CHAPTER V

Krylov Subspaces

14 Construction of a basis

With the Power method (see 8), we have generated the spanning vectors of a *Krylov Subspace*

$$\mathcal{K}^m(A; v_1) := \text{span}\{v_1, Av_1, \ldots, A^{m-1}v_1\}.$$

However, note that the Power method exploits only the two most recently computed vectors for the computation of an approximating eigenvalue. The result is that the speed of convergence depends on how fast the values of the polynomial $P_j(t) = (t/\lambda_1)^j$ decrease in absolute value for $t = \lambda_2$, for increasing j. Here we have assumed that the eigenvalues have been ordered as $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$. Of course, one may try to influence the convergence of the Power Method by working with shifts, either to separate the wanted eigenvalue more from the remaining spectrum, or to suppress eigenvector components of nearby eigenvalues. For this one needs, in fact, good guesses for these nearby eigenvalues, and, as we will see, the methods of Lanczos and Arnoldi do this automatically. These methods exploit the whole Krylov subspace, and they implicitly construct polynomials P_j that suppress the other eigenvalues in an optimal way.

From the definition of a Krylov subspace, it is clear that

$$\mathcal{K}^m(\alpha A + \beta I; v_1) = \mathcal{K}^m(A; v_1) \quad \text{for} \quad \alpha \neq 0. \tag{14.1}$$

This says that the Krylov subspace is spanned by the same basis if A is scaled and/or shifted. The implication is that Krylov subspace methods for the spectrum of the matrix A are invariant under translations for A.

Krylov subspaces play a central role in iterative methods for eigenvalue computations. In order to identify better solutions in the Krylov subspace we need a suitable basis for this subspace, one that can be extended inductively for subspaces of increasing dimension. The obvious basis v_1, Av_1,

..., $A^{m-1}v_1$ for $K^m(A;v_1)$ is not very attractive from a numerical point of view, since the vectors $A^j v_1$ point more and more in the direction of the dominant eigenvector for increasing j (the power method!), and hence the basis vectors become dependent in finite precision arithmetic.

Lanczos [1950] proposes to generate an orthogonal basis for the Krylov subspace, and shows that this could be done in a very economic way for symmetric A. Arnoldi [1951] describes a procedure for the computation of an orthonormal basis for unsymmetric matrices and for ease of presentation, we start with this one. We will see that Lanczos' algorithm follows as a special case.

If we have already an orthonormal basis $v_1, \ldots, v_j$ for $\mathcal{K}^j(A;v_1)$, then this basis is expanded by computing $\widetilde{v} = Av_j$, and by orthonormalizing this vector $\widetilde{v}$ with respect to $v_1, \ldots, v_j$. In principle the orthonormalization process can be carried out in different ways, but the most commonly used approach is to do this by a modified Gram-Schmidt procedure (Golub and Van Loan [1996]).

This leads to the algorithm for the creation of an orthonormal basis for $\mathcal{K}^m(A;v)$, as given in Fig. 5.

v is a convenient starting vector
$v_1 = v/\|v\|_2$
for $j = 1, .., m-1$
    $t = Av_j$
    for $i = 1, ..., j$
        $h_{i,j} = v_i^* t$
        $t = t - h_{i,j} v_i$
    end
    $h_{j+1,j} = \|t\|_2$
    $v_{j+1} = t/h_{j+1,j}$
end

Figure 5: The Arnoldi Method with Modified Gram-Schmidt

A more stable implementation for the construction of an orthonormal basis, useful for ill-conditioned matrices A, was suggested in Walker [1988]. Walker suggested to use Householder reflections rather than modified Gram-Schmidt. An alternative is to do two iterations with (modified) Gram-

Schmidt if necessary. This works as follows. If one wants to have a set of orthogonal vectors to almost working precision then one has to check, after the orthogonalization of a new vector with respect to the existing set, whether the remaining vector is significantly smaller in norm than before orthogonalization, say more than $\kappa < 1$ smaller. In that case we may have had cancellation effects, and we apply modified Gram-Schmidt once again. This is the basis for the refinement technique suggested in Daniel, Gragg *et al* [1976]. It leads to a set of vectors for which the mutual loss of orthogonality is limited to $1/\kappa$, in a relative sense. In the template in Fig. 6, we have included this technique in the Arnoldi algorithm.

> v is a convenient starting vector
> Select a value for κ, e.g., $\kappa = .25$
> $v_1 = v/||v||_2$
> **for** $j = 1, .., m-1$
> $t = Av_j$
> $\tau_{in} = ||t||_2$
>   **for** $i = 1, ..., j$
> $h_{i,j} = v_i^* t$
> $t = t - h_{i,j} v_i$
>   **end**
>   **if** $||t||_2/\tau_{in} \leq \kappa$
>   **for** $i = 1, ..., j$
> $\rho = v_i^* t$
> $t = t - \rho v_i$
> $h_{i,j} = h_{i,j} + \rho$
>   **end**
>   **endif**
> $h_{j+1,j} = ||t||_2$
> $v_{j+1} = t/h_{j+1,j}$
> end

Figure 6: The Arnoldi Method with refined Modified Gram-Schmidt

Note that, in exact arithmetic, the constants ρ in Fig. 6 are equal to zero. It is easily verified that, in exact arithmetic, the v_1, ..., v_m form an orthonormal basis for $\mathcal{K}^m(A; v)$ (that is, if the construction does not terminate at a vector $t = 0$).

The orthogonalization leads to relations between the v_j, that can be formulated in a compact algebraic form. Let V_j denote the matrix with columns v_1 up to v_j, $V_j := [v_1 \mid v_2 \mid \cdots \mid v_j]$, then it follows that

$$AV_{m-1} = V_m H_{m,m-1}. \tag{14.2}$$

The m by $m-1$ matrix $H_{m,m-1}$ is upper Hessenberg, and its elements $h_{i,j}$ are defined by the Arnoldi orthogonalization algorithm. We will refer to this matrix $H_{m,m-1}$ as the *reduced matrix* A, or more precisely, the matrix A reduced to the current Krylov subspace. From a computational point of view, this construction is composed from three basic elements: a matrix vector product with A, inner products, and updates. We see that this orthogonalization becomes increasingly expensive for increasing dimension of the subspace, since the computation of each $h_{i,j}$ requires an inner product and a vector update.

15 The Ritz values and Ritz vectors

The eigenvalues of the leading m by m part $H_{m,m}$ of the matrix $H_{m+1,m}$ are called the *Ritz values* of A with respect to the Krylov subspace of dimension m. We will see later, in our discussions on the Lanczos and Arnoldi methods, how they can be related to eigenvalues of A. If s is an eigenvector of $H_{m,m}$, then $y = V_m s$ is called a *Ritz vector* of A. Also these Ritz vectors can be related to eigenvectors of A.

All this is not so surprising, because of the relation with the Krylov subspace and the vectors generated in the Power Method. It seems obvious that we can base our expectations of the convergence of the Ritz values on our knowledge of convergence for the Power Method, something like, the largest Ritz value will converge faster to the largest eigenvalue of A than the Rayleigh quotient in the Power Method. But this kind of attitude is not very helpful for deeper insight. There is a very essential difference between Krylov subspace methods and the Power method and that is that the Krylov subspace is invariant for scaling of A and for shifting of A, that is, A and $\alpha A + \beta I$ generate exactly the same Krylov subspace. The reduced matrix $H_{m,m}$ is simply scaled and shifted to the reduced matrix $\widetilde{H}_{m,m} = \alpha H_{m,m} + \beta I_m$, asociated with $\widetilde{A} = \alpha A + \beta I$. With I_m we denote the m by m identity matrix.

This implies that for the Krylov subspace method the notion of largest (in absolute) value eigenvalue looses its special meaning as opposed to the

Power Method. Since the Krylov methods are shift invariant, the position of the origin in the spectrum is not relevant and we rather make the distinction between *exterior* and *interior* eigenvalues. The Krylov methods have no special preference for eigenvalues that are at about the same distance of the center of the spectrum, provided that these eigenvalues are about equally well separated from the others. In particular, when the real spectrum of a symmetric real matrix is symmetrically distributed with respect to $(\lambda_1 + \lambda_n)/2$, λ_1 being the largest eigenvalue of A and λ_n the smallest one, then for a starting vector that has also a symmetric weight distribution with respect to the corresponding eigenvectors, the convergence of the smallest Ritz value towards λ_n will be equally fast (or slow) as the convergence of the largest Ritz value to λ_1.

For complex spectra, one has to consider the smallest circle that encloses all eigenvalues. With proper assumptions about the starting vector, one may expect that the eigenvalues close to this circle will be approximated fastest and that the more interior eigenvalues will be approximated later in the Krylov process (that is, for larger values of m). For more general starting vectors this describes more or less the generic case, but with special starting vectors one can force convergence towards favoured parts of the spectrum. This forms the basis for restart strategies; we will discuss such strategies for the respective methods.

For an excellent overview of subspace methods, see SAAD [1992].

16 Ritz values and harmonic Ritz values

Once we have a basis for the Krylov subspace there are several possibilities to exploit this for eigenvalue approximations. For instance, when A is symmetric, then one can maximize the Rayleigh quotient x^*Ax/x^*x over the current subspace (and of subspaces thereof) in order to approximate eigenvalues (cf. Theorem 2). For symmetric matrices, these maxima are equal to the Ritz values.

One can also determine the minimum of x^*Ax/x^*A^2x over the Krylov subspace and convenient subspaces of it. Note that this would lead to approximations for eigenvalues of A^{-1} for vectors $y = Ax$, with x in the Krylov subspace. The inverses of these approximations are referred to as the *Harmonic Ritz values*. They are associated with *Harmonic Ritz vectors*, that is, the vectors for which the minima of the Rayleigh quotients are attained.

The relations between Ritz values and harmonic Ritz values and their computation, for symmetric matrices, are discussed in PAIGE, PARLETT and VAN DER VORST [1995]. See also BEATTIE [1998] for an enlightning discussion on the relations between Ritz values, harmonic Ritz values, and even *dual Harmonic Ritz values.* For the latter values one has to carry out one single step of inverse iteration (and A has to be positive definite), as they define the local minima of Rayleigh quotients $x^*x/x^*A^{-1}x$ over the Krylov subspace, generated with A. Together, these various Ritz values provide useful bounds on eigenvalues of A.

We will discuss how to compute the Ritz values and the harmonic Ritz values for the different methods in the respective chapters. The general principle that underlies the computation of these approximations for eigenvalues and eigenvectors, is that the entire computation can be carried out with information obtained for the Krylov subspace. That is, it can be done in low-dimensional subspaces that replace the high-dimensional space over which the given eigenproblem is formulated. The computations in low-dimensional subspaces can usually be carried out with relatively much lower costs than those involved with methods that attempt to solve the given eigenproblem directly. As an example of this principle, we will discuss in the next section the evaluation of the so-called refined eigenvectors.

17 Refined eigenvector approximations

The notion of *refined eigenvector approximations* has been introduced in JIA [1997], JIA [1998], and JIA [1999]. Jia suggests, for a given Ritz value θ, to minimize

$$\|Az - \theta z\|_2, \tag{17.1}$$

for normalized z from the given Krylov subspace. They are referred to as *refined Ritz vectors* in the original publications, but we prefer our terminology, because the refined vectors do not satisfy a Ritz(-Galerkin) condition. They lead, in general, to better eigenvector approximations, see the analysis in JIA and STEWART [1999].

The actual computation of the refined eigenvector approximations in the context of Krylov subspaces proceeds as follows. Suppose that $v_1, \ldots,$ v_m form an orthormal basis for $\mathcal{K}^m(A; v_1)$, and let V_m denote the n by m matrix with columns v_j. Then any vector z in this Krylov subspace can be expressed as $z = V_m y$, with $y \in \mathbb{R}^m$. Now we want to minimize (17.1) for z with $\|z\|_2 = 1$. Since V_m is orthonormal, this implies that $\|y\|_2 = 1$ a well, so that we have to determine

$$\min_{\|y\|_2=1} \|(AV_m - \theta V_m)y\|_2. \tag{17.2}$$

Clearly the vector y that realizes the minimum is the right singular vector of the matrix $(A - \theta I)V_m$, corresponding to the smallest singular value of that matrix (see Chapter 4). This vector can be rather easily computed from the available Krylov subspace information.

Equation (17.2) can be further simplified as

$$\begin{aligned} \|(AV_m - \theta V_m)y\|_2 &= \|(V_{m+1}H_{m+1,m} - \theta V_m)y\|_2 \\ &= \|V_{m+1}(H_{m+1,m} - \theta I_{m+1,m})y\|_2 \\ &= \|(H_{m+1,m} - \theta I_{m+1,m})y\|_2, \end{aligned} \tag{17.3}$$

so that y is simply the right singular vector corresponding to the smallest singular value of $H_{m+1,m} - \theta I_{m+1,m}$.

This leads to the refined eigenvector approximation $z_r = V_m y_s$ (corresponding to the Ritz value θ). In JIA and STEWART [1999] it has been shown that under rather natural assumptions the refined eigenpair θ, z_r converges to an eigenpair of the matrix A. There is no assumption necessary that A has only distinct eigenvalues or that A is diagonalizable.

We will also discuss the computation of refined eigenvector approximations in the context of the different methods in the respective chapters.

18 Block Krylov Methods

In Section 9 we have discussed the Simultaneous Iteration Method, as a useful extension for the Power Method. The Simultaneous Iteration Method was motivated by the expected faster convergence for clusters of eigenvalues. It also offers possibilities for efficient computation when the access of the matrix A is relatively expensive. The matrix A can then be used for a set of vectors, so that the costs of access are amortized somehow. Although obvious, the block variants of the Krylov subspace methods have not received much attention in litterature.

The main reason for this lack of attention may be that the Arnoldi method often has to be restarted in order to keep memory and computational overhead modest. For a restart, one has to select the best candidate vectors from the current Krylov subspace and it is quite straightforward to enhance such strategies with deflation in order to get convergence for (almost) multiple eigenvalues. We will see such strategies for Arnoldi's method in Chapter 31.

For the Hermitian matrices, the overhead in generating an orthogonal basis for the Krylov subspace is limited and there is no need for restart. This may explain that most attention for block methods has been devoted to the

Hermitian case. The first proposal for a Block (Lanczos) Krylov method was made in GOLUB and UNDERWOOD [1977]. RUHE [1979] suggested an alternative implementation for the block Lanczos method, that leads to more data locality, so that more operations can be carried out with data in local memory.

A careful implementation of Block Arnoldi has been discussed in BAGLAMA *et al* [1998], together with an interesting application. Saad gives a few possible versions of Block Arnoldi, including a block modified Gram-Schmidt version and a generalized version of Ruhe's approach: SAAD [1992, Chapter VI.5]. See also LEHOUCQ and MASCHOFF [2000], for a detailed discussion on a particular implementation of the Block Arnoldi Method.

We include here the Modified Gram-Schmidt variant published SAAD [1992, Chapter VI.5], because this variant shows very clearly the similarity between the single vector variant, described in Figure 5. The resulting Block Arnoldi Method is shown in Figure 7.

```
V_1 is a n by k unitary starting matrix
for j = 1,..,m-1
    W = A V_j
    for i = 1,...,j
        H_{i,j} = V_i^* W
        W = W - V_i H_{i,j}
    end
    QR Factor W: W = V_{j+1} H_{j+1,j}
end
```

Figure 7: The Block Arnoldi Method with Modified Gram-Schmidt

The starting matrix V_1 represents the set of k orthonormal starting vectors. The Block Arnoldi Method leads to an orthogonal basis for the block Krylov subspace

$$K^m(A;V_1) := \{V_1, AV_1, \ldots, A^{m-1}V_1\}.$$

It leads to the following analogue for (14.2)

$$AU_{m-1} = U_m H_{[m,m-1]}, \tag{18.1}$$

where $U_j := [V_1, V_2, \ldots, V_j]$, and $H_{[m,m-1]}$ is the block upper Hessenberg matrix of which the (i,j) block $h_{[i,j]}$ is defined as

$$h_{[i,j]} = H_{i,j}.$$

The further usage of this reduced upper Hessenberg matrix for eigenvalue computations is similar to the single vector case.

CHAPTER VI

The Lanczos Method

19 Reduction to tridiagonal form

Note that if A is symmetric, then so is

$$H_{m-1,m-1} = V_{m-1}^{*} A V_{m-1},$$

so that in this situation $H_{m-1,m-1}$ is tridiagonal:

$$AV_{m-1} = V_m T_{m,m-1}. \tag{19.1}$$

The matrix $T_{m,m-1}$ is an m by $m-1$ tridiagonal matrix, its leading $m-1$ by $m-1$ part is symmetric.

This means that in the orthogonalization process, each new vector has to be orthogonalized with respect to the previous two vectors only, since all other inner products vanish. The resulting three term recurrence relation for the basis vectors of $\mathcal{K}^m(A; v_1)$ is the kernel of the *Lanczos method* and some very elegant methods are derived from it. A template for the Lanczos method is given in Algorithm 8.

In the symmetric case the orthogonalization process involves constant arithmetical costs per iteration step: one matrix vector product, two inner products, and two vector updates. In exact arithmetic, this process must terminate after at most n steps, since then the n orthonormal vectors for the Krylov subspace span the whole space. In fact, the process terminates after k steps, if the starting vector has components only in the directions of eigenvectors corresponding to k *different* eigenvalues. In the case of finite termination, we have reduced the matrix A to triangular form with respect to an invariant subspace, and in fact, Lanczos' algorithm was initially viewed as a finite reduction algorithm for A.

Some years later it was recognized that this reduction could be done much better by techniques described by Givens and Householder, and Lanczos' method was discarded, because of its seemingly inferior behavior in

v is a convenient starting vector
$v_1 = v/||v||_2, \; v_0 = 0, \; \beta_0 = 0,$
for $j = 1, 2, ..., m-1$
  $t = Av_j - \beta_{j-1} v_{j-1}$
  $\alpha_j = v_j^* t$
  $t = t - \alpha_j v_j$
  $\beta_j = ||t||_2$
  $v_{j+1} = t/\beta_j$
end

Figure 8: The Lanczos algorithm

finite precision arithmetic, as a direct reduction algorithm for the reduction of A to tridiagonal form. PAIGE [1971] studies the method as an iterative technique and shows that it can be quite effective and lead to very accurate eigenvalue and eigenvector approximations. Before we discuss this in more detail, we will first consider the relation between eigenvalues of $T_{m,m} = V_m^* A V_m$ and those of A.

20 The Ritz values and Ritz vectors

According to the Minimax Theorem, see Theorem 2, we can characterize approximations θ_k, s_k, with $\theta_m \le \theta_{m-1} \le \ldots \le \theta_1$ from the Krylov subspace $\mathcal{K}^m(A; v_1)$, for eigenpairs of A, as

$$\theta_k = \max_{dim(S)=k} \left\{ \min_{0 \ne x \in S \subset \mathcal{K}^m(A;v_1)} \frac{x^* A x}{x^* x} \right\}, \qquad (20.1)$$

for $k = 1, 2, \ldots, m$. The x for which the minimum is taken, defines the approximate eigenvector s_k.

Because an $x \in \mathcal{K}^m(A; v_1)$ can be expressed as $V_m y$, in a unique way (V_m maps $y \in \mathbb{R}^m$ to $x = V_m y \in \mathcal{K}^m(A; v_1) \subset \mathbb{R}^n$), we have that

$$\frac{x^* A x}{x^* x} = \frac{y^* V_m^* A V_m y}{y^* V_m^* V_m y} = \frac{y^* T_{m,m} y}{y^* y}.$$

This shows that the θ_j are just the m eigenvalues of $T_{m,m}$, and if z_j is the corresponding eigenvector of $T_{m,m}$, then $s_j = V_m z_j$.

There is another characterization for the approximate eigenpairs of A. If we try to find an approximate eigenpair θ, s of A, such that its residual is orthogonal to the Krylov subspace (spanned by the columns of V_m), then clearly

$$V_m^*(As - \theta s) = 0, \tag{20.2}$$

or, since $s = V_m z$

$$T_{m,m} z - \theta z = 0, \tag{20.3}$$

and we see that the eigenpairs of $T_{m,m}$ also satisfy the Rayleigh-Ritz conditions, given in (20.2), for A with respect to the Krylov subspace (spanned by the columns of V_m). For this reason the values θ_j are called the *Ritz values* and the s_j are the corresponding *Ritz vectors* of A with respect to the given Krylov subspace.

The minimax property leads to the interesting observation that the ordered Ritz values are monotonically increasing to the ordered eigenvalues of A, that is, the largest Ritz value, for increasing m, increases monotonically to the largest eigenvalue of A. Since, in exact arithmetic, the Lanczos process is a finite process, we can strictly speaking, not regard the sequence of largest Ritz values as a converging process. However, in finite precision, the process will usually deliver accurate approximations that cannot be essentially improved, long before the finiteness of the process would play a role. In this sense, it behaves as a converging process that is terminated when finite precision effects inhibit further convergence. We will therefore, in the context of Ritz values and Ritz vectors speak of convergence, and we will consider an eigenpair as (practically) converged when its residual is in norm smaller than some given tolerance.

21 The polynomial characterization

There is still another characterization of the Ritz values. For this we return to the basic relation:

$$AV_{m-1} = V_m T_{m,m-1}.$$

If we denote the nonzeros in the i-th column of the tridiagonal matrix $T_{m,m-1}$ by $\beta_{i-1}, \alpha_i, \beta_i$, then we can also write down the famous three term recurrence relation for the basis vectors of the Krylov subspace:

$$\beta_i v_{i+1} = Av_i - \alpha_i v_i - \beta_{i-1} v_{i-1}. \tag{21.1}$$

Of course, any vector $v \in \mathcal{K}^i(A; v_1)$ can be expressed as a polynomial, of degree $i-1$ in A, acting on v_1. From equation (21.1) we conclude that the

v_i can be generated with polynomials $P_{i-1}(A)$:

$$v_i = P_{i-1}(A)v_1,$$

that satisfy the 3-term recurrence relation

$$\beta_i P_i(A) = (A - \alpha_i I)P_{i-1}(A) - \beta_{i-1}P_{i-2}(A). \tag{21.2}$$

The polynomial P_i is called a Lanczos polynomial.

This recurrence relation can be rewritten for scaled polynomials $P_i(A) = \mu_i \widetilde{P}_i(A)$, and by selecting the μ_i as $\mu_i = \mu_{i-1}/\beta_i$, with $\mu_0 = 1$, the recurrence for $\widetilde{P}$ becomes

$$\widetilde{P}_i(A) = (A - \alpha_i I)\widetilde{P}_{i-1}(A) - \beta_{i-1}^2 \widetilde{P}_{i-2}(A). \tag{21.3}$$

We see that this is exactly the same recurrence as for the determinants of the successive leading minors of $tI - T_{m,m}$, which shows that the zeros of $\widetilde{P}_i(t)$, and hence those of $P_i(t)$, are the eigenvalues of T_i.

Obviously, the polynomials $P_i(t)$ are orthogonal polynomials with respect to the inner product

$$[P_i, P_j] := (P_i(A)v_1, P_j(A)v_1).$$

If we write $A = Q^*\Lambda Q$, for the eigendecomposition of A, then clearly

$$[P_i, P_j] = (P_i(Q^*\Lambda Q)v_1, P_j(Q^*\Lambda Q)v_1) = (P_i(\Lambda)Qv_1, P_j(\Lambda)Qv_1). \tag{21.4}$$

This shows that the orthogonal polynomials P_i are fully characterized by the eigenvalues of A and by the vector Qv_1, that is the components of v_1 with respect to the eigenvector basis. We can now use the full theory of orthogonal polynomials for our analysis of the behavior of the Ritz values. We see that the method is invariant under orthogonal similarity transformations (at least in exact arithmetic), so that in our study of the convergence behavior of the Lanczos method it is no loss of generality if we carry out numerical experiments with diagonal matrices.

The inner product, with respect to which the polynomials are orthogonal, is a discrete inner product. As we will see, this has a consequence for the maximum degree of the associated orthogonal polynomials. Write the orthonormal set of eigenvectors of A as x_1, ..., x_n, with eigenvalues λ_1, ..., λ_n. Furthermore, we write v_1 with respect to this eigenvector basis as $v = \sum_k \gamma_k x_k$. Then, clearly

$$[P_i, P_j] = \sum_k P_i(\lambda_k)P_j(\lambda_k)\gamma_k^2,$$

and we conclude that we can take terms with equal λ_k together, so that the inner product is only determined by the different eigenvalues of A, for which $\gamma_k \neq 0$. This is equivalent with the observation, that multiplicity of an eigenvalue of A cannot be determined by the Lanczos method. Without loss of generality we will assume that the eigenvalues in the summation are different and that the corresponding weights γ_k are nonzero. Because the inner product is defined by the nonzero weights, we also conclude that the Lanczos method terminates when the degree of the polynomial exceeds the number of nonzero weights minus 1 (also a well-known fact). Moreover, from the theory of orthogonal polynomials we conclude that the two sets of Ritz values, corresponding to two successive iterations, interlace (SZEGÖ [1967]).

The interlacing is strict, because if two successive orthogonal polynomials have a zero in common, then, because of the 3-term recurrence relation, all higher degree orthogonal polynomials would have that zero. In particular $P_n(t)$ has that zero, and since this polynomial is proportional to the characteristic polynomial of A, this zero is an eigenvalue of A. This means that we have an invariant subspace, and the Lanczos method has terminated. An easy consequence of the interlacing property is that the largest (smallest) Ritz values of successive iteration steps form a strictly increasing (decreasing) sequence, with maximum (minimum) value the largest (smallest) eigenvalue for which $\gamma_k \neq 0$. A similar statement holds for the second largest (smallest) Ritz values of successive iterations. In general, the ℓ-th Ritz values of successive iterations form a strictly increasing sequence with the ℓ-th eigenvalue of A as its minimum. This minimum is attained within at most n iteration steps (in exact arithmetic).

The polynomial characterization also permits nice error bounds. In order to show this, we write $P_i(A)v_1$ as $\sum_j (\lambda_j - \theta_1^{(i)}) R_{i,1}(\lambda_j)\gamma_j x_j$ (remember that $P_i(t)$ has i zeros: the Ritz values $\theta_1^{(i)} > \theta_2^{(i)} > \ldots > \theta_i^{(i)}$), and because the i-th polynomial is orthogonal to all lower degree polynomials (and their combinations), we have that

$$0 = [P_i, R_{i,1}] = \sum_j (\lambda_j - \theta_1^{(i)}) R_{i,1}^2(\lambda_j)\gamma_j^2.$$

This gives the following relation for $\theta_1^{(i)}$:

$$\theta_1^{(i)} = \frac{\sum_j \lambda_j R_{i,1}^2(\lambda_j)\gamma_j^2}{\sum_j R_{i,1}^2(\lambda_j)\gamma_j^2}. \tag{21.5}$$

If we rewrite this in matrix notation, then we obtain

$$\theta_1^{(i)} = \frac{z^* A z}{z^* z},$$

with $z = R_{i,1}(A)v_1$. By combining this with our knowledge that the Ritz values are the (local) extrema of the Rayleigh quotient, we conclude that this gives us two nice results:

1. The Ritz vector corresponding to $\theta_1^{(i)}$ is given by $z = R_{i,1}(A)v_1$, where $R_{i,k}(A)$ is the reduced Lanczos iteration polynomial P_i, that is, the polynomial from which the linear factor $(A - \theta_k^{(i)} I)$ has been removed.
2. Because of the minimax property, the above characterization gives us an excellent means to find *a priori* lower bounds for $\theta_1^{(i)}$, namely, any polynomial of degree $i-1$ may replace $R_{i,k}$ and this will give a lower bound. We will consider this later in some more detail.

We now return to the expression given in (21.5) for $\theta_1^{(i)}$. Note that this expression is scaling invariant with respect to the polynomial $R_{i,1}$, and we scale it such that $\widetilde{R}_{i,1}(\lambda_1) = 1/\gamma_1$, that is we replace $R_{i,1}$ by

$$\widetilde{R}_{i,1}(\lambda) = R_{i,1}(\lambda)/(R_{i,1}(\lambda_1)\gamma_1).$$

Then we obtain the following expression for the distance of the largest Ritz value $\theta_1^{(i)}$ to the largest eigenvalue λ_1:

$$\lambda_1 - \theta_1^{(i)} = \frac{\sum_{j>1}(\lambda_1 - \lambda_j)\widetilde{R}_{i,1}^2(\lambda_j)\gamma_j^2}{1 + \sum_{j>1}\widetilde{R}_{i,1}^2(\lambda_j)\gamma_j^2}, \tag{21.6}$$

and obviously, because of the minimax property, this says that this distance is minimized over all polynomials ϕ of degree $i-1$, with the property that $\phi(\lambda_k) = 1$. As we will show now, the error in the Ritz value can be interpreted as the square of the norm of the error of some associated Conjugate Gradient process. In the Conjugate Gradient method for $Ay = b$, with starting vector $y_0 = 0$, the error in A norm, that is $(y - y_i, A(y - y_i))$ is mimimized over all $y_i = Q_{i-1}(A)b$. With Q_{i-1} we denote a polynomial of degree $i-1$. Hence, we see that $y - y_i = (I - Q_{i-1}(A)A)y$, so that with $G_i := I - AQ_{i-1}(A)$,

$$\|y - y_i\|_A = \|G_i(A)y\|_A$$

is minimized over all polynomials G_i of degree i, for which $G_i(0) = 1$.

Comparing this with (21.6), we see that the numerator in the right-hand of (21.6) can be interpreted as the error, in B norm, of a CG-process for a matrix B of order $n-1$ with $n-1$ eigenvalues $\lambda_1 - \lambda_j$ ($j > 1$), eigenvectors x_j, and with a solution y that can be expressed as

$$y = \sum_{j \geq 2} \gamma_j x_j.$$

Let Q_{i-1} denote the polynomial that minimizes the B-norm of this CG-process in step $i-1$, then clearly it leads to an upper bound for the approximation error in the largest Ritz value:

$$\begin{aligned} \lambda_1 - \theta_1^{(i)} &\leq \frac{\sum_{j>1} (\lambda_1 - \lambda_j) Q_{i-1}^2(\lambda_j)\gamma_j^2}{1 + \sum_{j>1} Q_{i-1}^2(\lambda_j)\gamma_j^2} \\ &\leq \sum_{j>1} (\lambda_1 - \lambda_j) Q_{i-1}^2(\lambda_j)\gamma_j^2 \\ &= \|x_{i-1} - x\|_B^2. \end{aligned} \tag{21.7}$$

This shows that $\lambda_1 - \theta_1^{(i)}$ is bounded from above by the square of the B-norm of the error in a related Conjugate Gradient process. As soon as $\theta_1^{(i)}$ has converged to some modest degree, then (21.6) implies that

$$1 + \sum_{j>1} \widetilde{R}_{i,1}^2(\lambda_j)\gamma_j^2 \approx 1,$$

and then the approximation error almost coincides with the CG-error. Or said in a slightly different way, the same kind (only marginally better) of convergence behavior that we observe for Conjugate Gradients (with a reduced and shifted spectrum) can be observed for the convergence behavior of the first Ritz value.

Numerical Experiment We consider the diagonal matrix of order 400 with eigenvalues $\lambda_1 = -0.034$, $\lambda_2 = -0.081$, $\lambda_3 = -0.0811$, $\lambda_4 = -0.127$, $\lambda_5 = -0.155$, and $\lambda_6 = -0.190$. The other 394 eigenvalues of this matrix are uniformly distributed in $[-1.2, -0.2]$. We select a v_1 that has equal components in all eigenvector directions: $\gamma_j = 1$. In Fig. 9 we see the values of the error $\theta_1^{(i)} - \lambda_1$, and the (squares of the) errors in B-norm of the associated CG-process. We see that the Ritz errors are nicely below the CG-errors, but ultimately they are close.

We also see that temporary slow-downs in the convergence behaviour of CG is reflected by a similar slowdown of the Ritz value, although not at

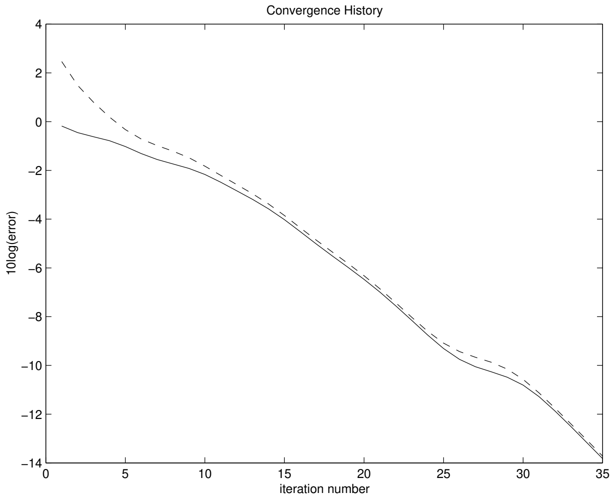

Figure 9: Ritz error (-line-) and CG-error (-dashed-)

the same iteration numbers and at the same level. Note that the associated CG process is NOT the CG process that defines the same Ritz values as the Lanczos process that we have used.

In VAN DER SLUIS and VAN DER VORST [1986], it was shown that the Conjugate Gradient process converges superlinearly, or more precisely, that Conjugate Gradients converges as a process from which the smallest eigenvalue has been removed, as soon as this smallest eigenvalue is modestly well approximated by the Ritz value of the corresponding Lanczos process (that is the Lanczos process defined by the iteration coefficients of CG).

In a similar way, we have superlinear convergence for the first Ritz value, as soon as the smallest eigenvalue of the shifted reduced problem is modestly well approximated by a Ritz value. Since the Ritz values are shift invariant, this implies that we may expect faster convergence in average, for the first Ritz value, as soon as the second Ritz value gets modestly close to the second eigenvalue.

Proceeding in this way, we see that we may expect a reduction of the average speed of convergence, each time when the next smallest eigenvalue gets a Ritz value in its neighbourhood. This explains nicely what we see in practice.

This interpretation also shows that the error can be bounded as an error in a CG-process, and for this upper bounds have been derived that decay with a linear factor that depends on the condition number of B, or, in our case $(\lambda_n - \lambda_1)/(\lambda_2 - \lambda_1)$.

The similarity between CG errors and Ritz errors had already been noticed and analyzed by SLEIJPEN and VAN DER SLUIS [1996, Section 6]. They show that the CG-error for a process with matrix A, characterized by the eigenvalues λ_j and weights γ_j, can be interpreted as a Ritz value for a related process that is characterized by eigenvalues $\bar{\lambda}_j$, augmented with $\bar{\lambda}_0 = 0$, and a set of weights $\bar{\gamma}_j$. They give explicit formulas for these eigenvalues and weights, and also lower bounds for the smallest Ritz value of this scheme. The approach in SLEIJPEN and VAN DER SLUIS [1996] is essentially the same as ours, but gives slightly more detailed results (at the cost of a more complicated analysis). We see that this correspondence between Ritz value convergence and CG-convergence allows us to transfer all the insights that result from analysis that has been carried out for convergence aspects of Ritz values (that is, for instance, misconvergence or stagnation, see VAN DER SLUIS and VAN DER VORST [1987a]) to CG-behavior.

22 Upper bounds

We are now also in the position to derive *a priori* upper bounds for the distance of a Ritz value to an eigenvalue. Such upper bounds were obtained in KANIEL [1966], PAIGE [1971], and SAAD [1980]. We return to equation (21.6), and we note that any polynomial of degree $i-1$, with value 1 in λ_k, leads to an upper bound. We consider the situation for the first eigenvalue in some more detail. Apparently, we have to look for a polynomial that is small in the eigenvalues $\lambda_2, \ldots, \lambda_n$, with respect to its value in λ_1. Typically, the shifted Chebyshev polynomials are used for this purpose and they produce the published upper bounds.

The *shifted Chebyshev polynomial* of degree i, is defined for an interval $[a, b]$, by

$$T_i(\lambda; a, b) := T_i\left(\frac{2\lambda - (a+b)}{b-a}\right).$$

The shifted Chebyshev polynomials satisfy the following properties (that follow easily from the properties of the original Chebyshev polynomials T_i, defined over $[-1,1]$):

$$\begin{cases} |T_i(\lambda;a,b)| \leq 1 \quad \text{for} \quad a \leq \lambda \leq b, \\ T_i(\lambda;a,b) = \frac{1}{2}\left[(\alpha+\sqrt{\alpha^2-1})^i + (\alpha+\sqrt{\alpha^2-1})^{-i}\right] \quad \text{for} \quad \lambda < a, \lambda > b, \end{cases} \tag{22.1}$$

with $\alpha = -1 + 2(\lambda - a)/(b-a)$.

If we replace $R_{i,1}(\lambda)$ in (21.6) by $T_{i-1}(\lambda;\lambda_n,\lambda_2)$, then we obtain the following upper bound (cf. SAAD [1980, Theorem 2]):

$$0 < \lambda_1 - \theta_1^{(i)} < \frac{(\lambda_1 - \lambda_n)\sum_{j\geq 2}\gamma_j^2}{T_{i-1}^2(\lambda_1;\lambda_n,\lambda_2)\gamma_1^2}. \tag{22.2}$$

This gives a more rapidly decaying upper bound when the gap between λ_1 and $[\lambda_n,\lambda_2]$ becomes larger. PARLETT [1980, p. 242] defines the gap ratio for λ_1 with respect to the remaining eigenvalues as $g_1 = (\lambda_1-\lambda_2)/(\lambda_2-\lambda_n)$. By neglecting the negative power in the expression for T_i, we see that

$$T_i(\lambda_1;\lambda_n,\lambda_2) > \frac{1}{2}e^{2i\sqrt{g_1}}. \tag{22.3}$$

This shows in a more quantitative way how the upper bound reduces when g_1 is bigger.

If the smallest eigenvalue is not well separated from the second one (an almost double eigenpair), then this upper bound is poor, but in that case, we can eliminate the influence of the second eigenvalue (at the price of a fairly large additional constant and a loss of one degree), by selecting (cf SAAD [1980, Theorem 4])

$$(\lambda_2 - \lambda)T_{i-2}(\lambda;\lambda_n,\lambda_3),$$

as a polynomial that produces an upper bound for (21.6). This choice leads to

$$0 < \lambda_1 - \theta_1^{(i)} < \frac{(\lambda_1-\lambda_n)\left[\frac{\lambda_2-\lambda_n}{\lambda_1-\lambda_2}\right]^2\sum_{j\geq 3}\gamma_j^2}{T_{i-2}^2(\lambda_1;\lambda_n,\lambda_3)\gamma_1^2}. \tag{22.4}$$

In VAN DER SLUIS and VAN DER VORST [1987a] an example is given for a situation, at the lower side of the spectrum (to which our formulas apply with a proper renumbering), with such a close eigenpair: $\lambda_n = 0.034$, $\lambda_{n-1} = 0.0341$, $\lambda_{n-2} = 0.082$, $\lambda_1 = 1.2$, and there the upper bound (22.2) shows hardly any convergence, whereas upper bound (22.4), although a fairly large overestimate, quite nicely predicts the (initial) speed of convergence, but is smaller than (22.2) only after about 25 iteration steps. In that phase of the process, the first Ritz value displays its superlinear convergence behavior (that is, the second and third eigenvalues get rather well approximated) and the exponential upper bound can, of course, not reflect this behavior (although we might select another polynomial for that phase, e.g., one in which a factor $(\lambda_3 - \lambda)$ is included).

The Chebyshev upper bounds suggest that the Lanczos method may converge much faster than the Power Method, which is in general the case. The Power Method leads to an approximation with respect to only the last Krylov vector, whereas the Lanczos method exploits the entire Krylov subspace. This is illustrated by a small example. In this example we compute the largest eigenvalue of a matrix, because the Power Method leads to approximations for the largest eigenvalue. The analysis (in particular the upper bounds) for the Lanczos method is similar for both sides of the spectrum. We take a diagonal matrix D of order 100 and with eigenvalues 1, 2, ..., 100. In order to include possible effects of rounding errors, we transformed this diagonal by a Householder matrix

$$Q = I - \frac{2}{w^* w} w w^*,$$

with $w = (1, 2, \ldots, 100)^*$. This leads to a matrix $A = Q^* D Q$ with eigenvalues 1, 2, ..., 100. The starting vector for the Lanczos method and the Power Method is chosen as the vector with all ones. The logarithm of the approximation errors in the largest Ritz value for Lanczos and the Power Method approximation are shown in Fig.10.

With the polynomial interpretation of the Lanczos process and with the properties of (discrete) orthogonal polynomials we can understand the local behavior of Ritz values in a qualitative way. In VAN DER SLUIS and VAN DER VORST [1987a] the situation for an almost double eigenpair λ_1, λ_2 has been analysed in great detail. The experimental observation that the first Ritz value converges initially to a point in between λ_1 and λ_2 could be substantiated by theory. By initial convergence we mean the situation that the first Ritz value moves, in successive Lanczos iterations, to a position in between the first two eigenvalues, and then makes very small moves in a number of successive Lanczos steps. This phenomenon is referred to as *misconvergence*, because it may easily lead a user of a Lanczos code to the conclusion that practical convergence has occurred and the fact that this Ritz value represents a close eigenpair may go unnoticed (cf PARLETT [1990]).

For practical computations in problems where an eigenvalue in the spec-

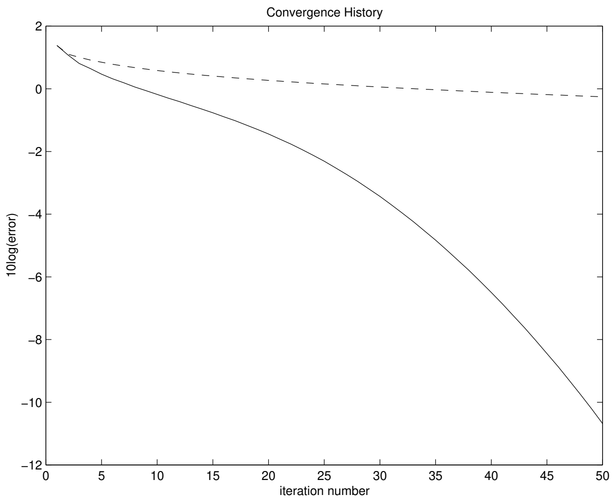

Figure 10: Lanczos error (-line-) and Power Method error (-dashed-)

trum may split, due to perturbations to the underlying physical problem, or may not, this is a dangerous and misleading situation. In VAN DER SLUIS and VAN DER VORST [1987a] it was proved that the phase of almost stagnation may take quite some steps, depending on how close the two eigenvalues are, but, most remarkably after the phase of stagnation the smallest Ritz values resumes its original progression to the smallest eigenvalue (that is the rate of convergence is comparable to what happened before the stagnation phase), and at the same time the second Ritz value moves to the second eigenvalue with virtually the same rate of convergence.

This second Ritz value had moved first to the third eigenvalue (as the process considered the close eigenpair as a single eigenvalue with the combined weight of the two closeby ones), and when the second one "discovers" the second eigenvalue, its place is almost immediately taken by the third Ritz value. We can understand this sudden "replacement" from the Lanczos recursion

$$\beta_i P_i = (A - \alpha_i I) P_{i-1} - \beta_{i-1} P_{i-2},$$

which says that if Ritz values are located close to the same eigenvalue in two consecutive steps of the Lanczos process (even though to the 'wrong' eigenvalue, that is, the eigenvalue has not the same index as these Ritz values), say in step $i-2$ and $i-1$, then the corresponding Lanczos polynomials P_{i-2}

and P_{i-1}, have zeros $\theta_k^{(i-2)}$, $\theta_k^{(i-1)}$, respectively, with $\theta_k^{(i-2)} \approx \theta_k^{(i-1)}$. This implies that also P_i should have a nearby zero, unless β_{i+1} is pathologically small.

Some other nice and elegant results can be found in VAN DER SLUIS and VAN DER VORST [1987a]. We mention the result that if we increase (decrease) a weight γ_j in absolute value, then the Ritz values are attracted (repelled) to the eigenvalue λ_j. This seems plausible, but a remarkable aspect is (which could also be proved) that the Ritz values never pass the next closest one of the Ritz values corresponding to the unchanged weights. By pulling a weight one can move a Ritz value in the direction of the corresponding eigenvalue, but never beyond the next Ritz value of the scheme corresponding to the original weights (at least, if there is a next one in between the Ritz value that we attempt to attract and the concerned eigenvalue). For more details and a proof see VAN DER SLUIS and VAN DER VORST [1987a, Theorem 6.23].

Computable error bounds The basic relation in matrix form

$$AV_m = V_{m+1} T_{m+1,m}$$

can also be formulated as

$$AV_m = V_m T_{m,m} + \beta_m v_{m+1} e_m^T. \tag{22.5}$$

In this formulation e_m denotes the m-th canonical basis vector in $\mathbb{R}^m$:

$$e_m = (0, 0, \ldots, 0, 1)^T.$$

The Ritz values and Ritz vectors of A follow from the eigenvalues and eigenvectors of $T_{m,m}$, and for that reason the formulation in (22.5) is convenient. Let θ, y denote an eigenpair of $T_{m,m}$, then it follows that

$$\begin{aligned} AV_m y &= V_m T_{m,m} y + \beta_m v_{m+1} e_m^T y \\ &= \theta V_m y + \beta_m y_{(m)} v_{m+1}, \end{aligned}$$

and $y_{(m)}$ denotes the last (mth) coordinate of y.
From this we conclude for the Ritz pair $\theta, s = V_m y$ that

$$r_{\theta,s} := As - \theta s = \beta_m y_{(m)} v_{m+1}. \tag{22.6}$$

The perturbation results for symmetric matrices (see Ch. 5) then lead to the observation that there is an eigenvalue λ_i of A such that

$$|\lambda_i - \theta| \leq ||r_{\theta,s}||_2 = |\beta_m y_{(m)}|.$$

If λ_i is well separated from the other eigenvalues, then the error bounds for Rayleigh quotients (note that the Ritz values are Rayleigh quotients with respect to the corresponding Ritz vectors) suggest even sharper results. Let

$$\mathrm{gap}_i := \min_{j \neq i} |\lambda_j - \theta| \tag{22.7}$$

and let x_i denote the normalized eigenvector corresponding to λ_i, with ϕ the angle between s and x_i. Then (cf. PARLETT [1980, p. 222], see also Chapter 5)

$$|\lambda_i - \theta| \leq \frac{||r_{\theta,s}||_2^2}{\mathrm{gap}_i}, \quad \sin\phi \leq \frac{||r_{\theta,s}||_2^2}{\mathrm{gap}_i}. \tag{22.8}$$

This requires knowledge of the gap, but in practical situations this can be easily estimated. Usually, if $\theta_i^{(m)}$ has arrived, for some m, close enough to an eigenvalue λ_j, then the nearby Ritz values $\theta_{i-1}^{(m)}$ and $\theta_{i+1}^{(m)}$ are also close to the eigenvalues λ_{j-1} and λ_{j+1}, respectively (the closeness can be controlled by the computable norms of the respective residuals). In the words of Parlett (PARLETT [1980, p. 223]): "*there comes a time when it is very unlikely that there remain undetected [eigenvalues] among the detected [eigenvalues]*". We interpret that this has happened for values of m for which a Ritz value and its neighbouring Ritz values seem to have arrived near eigenvalues of A. Parlett suggests to approximate the gap by

$$\min_j (|\theta_i^{(m)} - \theta_j^{(m)}| - ||r_{\theta_j,s_j}||_2), \tag{22.9}$$

where s_j is the Ritz vector corresponding to $\theta_j^{(m)}$. This result can be obtained by replacing λ_j in (22.7) with

$$\theta_j^{(m)} \pm \|r_{\theta_j,s_j}\|_2.$$

Note that this bound is only useful for rather well-separated eigenvalues; for clustered eigenvalues we can only derive such bounds in terms of gaps between the cluster and the remaining part of the spectrum. For this situation we refer to the discussion in PARLETT [1980, Chapter 11-7-1].

23 A Generalized Lanczos Scheme

In VAN DER VORST [1982] it has been shown that the Lanczos process can also be used for matrices A of the form $A = CB$, where B is symmetric positive definite and C is either symmetric or skew-symmetric. This can be accomplished by using another inner product for the Lanczos process:

$$[x, y] := (x, By),$$

which defines a proper inner product $[\,,\,]$, since B is symmetric positive definite.

We consider first the case $C = C^T$. Then we note that

$$[Av, w] = (CBv, Bw) = (v, BCBw) = [v, Aw],$$

which shows that A is symmetric with respect to the new inner product. Thus we can carry out the Lanczos process with this new inner product, which leads again to

$$AV_i = V_{i+1}T_{i+1,i},$$

with a tridiagonal matrix $T_{i+1,i}$, but now the orthogonality constraint is with respect to the B-inner product, implying that V_i is orthonormal, with respect to the new inner product $[\,,\,]$.

Likewise, Rayleigh quotients have to be interpreted with respect to this inner product as well. The alternative for this is to factorize B as LL^T, a Cholesky splitting, and to observe that $CBx = \lambda x$, with the transformation $x = L^{-T}y$, is equivalent with $L^TCLy = \lambda y$. Of course, we can now apply the Lanczos method for L^TCL in an obvious manner, but this requires the factorization of B, which is avoided with the B-inner product. This is attractive, for instance if one wants to compute the eigenvalues of preconditioned operators. With our approach, it is sufficient to know that the preconditioner is positive definite, and it is not necessary to split it explicitly.

The eigenvalues θ and eigenvectors y of $T_{i,i}$ lead directly to the approximated eigenpairs of A:

$$AV_iy = V_{i+1}T_{i+1,i}y = \theta V_iy + \beta_{i+1}v_{i+1}(y)_i,$$

so that $s = V_iy$ is the Ritz vector for Ritz value θ, and the residual for this eigenpair is given by $\beta_{i+1}v_{i+1}(y)_i$. The scalar $(y)_i$ denotes the i-th coordinate of y.

For the case $C = -C^T$ and $B = I$, we first observe that the orthogonalization procedure for the Krylov subspace leads to

$$CV_i = V_{i+1}H_{i+1,i},$$

so that $V_i^T C V_i = H_{i,i}$. Because C is skewsymmetric, the reduced matrix $H_{i,i}$ must also be skewsymmetric, which shows that the Lanczos process for a skewsymmetric matrix C, leads to a skewsymmetric tridiagonal matrix $T_{i,i}$. The eigenvalues of this skewsymmetric tridiagonal matrix can be computed by the convenient procedures for a Hermitian matrix, by realizing that $\imath T_{i,i}$, with $\imath = \sqrt{-1}$, is Hermitian. The eigenvalues of that matrix, which are all real, have to be multiplied with $\imath$ to give the desired Ritz values of C.
For more general positive definite B, we proceed as in the symmetric case, as described above. Note that CB is now skewsymmetric with respect to the B-inner product.

24 Lanczos in finite precision

In finite precision the Lanczos algorithm behaves quite so as predicted for exact arithmetic *until* an eigenvalue has converged to full machine precision. Almost exactly at that point the orthogonality among the newly computed basis vectors is lost. One might say that convergence of the Lanczos process and loss of orthogonality go hand in hand. This phenomenon was initially not well understood, and it was Paige who in 1971 (Paige [1971]) gave an analysis for it. As it turns out, the loss of orthogonality is lost only in the direction of the converged Ritz vector(s). The main effect is that, because the converged Ritz vector becomes part again of the new Lanczos vectors, the corresponding eigenvector and its associated eigenvalue are discovered by the algorithm again.

This history repeats itself and leads to the so-called *multiplets*: groups of (almost) multiple Ritz values (Parlett [1980]). Note that these multiple Ritz values have nothing to do with possible multiple eigenvalues of the given matrix, they occur simply as a result of earlier convergence. How can we detect the converged eigenvalues and separate them from those that are the result of rounding errors? A simple trick, that was advocated by Cullum and Willoughby [1986], is the following. Consider $T_{i,i}$ and strip the first column and row of this matrix, call the remaining $i-1$ by $i-1$ part $\underline{T}_{i-1}$. Compute the eigenvalues of $T_{i,i}$ and $\underline{T}_i$. Those that differ less than a small multiple times machine precision are the unwanted eigenvalues (that is, the ones due to loss of orthogonality), and remove them from the eigenvalue set of $T_{i,i}$.

The justification for this strategy is the following. The Ritz vectors $s_j^{(i)}$ of A with respect to $K_i(A; v_1)$ can be written as $s_j^{(i)} = V_i y_j$, where y_j is an eigenvector of $T_{i,i}$. The columns of V_i form a basis for the Krylov subspace, so that, as long as orthogonality has not been lost, each of them represents an essential component of the eigenvector, essential in the sense that if we remove a vector column v_j, then the resulting error in the eigenvector can not be compensated by the other columns. For convergence it is necessary that the last components of y_j go to zero when i increases.

Now note that $\underline{T}_i$ is the matrix that corresponds to the subspace from which v_1 has been removed, Consequently, none of the eigenvectors should have converged (nor their corresponding eigenvalues), and, necessarily, those eigenvalues that are the same as for $T_{i,i}$ correspond to invariant subspaces introduced by rounding errors. Whether these eigenvalues provide useful information, depends upon the situation. For the unsymmetric Lanczos process, rounding errors may lead to invariant subspaces for a perturbed matrix A that has little in common with the unperturbed A.

The main negative effect of the loss of orthogonality, in the symmetric case, seems to be a delay in the convergence for the remaining eigenvalues. If some of the extremal eigenvalues are well separated, in a relative sense, from the spectrum, then they are approximated by Ritz values in an early phase of the iteration process. After convergence to machine precision, the loss of orthogonality sets in and after a while the detected eigenvalue is rediscovered. This implies that the computed basis for the Krylov subspace contains almost dependent vectors, and this leads to a loss of efficiency (in particular if this happens long before other wanted eigenvalues have been discovered). If one wants the eigenvectors as well, then one has to store the basis vectors and also in this case, one wants to avoid working with unnecessarily many vectors. The subspace can be kept at an effective size by a process that is called *selective orthogonalization*. This comes down to keeping future Lanczos vectors also explicitly orthogonal to converged Ritz vectors. The technique has been investigated in Parlett and Scott [1979]; for an extensive discussion see also Parlett [1980].

25 Alternative approximations

The Krylov subspace offers more possibilities for the computation of useful approximations for eigenvalues and eigenvectors. In Paige, Parlett and van der Vorst [1995], the *harmonic Ritz values* are discussed, together with some interesting properties. The harmonic Ritz values are defined as the inverses of the Ritz values of A^{-1} with respect to the subspace $AK^i(A; v_1)$. Using the Rayleigh quotient characterization, the Ritz values of A^{-1} can also be defined as the local extrema of (cf Beattie [1998], see

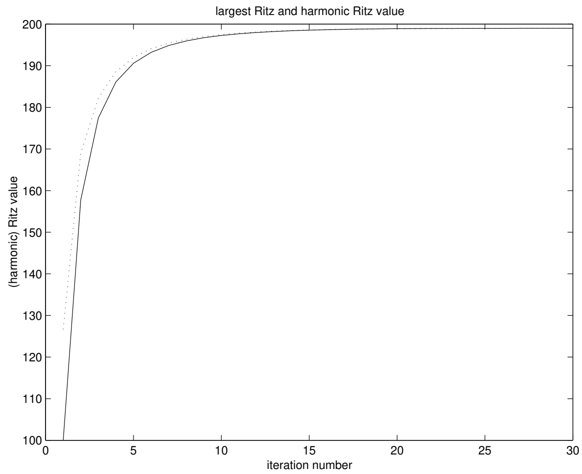

Figure 11: Largest Ritz value (-line-) and largest harmonic Ritz value (-dotted-)

also Th. 2)

$$\begin{aligned} \rho(A^{-1}, Ax) &= \frac{(A^{-1}Ax)^T Ax}{(Ax)^T Ax} \\ &= \frac{x^T Ax}{x^T A^2 x}, \end{aligned} \tag{25.1}$$

for nontrivial $x \in K^i(A, v_1)$.

For the computation of the local extrema, it is useful to have an orthonormal basis for $AK^i(A, v_1)$. This can be made using (19.1) as follows

$$V_i^T AAV_i = T_{i+1,i}^T V_{i+1}^T V_{i+1} T_{i+1,i} \tag{25.2}$$

$$= T_{i+1,i}^T T_{i+1,i}. \tag{25.3}$$

Note that $T_{i+1,i}^T T_{i+1,i}$ is positive definite, since no column of V_i is in the kernel of A. This implies that we can compute a Choleski decomposition

$$T_{i+1,i}^T T_{i+1,i} = U^T U,$$

with U an i by i upper triangular matrix. This implies that the columns of the matrix AV_iU^{-1} are orthonormal, so that these columns can be used

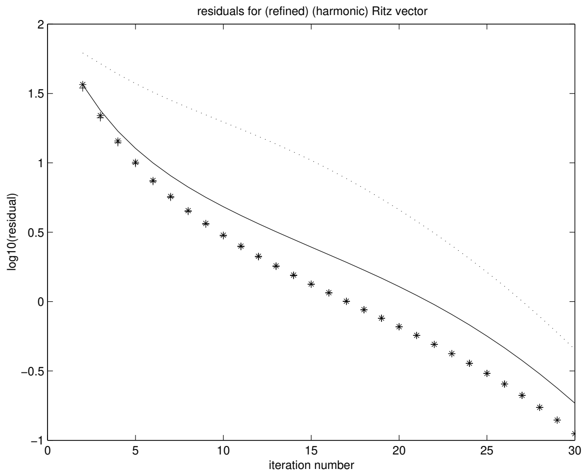

Figure 12: Residuals for largest Ritz vector (-line-) and largest harmonic Ritz vector (-dotted-), refined Ritz vector (+), and refined harmonic Ritz vector (*)

as an orthormal basis for $AK^i(A; v_1)$. Any vector z in this subspace can be expressed as $z = AV_iU^{-1}y$, with $y \in \mathbb{R}^i$.

For these vectors z, the Rayleigh quotient $\rho(A^{-1}, z)$ takes the form

$$\rho(A^{-1}, z) = \frac{y^T U^{-T} V_i^T A A^{-1} A V_i U^{-1} y}{y^T U^{-T} V_i^T A A V_i U^{-1} y} \tag{25.4}$$

$$= \frac{y^T U^{-T} T_{i,i} U^{-1} y}{y^T y}. \tag{25.5}$$

This shows that the Ritz values μ of A^{-1} are the eigenvalues of the eigenproblem

$$U^{-T} T_{i,i} U^{-1} y = \mu y.$$

With the transformation $u = U^{-1}y$, it follows that

$$T_{i,i} u = \mu U^T U u,$$

or

$$T_{i,i} u = \mu T_{i+1,i}^T T_{i+1,i} u. \tag{25.6}$$

Since we are primarily interested in eigenvalues of A, we conclude that the *harmonic Ritz values* $\theta_H := \mu^{-1}$ are the eigenvalues of the generalized eigenproblem

$$T_{i+1,i}^T T_{i+1,i} u = \theta_H T_{i,i} u. \tag{25.7}$$

The eigenvectors u lead to *harmonic Ritz vectors*

$$x_H = AV_i u.$$

Note that these harmonic Ritz values and vectors are formally associated with A^{-1}, so that it is natural to consider the residual

$$r = A^{-1} x_H - \frac{1}{\theta_H} x_H,$$

which, using he expression for x_H, can be written as

$$r = V_i u - \frac{1}{\theta_H} AV_i u.$$

This shows that the pair $\theta_H, V_i u$ can be considered as an eigenpair approximation for A, with residual $\widetilde{r} = \theta_H r$.

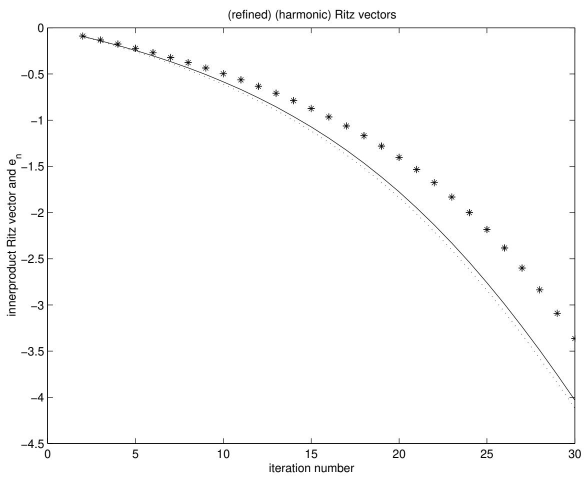

Figure 13: Log of the error components in the direction of e_n of Ritz vector (-line-), harmonic Ritz vector (-dotted-), and refined harmonic Ritz vector (*)

The harmonic Ritz values and vectors can be computed from the generalized eigenproblem (25.7). This equation can be reformulated as

$$\left(T_{i,i}^2 + \beta_i^2 e_i e_i^T - \theta_H T_{i,i}\right) u = 0,$$

which shows that the harmonic Ritz values satisfy the equation

$$\det\left(T_{i,i}(T_{i,i} - \theta_H I) + \beta_i^2 e_i e_i^T\right) = 0. \tag{25.8}$$

It was this type of equation that was used in LEHMANN [1966] for the definition of the so-called *Lehmann intervals*.

In order to explain these Lehmann intervals, we have to consider the properties of the harmonic Ritz values in some more detail. This will teach us that they provide useful information in particular for interior eigenvalues of A.

The Ritz values are shift invariant in the sense that if the matrix A is shifted to $A - \sigma I$, then the Ritz values θ are shifted to Ritz values $\theta - \sigma$ corresponding to $A - \sigma I$. The Ritz vectors do not depend on the shift at all. Because

$$(A - \sigma I)V_i = V_{i+1}(T_{i+1,i} - \sigma I_{i+1,i})$$

($I_{i+1,i}$ defines a part of the $i+1$ by $i+1$ identity matrix), we conclude that the harmonic Ritz values are not shift invariant since they are the zeros of

$$\det\left((T_{i,i} - \sigma I)(T_{i,i} - (\sigma + \theta_H)I) + t_{i+1,i}^2 e_i e_i^T\right) = 0. \tag{25.9}$$

It is obvious that this lead to a more complicated dependence on σ. From the Minimax Theorem 2 it follows that the Ritz values for $(A - \sigma I)^{-1}$ are the local maxima of the Rayleigh quotient $\rho((A - \sigma I)^{-1}, (A - \sigma I)x)$, for $x \in K^i(A; v_1)$. Therefore, they are always below (or above) the corresponding positive (or negative) eigenvalues of $(A - \sigma I)^{-1}$ (cf. the interlace property, see Ch. 21).

Let A be symmetric positive definite and let us assume that A has n different eigenvalues. We denote the r eigenvalues of A smaller than σ as

$$\lambda_{-r} < \lambda_{-r+1} < \ldots < \lambda_{-1} < \sigma.$$

Likewise, let the eigenvalues larger than σ be denoted by

$$\sigma < \lambda_1 < \lambda_2 < \ldots < \lambda_{n-r}.$$

Write $\vartheta^{(i)}$ for the harmonic Ritz values of $A - \sigma I$, with respect to the i dimensional subspace $(A - \sigma I)K^i(A, v_1)$. Then the Minimax Theorem applied to $(A - \sigma I)^{-1}$ for the largest and the smallest eigenvalue, leads to the following results:

$$\vartheta_{-1}^{(i)} + \sigma < \lambda_{-1} \quad \text{and} \quad \lambda_1 < \sigma + \vartheta_1^{(i)}.$$

The $\vartheta_j^{(i)}$ define the so-called *Lehmann intervals*:

$$[\sigma, \sigma + \vartheta_j^{(i)}],$$

for which LEHMANN [1963] and LEHMANN [1966] prove the following theorem (the above relations for $\lambda_{\pm 1}$ and $\vartheta_{\pm 1}^{(i)}$ are special cases).

Theorem 4 *(*LEHMANN *[1966]) Each interval $[\sigma, \sigma + \vartheta_j^{(i)}]$, $j = 1, \ldots, i - t$ contains at least j eigenvalues of A. Each interval $[\sigma + \vartheta_{-j}^{(i)}, \sigma]$, $j = 1, \ldots, t$ contains at least j eigenvalues of A. Moreover, in the absence of extra information no smaller intervals have these properties.*

Another interesting property of the harmonic Ritz values is that the $\vartheta_{-j}^{(i)} + \sigma$ convergence monotonically, for increasing i, from below to λ_{-j}. Also, the $\vartheta_j^{(i)} + \sigma$ converge monotonically from above, for increasing i, to λ_j. This can be deduced right away from similar monotonicity properties of the Ritz values associated with $(A - \sigma I)^{-1}$. For details on this see PAIGE, PARLETT and VAN DER VORST [1995].

Numerical Example We take the numerical example from PAIGE, PARLETT and VAN DER VORST [1995]. A is the diagonal matrix of order 100, with diagonal elements (eigenvalues) $a_{k,k} = 2k - 1$. As a starting vector for the Lanczos process we take the vector with all ones (and normalize this vector). In Figure 11 we have plotted the largest Ritz value and the largest harmonic Ritz value (the latter being only slightly better, as expected (see BEATTIE [1998])).

The situation is different, however, if we consider the residual vectors. In Figure 12 we have plotted the 2-norms of the residuals for the largest Ritz pair and the largest harmonic Ritz pair. In this picture we have also plotted the 2-norm of the residual of the refined approximation (see Section 27), computed with the largest harmonic Ritz vector. This is apparently the best result (in this case).

We have also plotted, in Figure 13, the log of the difference between the inner product of the (largest) Ritz vector with the largest eigenvector and the limit value 1, and similarly for the largest harmonic Ritz vector and the largest eigenvector. We have also included the inner product of the refined approximation with the largest eigenvector (see Section 27).

We see that, as expected from the Rayleigh quotient argument for the harmonic Ritz vector, the harmonic Ritz vector makes the smallest angle with the largest eigenvector of A (although, in this case, only marginally better than the Ritz vector). See Section 27 for an example of the situation for an interior eigenvalue.

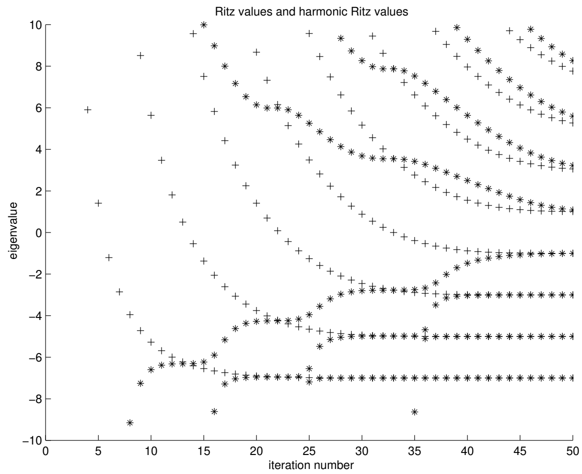

Figure 14: Ritz values (+) and harmonic Ritz values (*) for $A - \sigma I$

Now we take $\sigma = 8$, so that $A - \sigma I$ has four negative eigenvalues. In Fig. 14, we have plotted for successive values of i the harmonic Ritz values for $A - \sigma I$, and the Ritz values for the same matrix (these are the shifted Ritz values of A). We see that the respective harmonic Ritz values converge monotonically towards the eigenvalues near the shift, and no harmonic Ritz value ever appears inside the interval $[\lambda_{-1} - \sigma, \lambda_1 - \sigma] = [-1, 1]$. Ritz values may be in this interval from time to time, but they are systematically converging to the exterior eigenvalues of $A - \sigma I$. This means that a Ritz value that is occasionally close to some eigenvalue close to σ does not necessarily have a significant component of its corresponding eigenvector in the direction of the eigenvector corresponding to the nearby eigenvalue.

For the harmonic Ritz value and corresponding harmonic Ritz vector this is different. Since ϑ_{-j} represents a maximum of a Rayleigh quotient in relation with λ_{-j}, the corresponding harmonic Ritz vector may be expected to have a significant component in the direction of the eigenvector associated with λ_{-j} (significant compared with the occasionally nearby Ritz value).

26 Accuracy of harmonic Ritz values

We finish our discussion on harmonic Ritz values by remarking that (25.7) has the disadvantage that we need to work with $T_{i+1,i}^T T_{i+1,i}$, which may lead to inaccuracies if $T_{i,i}$ has a relatively large condition number. Kahan found a more accurate way for determining the harmonic Ritz values (see PARLETT [1980, Chapter 10], PAIGE, PARLETT and VAN DER VORST [1995]). We consider again the situation for a shift σ. We define

$$\delta_i^{-1} := e_i^T (T_{i,i} - \sigma I)^{-1} e_i.$$

One can show that δ_i is the (i,i) element of the diagonal matrix Δ_i in the triangular factorization $T_{i,i} - \sigma I = L_i \Delta_i L_i^T$. In this factorization L_i is a unit diagonal lower triangular matrix. Kahan showed that the matrix

$$\bar{T}_{i+1,i+1} := \begin{bmatrix} T_{i,i} - \sigma I & \beta_i \\ \beta_i & \beta_i^2/\delta_i \end{bmatrix}, \tag{26.1}$$

with $\beta_i = t_{i+1,i}$, has one zero eigenvalue, and the other i eigenvalues are the harmonic Ritz values for $A - \sigma I$.

Note that because of (26.1), the harmonic Ritz values and the Ritz values (for $A - \sigma I$) interlace.

Kahan's result follows from evaluating the determinant $\det(\bar{T}_{i+1,i+1} - \vartheta I)$: this determinant is, apart from a factor *constant* $\times\, \vartheta$, equal to the determinant in (25.9).

Note that the harmonic Ritz values are very sensitive to the shift σ. Indeed, when σ coincides with an eigenvalue of A, then this eigenvalue cannot be detected by the harmonic Ritz values. More precisely, because the harmonic Ritz values are the eigenvalues of the projection of $(A - \sigma I)^{-1}$ with respect to the space $(A - \sigma I)\mathcal{K}^m(A - \sigma I, v_1)$, and since this subspace contains no components in the kernel of $A - \sigma I$, we conclude that vectors in this kernel are not recovered. Also, one may expect that if σ is close to an eigenvalue, then this eigenvalue can only be sufficiently well approximated by a harmonic Ritz value if the dimension m of the Krylov subspace is rather large. In other words, the harmonic Ritz value corresponding to an eigenvalue λ goes to infinity if we let σ approach λ.

The harmonic Ritz vectors seem to depend less critically on the variation in σ, and for this reason it has been advocated to rely on the harmonic

Ritz vector rather than on the harmonic Ritz value, see SLEIJPEN, VAN DER VORST and MEIJERINK [1998, Section 4]. In particular, it is suggested in that paper to recompute the approximate eigenvalue as the Rayleigh quotient for the computed harmonic Ritz vector, rather than using the harmonic Ritz value. The most suitable harmonic Ritz vector in step m is selected as the one that makes the smallest angle with the best approximation so far (for instance, the selected eigenvector approximation in step $m-1$).

We will illustrate this with an example of what may happen.

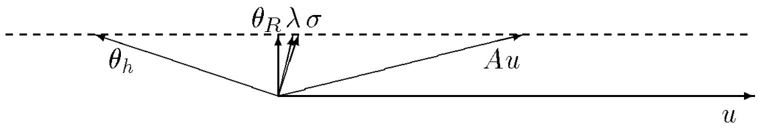

Figure 15: Various possibilities for the residual

In Fig. 15 we have sketched a situation that may occur in practice. Suppose that we have a target value σ close to an eigenvalue λ. In the picture we have drawn the harmonic Ritz vector $u = V_m y$, the vector Au, and the residual vectors for various values of the (approximated) eigenvalue. Note that these residual vectors are on a line indicated by the dashed line: $Au - \alpha u$.

The residual vector for the target value σ is indicated by a σ. The residual for the actual (but unknown) eigenvalue will most likely be somewhat smaller: indicated by λ. The residual $Au - \theta_h u$, corresponding to the harmonic Ritz value θ_h, is orthogonal to the space $(A-\sigma I)\mathcal{K}^m(A-\sigma I; v_1)$ (and in particular to $Au - \sigma u$). This vector is indicated by θ_h. If σ is close to λ, and u is a good approximation for the corresponding eigenvector, then necessarily the residual vector $Au - \theta_h u$ must be relatively large (in the limit, when σ approaches λ, the value of θ_h goes to $\pm\infty$).

With θ_R, we have indicated the residual computed with the harmonic Ritz vector u and the Rayleigh quotient for this vector. By construction this vector has minimal length for all possible residuals with u, and the picture suggests that θ_R represents a much better approximation for the wanted eigenvalue λ than any other of the combinations.

27 Refined eigenvector approximations

In Section 17 we have discussed Jia's ideas for further refinement of eigenvector approximations (originally proposed in JIA [1997]). One has to be careful with applying this idea for the further improvement of (harmonic) Ritz vectors associated with interior eigenvalues. As we have argued before, the harmonic Ritz vector associated with a θ_h closest to a target τ gives in some sense the best component for an eigenvector associated with an eigenvalue close to the target. If we attempt to get a better approximation for the eigenvector, by minimizing

$$||AV_i s - \theta_h V_i s||_2, \tag{27.1}$$

for s, then this may lead to a smaller residual with $\widehat{u} = V_i s$, but this u may be a composite of eigenvectors farther away from the target, and, in fact, there is no need that it has a significant component in any eigenvector direction close to the target. This is also not necessarily the case for the Rayleigh quotient θ_R:

$$\min_{||s||_2} ||AV_i s - \theta_R V_i s||_2,$$

which leads to $\widetilde{u} = V_i s$.

The pair θ_h, u, or even better, θ_R, u, represents good information for a restart, while the pairs $\theta_R, \widetilde{u}$ and $\theta_h, \widehat{u}$ may be *not* as useful for restart. We will illustrate this by a numerical example.

Numerical example Without essential limitations for our example, we choose a diagonal matrix A. The matrix has diagonal elements

$$-60, -59, \ldots, -13, -12, -1, 1, 12, 13, \ldots, 59, 60,$$

that is, the eigenvalues in between -60 and -12, and in between 12 and 60, are equally spaced with spacing 1. The order of the matrix is 100. For the Lanczos process we have taken a starting vector of all ones.

We are interested in the smallest nonnegative Ritz value θ_{+1} and the smallest nonnegative harmonic Ritz value θ^h_{+1}, as an approximation for the interior eigenvalue $+1$. The corresponding approximate Ritz vectors are u_{+1} and u^h_{+1}, respectively. They are approximations for the eigenvector corresponding to $+1$, that is, the canonical vector e_{51}.

With u^h_{+1} we have also computed the Rayleigh quotient

$$\theta_R = {u^h_{+1}}^T A u^h_{+1}$$

(note that $||u^h_{+1}||_2 = 1$).

With θ_{+1} and θ_R, we have also computed the *refined* eigenvector approximations, by solving the minimization problem (27.1). This leads to the vectors u_r and u_R, respectively.

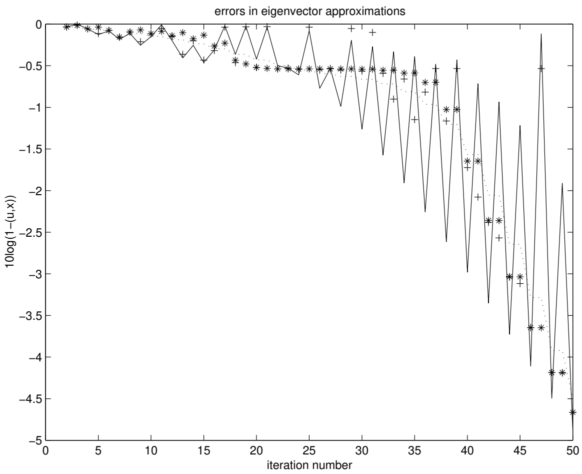

Figure 16: The value of $\log_{10}(1 - |u^T e_{51}|)$ for the standard Ritz vector $(-)$, the refined vector corresponding to the Ritz value θ_1 $(+)$, the harmonic Ritz vector (...), and the refined vector for the Rayleigh quotient θ_R $(*)$

The absolute values, of the inner product of e_{51} and the successive eigenvector approximations, should ideally converge to 1, for increasing dimension of the Krylov subspace.

In Figure 16, we have plotted the 10-log of the difference between the absolute inner products and 1 for the four different eigenvector approximations as a function of the Krylov subspace dimension.

As expected, the standard Ritz vector approximation, whose deviation of the inner product $u_{+1}^T e_{51}$ with respect to 1 is marked with the drawn line, shows a rather irregular convergence behavior. The corresponding errors in the inner products for the refined approximation u_r, have been marked with a '+'. They are on the average not that much better, and they still display irregular convergence behavior. Note that, for every odd subspace dimension, there is always a standard Ritz value equal to 0, that is, relatively close to the smallest positive eigenvalue $+1$. This may help to explain the very jumpy convergence history for the standard Ritz vector inner product, and also for its refined counterpart.

The harmonic Ritz vector (the errors in the inner product with e_{51} have been marked with dots) shows a regular monotonically decreasing convergence behavior, ideal for restart purposes.

Finally, if we compute the refined approximate eigenvector with θ_R, then we obtain the '$*$'-s. They are not monotonic, nor always better than the harmonic Ritz vector, but for larger subspace dimension they give some improvement.