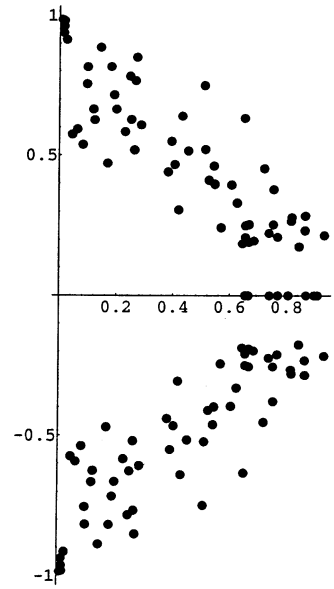

Figure 10. One run, Ritz and rho.

briefly describe preconditioning for eigenvalue problems, then give examples that show the potential of preconditioning for interior problems and show that Interior Rayleigh–Ritz is important.

The GD method builds a subspace iteratively. At every iteration the best current approximate eigenpair (ρ, y) is computed. Then the new vector for the expanded subspace is

$$v = (M - \alpha I)^{-1}(B - \rho I)y$$

where M is an approximation of B, and α is generally chosen to be near the desired eigenvalues. This vector v is not itself an approximate eigenvector, but can be viewed as a correction to the approximate eigenvector y. Alternatively, we can view this method as approximately generating a Krylov subspace. Suppose ρ approaches the eigenvalue λ with associated eigenvector z. Then the new portion of the subspace will be approximately Krylov with operator $N \equiv (M - \alpha I)^{-1}(B - \lambda I)$. This operator has z as an eigenvector, and the corresponding eigenvalue is 0. It is hoped that this eigenvalue will be well separated in the spectrum of N.

If $M - \alpha I$ is a good approximation of $B - \lambda I$, the other eigenvalues of N will be mostly compressed around 1. So, the preconditioning effectively changes the distribution of eigenvalues and makes convergence better for the desired eigenvalue.

4.1. The GD method

0. Choose an initial subspace Span$\{v_1, v_2, ..., v_k\}$.
   Iterate for $j = k + 1, k + 2, ...$:
1. Use either standard Rayleigh–Ritz or the Interior Rayleigh–Ritz procedure to extract an approximate eigenpair (ρ, y) from the subspace.
2. Choose α. Let $v_j = (M - \alpha I)^{-1}(B - \rho I)y$.

© 1998 John Wiley & Sons, Ltd.

3. Let the new expanded subspace be $Span\{v_1, v_2, ...v_{j-1}, v_j\}$, and go to 1.

There is a correspondence between the two problems of computing interior eigenvalues with the GD method and solving indefinite systems of linear equations with a preconditioned Krylov subspace method. For this discussion we assume that eigenvalues near zero are desired; for eigenvalues away from zero the correspondence would be to linear equations with a shifted matrix. When solving linear equations with B, the preconditioned matrix is $M^{-1}B$. Here M is an approximation of B, and left preconditioning is being used. A simple criterion for effective preconditioning is that the spectrum of $M^{-1}B$ be compressed around 1 and be away from 0. For computing the eigenvalue nearest 0, say λ, the same preconditioner can be used. Then the operator for GD will approach $N = M^{-1}(B - \lambda I)$. The desired eigenvalue of N is at 0, and we want the others compressed about 1, away from the one at 0. This is similar to the criterion for linear equations. The two operators $M^{-1}B$ and $M^{-1}(B - \lambda I)$ are similar, because λ is small. Their spectra have some distinct differences (specifically, the eigenvalue at 0), but are likely to be mostly the same. So, there is reason to try preconditioners developed for indefinite linear equations on interior eigenvalue problems.

There are several choices in implementing the GD method. For the examples that follow, complex basis vectors are used for the subspace (see [27] for an approach that uses real vectors). The α value is set equal to σ, the initial eigenvalue estimate, and is not changed. The Interior Rayleigh–Ritz version of the method orthogonalizes these vectors with respect to the $(B-\sigma I)^*(B-\sigma I)$ norm, so the right-hand side of equation (2.4) is the identity matrix. When standard Rayleigh–Ritz is used in Step 1 of the algorithm, we call the method standard GD. With Interior Rayleigh–Ritz, it is called Interior GD. The first example demonstrates that preconditioning can make a difference for interior eigenvalue problems.

Example 4.
Let B be the same matrix as in Example 2. All of the versions of Arnoldi converged very slowly for this matrix. To reach the residual norm of 10^{-6} required over 1 000 iterations (25 runs out to dimension 50) for Interior Arnoldi. Now we use GD with diagonal preconditioning and with $\alpha = \sigma = 1.0$. For standard GD, the approximate eigenpair (ρ, y) is chosen as the Ritz pair with Ritz value nearest σ. For Interior GD, the (ρ, y) corresponding to the θ value closest to σ is used. The convergence test is $||r|| < 10^{-6}$. The standard GD method converges in 17 iterations, a tremendous improvement. The interior version converges in 16 iterations. So the interior method is not really needed in this case. The tremendous improvement brought by preconditioning is due to the structure of the matrix. The main diagonal is a good approximation of the whole matrix.

We next examine a more realistic preconditioning situation. The purpose here is not to give a thorough study of preconditioners for interior eigenvalues. Instead, we wish to demonstrate that some of the techniques that have been developed for preconditioning indefinite systems of linear equations can be applied to interior eigenvalue problems.

Example 5.
We consider a test matrix from the Harwell–Boeing Sparse Matrix Collection [10]. The matrix is SHERMAN5, which is from an oil reservoir simulation problem. It has dimension 3 312 and an average of six non-zero elements per row. It is well suited for preconditioning because the larger elements are on or near the main diagonal. The matrix is indefinite with many eigenvalues on both the negative and positive sides. We compute the three eigenvalues nearest zero: 0.047, 0.125, and 0.403. The next eigenvalue is 0.580 and the matrix has

© 1998 John Wiley & Sons, Ltd.

Table 1. SHERMAN5 matrix

	Interior GD iterations	Standard GD iterations
Arnoldi (no preconditioning)	5 000+	—
ILUT, NFILL=5	87	453
ILUT, NFILL=20	44	91

elements of magnitude near 1 000. With no preconditioning, this is a difficult problem. The three desired eigenvalues are on the interior and are not well separated compared with the rest of the spectrum. We attempt to compute these eigenvalues so that the residual norms are below 10^{-5}.

For a comparison to an approach without preconditioning, the Interior Arnoldi with Eigenvectors method is used (Arnoldi with standard Rayleigh–Ritz was not tested). The method is restarted when the subspace reaches dimension 30. Eight approximate eigenvectors are retained at each restart. We allow 250 runs or 5 508 iterations, but the residual norm for the third eigenvalue is still above 10^{-5}. It is at 0.18×10^{-4} but is barely improving.

Next we try an incomplete factorization preconditioner [24, 37]. Saad's ILUT [37] is used. This preconditioner monitors the size of prospective fill-in elements and keeps the larger ones. This is an attempt to have a fairly accurate factorization, while controlling expense and storage. With this tolerance check for the size of elements that are dropped during the factorization, the method can be more effective for indefinite problems [37]. The particular ILUT program used is from QMRPACK [16]. It has row storage, while the Harwell-Boeing matrices are stored by columns. So, we used the transpose of an incomplete factorization of B^{T} as the preconditioner. We let $\alpha = \sigma = 0$. So, no shifting is done before factoring. This preconditioner is never changed, so only one incomplete factorization is needed.

For standard GD, the approximate eigenpair (ρ, y) is chosen as the Ritz pair with Ritz value nearest σ. For Interior GD, the (ρ, y) corresponding to the θ value closest to σ is used. The maximal dimension of the subspace is 20, and five approximate eigenvectors are retained at the restart. The initial starting vector is the right-hand side vector given for SHERMAN5 in the Harwell–Boeing collection.

Two tests are done. The parameters in the ILUT program are first chosen as NFILL$=5$, TOL=10^{-3} and then allowing more fill-in, NFILL$=20$, TOL=10^{-3}. The effect of the preconditioning is remarkable; see Table 1. For Interior GD, only 44 iterations are required to find the three eigenvalues with the second preconditioner. With standard GD, 91 iterations are needed. Let us examine this more closely. Standard GD actually finds the first eigenvalue quicker than the interior method, 14 iterations to 19. It is not clear why this happens, but initial convergence for the GD method is more unpredictable than for Krylov subspace methods. For the next two eigenvalues, standard GD has frequent appearances of spurious Ritz values that detract in two different ways. For some iterations, the Ritz value closest to the origin (other than the one or two that have already converged) is spurious. With the corresponding spurious Ritz vector chosen for Step 2 of the GD method, that iteration is essentially wasted. At some other iterations, the best Ritz pair is the one closest to the origin and so, as we would wish, it is chosen as the Ritz vector for Step 2. But a nearby spurious Ritz value reduces the accuracy of this Ritz vector and there is little improvement for this iteration. We give examples of these two effects: at iteration 23, the spurious Ritz value

© 1998 John Wiley & Sons, Ltd.

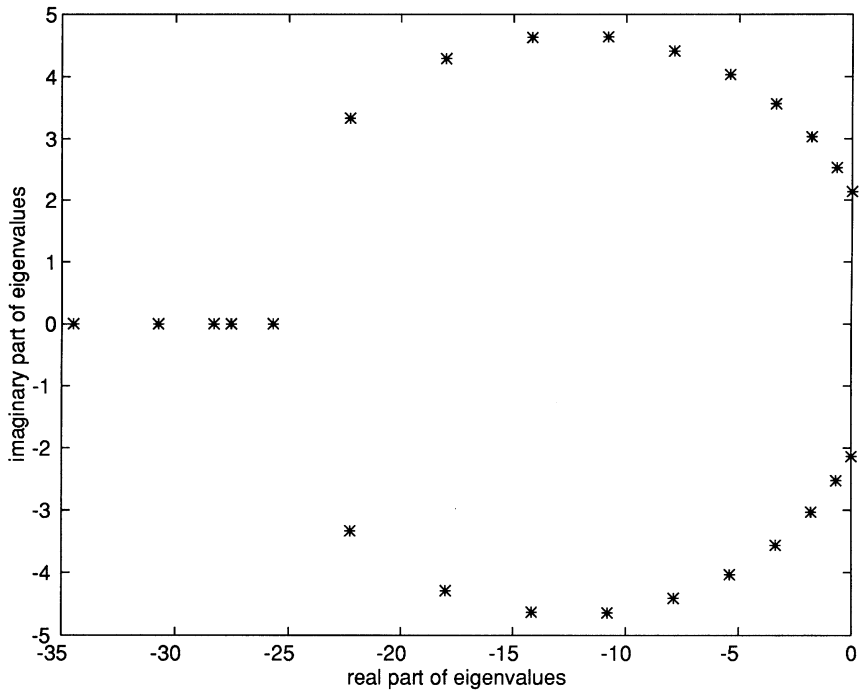

Figure 11. The 25 rightmost eigenvalues of BWM2000.

$0.1221 + 0.0013i$ (with less accurate Ritz vector) is chosen instead of the more accurate Ritz value $0.1259 - 0.0002i$. Then for iterations 24 to 28, the best Ritz value 0.1255 is chosen, but there are a couple of spurious values nearby. At step 24, they are $0.219 - 0.304i$ and $0.220 + 0.334i$, but they move around. There is no improvement during these iterations. So, even if the best Ritz vector could always be chosen, standard Rayleigh–Ritz would not always work well for the GD method, because it is important to have the best possible approximate eigenvector at every iteration.

Example 6.
Let the matrix be BWM2000 from the Test Matrix Collection for Non-Hermitian Eigenvalue Problems [3]. This matrix was generated from modeling a chemical reaction [38]. Figure 11 shows the 25 smallest eigenvalues. There are other eigenvalues past -10^5, so the small ones are not well separated relative to the whole spectrum. The elements of the starting vector are chosen randomly from the interval $(-1,1)$. For standard GD, the approximate eigenpair (ρ, y) is chosen as the Ritz pair with Ritz value nearest σ. For Interior GD, the (ρ, y) is the one corresponding to the θ value closest to σ, but with imaginary part greater than -1. This is so only one of a complex conjugate pair is computed. The convergence test is $||r|| < 10^{-5}$. The maximal dimension of the subspace is 25, and 10 Ritz vectors are retained at the restart.

For the preconditioner, we just consider ILUT with NFILL=60 and TOL=10^{-6}. Lesser values of NFILL were much less effective. We do three different comparisons. First, the five eigenvalues nearest to the origin are computed (counting a complex conjugate pair as one

© 1998 John Wiley & Sons, Ltd.

Table 2. BWM2000 matrix

	Interior GD iterations	Standard GD iterations
$\sigma = 0, \alpha = 0$	92	85
$\sigma = 0, \alpha = 5$	52	71
$\sigma = -20, \alpha = 5$	465	—

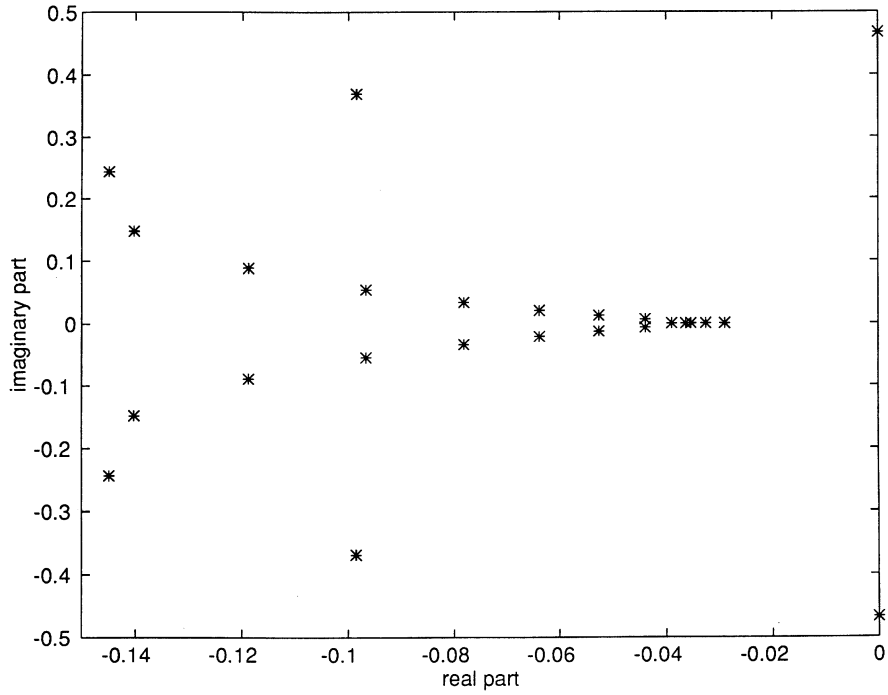

Figure 12. Reciprocals of 25 eigenvalues of BWM2000.

eigenvalue). Also, the matrix is not shifted before the incomplete factorization. So $\sigma = 0$ and $\alpha = 0$. The five eigenvalues are exterior, since they are on the edge of the spectrum. However, they can also be viewed as being interior eigenvalues because the smallest ones are somewhat between the ones further out. See Table 2 for the results. Standard GD finds the five eigenvalues in 85 iterations, which is slightly faster than Interior GD. This is in spite of the fact that standard GD has some trouble with spurious Ritz values. Spurious values interfere at 13 of the iterations, all during computation of the first two eigenvalues. It is interesting that spurious Ritz values appear even for these rightmost eigenvalues. Figure 12 plots the reciprocals of the 25 smallest eigenvalues. As these values are computed, starting with the one farthest from the origin, none of them are between values that are not yet computed. This shows why Interior Rayleigh–Ritz does not get spurious values here as does standard Rayleigh–Ritz.

Next the preconditioner is changed. This time the matrix is shifted. The preconditioner is the transpose of the ILUT factorization of $B^{\mathrm{T}} - 5I$, again with NFILL=60 and TOL=10^{-6}. This matrix has eigenvalues shifted to the left by five, so it is a little further from being

© 1998 John Wiley & Sons, Ltd.

indefinite. For this test, Interior GD converges faster than standard GD, 52 iterations compared with 71. It is not clear why this happens, since this time standard GD does not have problems with spurious Ritz values.

The last test is for computing eigenvalues further in the interior. The five eigenvalues nearest to -20 are sought. The same factorization of $B^{\mathrm{T}} - 5I$ is used. Interior GD finds the eigenvalues in 465 iterations. Standard GD does not find any in 500 iterations. There are many spurious Ritz values that interfere. This test shows the importance of using Interior Rayleigh–Ritz with preconditioning methods for interior eigenvalues.

Finally, we mention that a factorization of $B^{\mathrm{T}} + 20I$ was also tried for computing the eigenvalues near -20, but was not effective. This matrix is fairly indefinite, and other preconditioners should be tried. See [6, 38, 41] for other approaches, such as pivoting or incomplete QR factorization.

5. Conclusion

A modified version of the Rayleigh–Ritz procedure for non-symmetric eigenvalue problems has been discussed. It attempts to find estimates of eigenvalues near a specified value. Not only are better approximate eigenvalues (the ρ values) produced, but identification of which approximations are most significant is easy even at early stages by considering the harmonic θ values. The Interior Rayleigh–Ritz procedure is particularly important in the non-symmetric case, because restarting is generally used, and it is important to restart with the best possible eigenvector approximations.

If an approximation of the boundary of the spectrum is needed, approximate eigenvalues from this technique can be combined with standard Arnoldi Ritz values to give a better picture. Much work could be done on further developing this. A procedure for choosing the σ's would be desirable.

Krylov subspace methods, such as the Arnoldi method, often have difficulty computing interior eigenvalues, so it is important to also consider preconditioning methods. Preconditioners that have been developed for indefinite linear equations problems can be used, but further study is needed.

Also of interest is an interior version of the non-symmetric Lanczos algorithm. A two-sided interior Rayleigh–Ritz approach is possible (see [54]).

Acknowledgements

The authors would like to thank the referees for their many helpful suggestions, including the suggestion of making an analogy to SYMMLQ versus MINRES, as is done in Section 2. Also thanks to Valeria Simoncini and Gerard Sleijpen for helpful discussions. The research of the first author was supported by the National Science Foundation under contract CCR-9102221.

REFERENCES

1. S. F. Ashby, T. A. Manteuffel and P. E. Saylor. Adaptive polynomial preconditioning for Hermitian indefinite linear systems. *BIT*, 29, 583–609, 1989.
2. W. E. Arnoldi. The principle of minimized iteration in the solution of the matrix eigenvalue problem. *Quart. Appl. Math.*, 9, 17–29, 1951.

© 1998 John Wiley & Sons, Ltd.

3. Z. Bai, D. Day, J. Demmel and J. Dongarra. A test matrix collection for non-hermitian eigenvalue problems. LAPACK Working Note, 1996.
4. J. G. L. Booten, P. M. Meijer, H. J. J. te Riele and H. A. van der Vorst. Parallel Arnoldi method for the construction of a Krylov subspace basis: an application in magnetohydrodynamics. *Proceedings of High Performance Computing and Networking, Vol. 2; Lecture Notes in Computer Science 797*, Wolfgang Gentzsch and Uwe Harms, editors, Springer-Verlag, New York, NY, 1994, pp. 196–201.
5. D. Calvetti and L. Reichel. An adaptive Richardson iteration method for indefinite linear systems. *Numer. Algorithms*, 12, 125–150, 1996.
6. E. Chow and Y. Saad. Approximate inverse preconditioners for general sparse matrices. *SIAM J. Sci. Comput.*, 18, 1657–1675, 1997.
7. A. K. Cline, G. H. Golub and G. W. Platzman. Calculation of normal modes of oceans using a Lanczos method. J. R. Bunch and D. J. Rose, editors, *Sparse Matrix Computations*, pages 409–426, Academic Press, New York, NY, 1976.
8. B. N. Datta and Y. Saad. Arnoldi's methods for large Sylvester-like observer matrix equations, and an associated algorithm for partial spectrum assignment. *Linear Algebra Appl.*, 154–156, 225–244, 1991.
9. E. R. Davidson. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.*, 17, 87–94, 1975.
10. I. S. Duff, R. G. Grimes and J. G. Lewis. Sparse matrix test problems. *ACM Trans. Math. Soft.*, 15, 1–14, 1989.
11. T. Ericsson and A. Ruhe. The spectral transformation Lanczos method for the numerical solution of large sparse generalized symmetric eigenvalue problems. *Math. Comp.*, 35, 1251–1268, 1980.
12. B. Fischer. On polynomial iterative methods for indefinite linear systems. Presentation at The Copper Mountain Conference on Iterative Methods, 1990.
13. R. Freund. On polynomial preconditioning and asymptotic convergence factors for indefinite Hermitian matrice. *Linear Algebra Appl.*, 154–156, 259–288, 1991.
14. R. W. Freund. Quasi-kernel polynomials and their use in non-Hermitian matrix iterations. *J. Comput. Appl. Math.*, 43, 135–158, 1992.
15. R. W. Freund, G. H. Golub and N. M. Nachtigal. Iterative solution of linear systems. *Acta Numerica*, 1, 57–100, 1992.
16. R. W. Freund and N. M. Nachtigal. QMRPACK: a package of QMR algorithms. *ACM Trans. Math. Soft.*, 22, 46–77, 1996.
17. S. Goossens and D. Roose. Ritz and harmonic Ritz values and the convergence of FOM and GMRES. Report TW257, K.U. Leuven, 1997.
18. A. Jennings and W. J. Stewart. A simultaneous iteration algorithm for real matrices. *J. Math. Inst. Appl.*, 15, 351–361, 1980.
19. A. V. Knyazev. Convergence rate estimates for iterative methods for mesh symmetric eigenvalue problem. *Soviet J. Numer. Anal. and Math. Model.*, 2, 371–396, 1987.
20. A. V. Knyazev. The Temple–Lehmann methods in iterative algorithms. *Soviet J. Numer. Anal. and Math. Model.*, 5, 265–273, 1990.
21. C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bur. Stand.*, 45, 255–282, 1950.
22. T. A. Manteuffel and J. Otto. On the roots of orthogonal polynomials and residual polynomials associated with a conjugate gradient method. *Num. Lin. Alg. with Appl.*, 1, 449–475, 1994.
23. T. A. Manteuffel and G. Starke. On hybrid iterative methods for nonsymetric systems of linear equations. *Numer. Math.*, 73, 489–506, 1996.
24. J. A. Meijerink and H. A. van der Vorst. An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix. *Math. Comp.*, 31 , 148–162, 1977.
25. H. Mittelman, C. C. Law, D. F. Jankowski and G. P. Neitzel. A large, sparse, and indefinite generalized eigenvalue problem from fluid mechanics. *SIAM J. Sci. Statist. Comput.*, 13, 411–424, 1992.
26. R. B. Morgan. Computing interior eigenvalues of large matrices. *Linear Algebra Appl.*, 154–156, 289–309, 1991.
27. R. B. Morgan. Generalizations of Davidson's method for computing eigenvalues of large nonsymmetric matrices. *J. Comput. Phys.*, 101, 287–291, 1992.

© 1998 John Wiley & Sons, Ltd.

28. R. B. Morgan. On restarting the Arnoldi method for large non-symmetric eigenvalue problems. *Math. Comp.*, 65, 1213–1230, 1996.
29. R. B. Morgan and D. S. Scott. Preconditioning the Lanczos algorithm for sparse symmetric eigenvalue problems. *SIAM J. Sci. Comput.*, 14, 585–593, 1993.
30. R. B. Morgan. A restarted GMRES method augmented with eigenvectors. *SIAM J. Matrix Anal. Appl.*, 16, 1154–1171, 1995.
31. R. B. Morgan and D. S. Scott. Generalizations of Davidson's method for computing eigenvalues of sparse symmetric matrices. *SIAM J. Sci. Statist. Comput.*, 7, 817–825, 1986.
32. N. M. Nachtigal, L. Reichel and L. N. Trefethen. A hybrid GMRES algorithm for non-symmetric linear systems. *SIAM J. Matrix Anal. Appl.*, 13, 796–825, 1992.
33. C. C. Paige, B. N. Parlett and H. A. van der Vorst. Approximate solutions and eigenvalue bounds from Krylov subspaces. *Num. Lin. Alg. with Appl.*, 2, 115–133, 1995.
34. C. C. Paige and M. A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, 12, 617–629, 1975.
35. B. N. Parlett. *The Symmetric Eigenvalue Problem*. Prentice-Hall, Englewood Cliffs, NJ, 1980.
36. B. N. Parlett, D. R. Taylor and Z. A. Liu. A look-ahead Lanczos algorithm for unsymmetric matrices. *Math. Comp.*, 44, 105–124, 1985.
37. Y. Saad. ILUT: a dual threshold incomplete LU factorization. *Num. Lin. Alg. with Appl.*, 1, 387–402, 1994.
38. Y. Saad. *Iterative Methods for Sparse Linear Systems*. PWS Publishing, Boston, MA, 1996.
39. Y. Saad. *Numerical Methods for Large Eigenvalue Problems*. Halsted Press, New York, NY, 1992.
40. Y. Saad. Numerical solution of large non-symmetric eigenvalue problems. *Comput. Phys. Commun.*, 53, 71–90, 1989.
41. Y. Saad. Preconditioning techniques for non-symmetric and indefinite linear systems. *J. Comput. Appl. Math.*, 24, 89–105, 1988.
42. Y. Saad. Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices. *Linear Algebra Appl.*, 34, 269–295, 1980.
43. Y. Saad and M. H. Schultz. GMRES: a generalized minimal residual algorithm for solving non-symmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7, 856–869, 1986.
44. P. E. Saylor and D. C. Smolarski. Implementation of an adaptive algorithm for Richardson's method. *Linear Algebra Appl.*, 154–156, 615–646, 1991.
45. D. S. Scott. The advantages of inverted operators in Rayleigh–Ritz approximations. *SIAM J. Sci. Statist. Comput.*, 3, 68–75, 1982.
46. A. Sidi. Application of vector-valued rational approximations to the matrix eigenvalue problem and connections with Krylov subspace methods. *SIAM J. Matrix Anal. Appl.*, 16, 1341–1369, 1995.
47. V. Simoncini. Ritz and pseudo-Ritz values using matrix polynomials. *Linear Algebra Appl.*, 241–243, 787–801, 1996.
48. G. L. G. Sleijpen and H. A. van der Vorst. A Jacobi–Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 17, 401–425, 1996.
49. D. C. Sorensen. Implicit application of polynomial filters in a k-step Arnoldi method. *SIAM J. Matrix Anal. Appl.*, 13, 357–385, 1992.
50. A. Stathopoulos, Y. Saad and F. Fischer. Robust preconditioning of large, sparse, symmetric eigenvalue problems. *J. Comput. Appl. Math.*, 64, 197–215, 1995.
51. H. L. Tanaka and E. C. Kung. Normal mode energetics of the general circulation during FGGE winter. *J. Atmos. Sci.*, 45, 3723–3736, 1988.
52. L. N. Trefethen. Approximation theory and linear algebra. In J. C. Mason and M. G. Cox, editors, *Algorithms for Approximations, II*, pages 336–360, Chapman and Hall, New York, NY, 1990.
53. R. Varga. Algorithms for iteratively solving large non-symmetric matrix equations. Presentation at The Copper Mountain Conference on Iterative Methods, 1990.
54. M. Zeng. *Finding interior eigenvalues of large non-symmetric matrices*. Ph.D. thesis, University of Missouri, Columbia, MO, 1996.

© 1998 John Wiley & Sons, Ltd.