University of Maryland College Park

Institute for Advanced Computer Studies TR–97–74
Department of Computer Science TR–3839

On the Convergence of a New Rayleigh Quotient Method with Applications to Large Eigenproblems*

D. P. O'Leary†
G. W. Stewart†

October 1997

ABSTRACT

In this paper we propose a variant of the Rayleigh quotient method to compute an eigenvalue and corresponding eigenvectors of a matrix. It is based on the observation that eigenvectors of a matrix with eigenvalue zero are also singular vectors corresponding to zero singular values. Instead of computing eigenvector approximations by the inverse power method, we take them to be the singular vectors corresponding to the smallest singular value of the shifted matrix. If these singular vectors are computed exactly the method is quadratically convergent. However, exact singular vectors are not required for convergence, and the resulting method combined with Golub–Kahan–Krylov bidiagonalization looks promising for enhancement/refinement methods for large eigenvalue problems.

*This report is available by anonymous ftp from `thales.cs.umd.edu` in the directory `pub/reports` or on the web at `http://www.cs.umd.edu/ stewart/`.

†Computer Science Department and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742. This work was supported by the National Science Foundation under grant CCR-95-03126.

On the Convergence of a
New Rayleigh Quotient Method with
Applications to Large Eigenproblems

D. P. O'Leary*
G. W. Stewart*

Abstract

In this paper we propose a variant of the Rayleigh quotient method to compute an eigenvalue and corresponding eigenvectors of a matrix. It is based on the observation that eigenvectors of a matrix with eigenvalue zero are also singular vectors corresponding to zero singular values. Instead of computing eigenvector approximations by the inverse power method, we take them to be the singular vectors corresponding to the smallest singular value of the shifted matrix. If these singular vectors are computed exactly the method is quadratically convergent. However, exact singular vectors are not required for convergence, and the resulting method combined with Golub–Kahan–Krylov bidiagonalization looks promising for enhancement/refinement methods for large eigenvalue problems.

1. Introduction

The starting point for the algorithm analyzed in this paper is the following variant of the Rayleigh quotient method. Let A be of order n, and let λ be a simple eigenvalue of A with right and left eigenvectors x and y^{H}. Let $\tilde{v}$ and $\tilde{w}^{\mathrm{H}}$ be approximations to x and y^{H}, and let τ be an approximation to λ. Then new approximations $\hat{v}$, $\hat{w}^{\mathrm{H}}$, and $\hat{\tau}$ are generated as follows:

$$\begin{array}{ll} 1. & \hat{v} = (A-\tau I)^{-1}\tilde{v} \\ 2. & \hat{w}^{\mathrm{H}} = \tilde{w}^{\mathrm{H}}(A-\tau I)^{-1} \\ 3. & \hat{\tau} = \hat{w}^{\mathrm{H}}A\hat{v}/\hat{w}^{\mathrm{H}}\hat{v} \end{array} \tag{1.1}$$

This procedure can, of course be iterated. The quantity $\hat{\tau}$ is called the generalized Rayleigh quotient of A at $\hat{v}$ and $\hat{w}^{\mathrm{H}}$. Ostrowski [4] showed that under weak conditions on $\hat{v}$ and $\hat{w}^{\mathrm{H}}$ the shift τ converges cubically to λ provided that the initial shift is sufficiently near λ. There are two reasons for the fast convergence. First, steps 1 and 2 in (1.1) improve earlier approximations to the right and left eigenvectors. Second, this improvement is magnified by the generalized Rayleigh quotient, which is more accurate than an ordinary Rayleigh quotient formed from a single vector.

In this paper we will be concerned with a variant of this method in which the approximations $\tilde{v}$ and $\tilde{w}^{\mathrm{H}}$ are determined in a different way. We begin by noting that

*Computer Science Department and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742. This work was supported by the National Science Foundation under grant CCR-95-03126.

if $\tau = \lambda$ then $A - \tau I$ has a zero singular value, with right and left singular vectors x and y^{H}. Consequently, if τ is near λ, the right and left singular vectors v and w corresponding to the smallest singular value σ of $A - \tau I$ should approximate x and y^{H}. (We will make this statement more precise in Theorem 2.1.) For brevity we will call these singular vectors the inferior singular vectors of $A - \tau I$. In practice, we do not compute the inferior singular vectors exactly but instead approximate them. This suggests the following procedure, which can also be iterated.

$$\begin{array}{ll} 1. & \text{Let } \tilde{v} \text{ and } \tilde{w}^{\mathrm{H}} \text{ be approximations to the right and left inferior} \\ & \quad \text{singular vectors of } A - \tau I \\ 2. & \hat{\tau} = \tilde{w}^{\mathrm{H}} A \tilde{v} / \tilde{w}^{\mathrm{H}} \tilde{v} \end{array} \qquad (1.2)$$

Because we do not improve on previous vectors in step one, the scheme is slower than (1.1). But, as we will show, it converges quadratically if the singular vectors are exact, and otherwise it can still be fast. We will call the method the singular vector Rayleigh quotient (SVRQ) method.

At first glance the SVRQ method does not seem to have much to recommend it. It is more difficult to compute singular vectors than to solve linear systems, and consequently a SVRQ step (1.2) requires more work than a step of the original algorithm (1.1). And as we have noted, the new method is slower. Nonetheless, the method may be useful in finding eigenpairs of large matrices.

Specifically, over the past decade new algorithms have been developed to solve large eigenvalue problems by building up approximations to the eigenspaces of eigenvalues lying in a neighborhood of the complex plane. These algorithms (e.g., see [3, 5, 1]) generally begin with subspaces $\mathcal{V}$ and $\mathcal{W}$. The space $\mathcal{V}$ approximates a right eigenspace of A (the space $\mathcal{W}$ usually does not approximate a corresponding left eigenspace). In an enhancement step, the spaces $\mathcal{V}$ and $\mathcal{W}$ are expanded in such a way as to improve the approximations they contain. Since storage considerations limit the dimensions of the spaces, enhancement is followed by a refinement step in which unwanted vectors are purged from the spaces.

The enhancement step generally requires the solution of equations involving $A - \tau I$, where τ is a shift chosen during the refinement step.[1] If A is large, these systems cannot be solved directly, and iterative methods such as GMRES must be employed. Unfortunately, these iterative methods are computationally expensive and consume valuable storage. Moreover, although potentially useful information is generated in the course of the iteration, it is not easy to fold it into the algorithm. Consequently, the information is usually discarded and only the approximate solution is retained.

If we regard steps 1 and 2 in the algorithm (1.1) as enhancement steps, and step 3 as a refinement step (the analogies are not at all far-fetched), then the advantage of the new algorithm (1.2) becomes evident. It is true that (1.2) replaces the iterative

[1] The Jacobi–Davidson method works with a projected version of $A - \tau I$.

solution of a large nonsymmetric system with the iterative determination of inferior singular vectors. But there are effective, well-understood Krylov sequence methods for the singular value decomposition. In the present application the Golub–Kahan–Lanczos (GKL) bidiagonalization method is a natural.[2] This method generates two sequences of orthogonal vectors spanning Krylov subspaces defined by

$$\begin{array}{ccc} \hat{v}, & [(A-\tau I)^{\mathrm{H}}(A-\tau I)]\hat{v}, & [(A-\tau I)^{\mathrm{H}}(A-\tau I)]^2\hat{v}, \ \ldots \\ (A-\tau I)\hat{v}, & (A-\tau I)[(A-\tau I)^{\mathrm{H}}(A-\tau I)]\hat{v}, & (A-\tau I)[(A-\tau I)^{\mathrm{H}}(A-\tau I)]^2\hat{v}, \ \ldots \end{array}$$

The vectors in the first sequence contain approximations to the right singular vectors, while the vectors in the second contain approximations to the left singular vectors, which makes them natural candidates to add to $\mathcal{V}$ and $\mathcal{W}$. Moreover, since the singular subspaces also contain approximations to eigenvectors corresponding to eigenvalues near λ (see Theorem 2.1), the refinement step will benefit from the fact that we have approximations to both right and left eigenspaces.

As a first step in applying this idea to Krylov subspace methods, we give a convergence analysis of algorithm (1.2). To anticipate our results, we will show that if the singular vectors are computed exactly, then the method converges quadratically whenever the initial value of τ is sufficiently near λ and that the size of the convergence region is controlled by the condition numbers of λ and x. If the singular vectors are only approximated, then we give conditions under which convergence rate can be maintained.

This paper is organized as follows. In the next section we establish the relationship between the inferior singular vectors and the eigenvectors corresponding to small eigenvalues. Then we introduce a decomposition associated with a simple eigenvalue and establish a result on the accuracy of generalized Rayleigh quotients. In §3 we study the convergence of algorithm (1.2). In the final sections we discuss the results and draw conclusions. Throughout this paper $\|\cdot\|$ denotes the Euclidean vector norm and the subordinate spectral matrix norm.

2. Singular Subspaces, Eigenspaces, and Generalized Rayleigh Quotients

In this section we first study the relation between certain eigenspaces and singular spaces.

Theorem 2.1. *Let A be of order n. Let $X \in \mathbf{C}^{n\times p}$ have orthonormal columns and satisfy*

$$AX = XE, \tag{2.1}$$

[2] We use the appellation Golub–Kahan–Lanczos bidiagonalization to stress the fact that the method is based on Krylov sequences and to distinguish it from the Golub–Kahan reduction to bidiagonal form by orthogonal transformations. Actually both methods are due to Golub and Kahan [2].

where $E = X^{\mathrm{H}}AX$. Let A have the singular value decomposition

$$\begin{pmatrix} W_1^{\mathrm{H}} \\ W_2^{\mathrm{H}} \end{pmatrix} A(V_1\ \ V_2) = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix},$$

where Σ_1 is nonsingular of order p and the singular values are in descending order. If we denote by $\Theta(V_1, X)$ the diagonal matrix of canonical angles between the column spaces of V_2 and X, then

$$\|\sin\Theta(V_2, X)\| \le \frac{\|E\|}{\sigma_p}. \tag{2.2}$$

Proof. The sines of the canonical angles between X and V_2 are the singular values of $V_1^{\mathrm{H}}X$ (see [7, §I.5.2]). Multiplying (2.1) by W_1^{H} and using the fact that $W_1^{\mathrm{H}}A = \Sigma_1 V_1^{\mathrm{H}}$, we find that

$$W_1^{\mathrm{H}}XE = W_1^{\mathrm{H}}AX = \Sigma_1 V_1^{\mathrm{H}}X.$$

The inequality (2.2) now follows on multiplying by Σ_1^{-1} and taking norms. ■

A related result holds in which the spectral norm in (2.2) is replaced by the Frobenius norm. In plain words, the theorem says that if an invariant subspace of A has a small spectrum and the rest of the spectrum is well behaved in the sense that σ_p is larger than $\|E\|$, then as E approaches zero the invariant subspace and the corresponding singular subspace approach one another. This is the foundation of our algorithm, although we will use the result only for $p = 1$ [see (3.4)].

Next we introduce a decomposition associated with a simple eigenvalue and use it to assess the accuracy of the generalized Rayleigh quotient in algorithm (1.2). First the decomposition.

Theorem 2.2. *Let A be of order n. Let λ be a simple eigenvalue of A with right eigenvector x normalized so that $\|x\| = 1$ and left eigenvector y^{H} normalized so that $y^{\mathrm{H}}x = 1$. Then there are $n\times(n-1)$ matrices X and Y such that*

$$\begin{pmatrix} y^{\mathrm{H}} \\ Y^{\mathrm{H}} \end{pmatrix} (x\ \ X) = \begin{pmatrix} 1 & 0 \\ 0 & I \end{pmatrix}$$

and

$$\begin{pmatrix} y^{\mathrm{H}} \\ Y^{\mathrm{H}} \end{pmatrix} A(x\ \ X) = \begin{pmatrix} \lambda & 0 \\ 0 & L \end{pmatrix},$$

where

$$L = Y^{\mathrm{H}}AX = Y^{\mathrm{H}}AY.$$

Moreover

$$\|x\| = \|Y^{\mathrm{H}}\| = 1 \quad \text{and} \quad \|y^{\mathrm{H}}\| = \|X\| \equiv \kappa. \tag{2.3}$$

For a proof see [6]. The theorem states that the eigenvalue λ can be uncoupled from the rest of A by a similarity transformation and that the transformation has certain special properties, which we will use in the sequel. Note that there are block versions of this theorem in which x and y^{H} are replaced by matrices spanning left and right eigenspaces of A (see [7, §V.1]).

The number κ in (2.3) is a condition number for the eigenvalue λ [7, §IV.2.2]. Specifically, for sufficiently small E there is a unique eigenvalue $\tilde{\lambda}$ of $A+E$ such that

$$\tilde{\lambda} = \lambda + y^{\mathrm{H}}Ex + O(\|E\|^2).$$

It follows on taking norms that

$$|\tilde{\lambda} - \lambda| \leq \kappa\|E\| + O(\|E\|^2).$$

In other words κ plays the traditional role of a condition number by bounding the effects on the eigenvalue λ of errors in A.

We now consider the accuracy of the generalized Rayleigh quotient $\tilde{w}^{\mathrm{H}}A\tilde{v}/\tilde{w}^{\mathrm{H}}\tilde{v}$. We begin with the observation that in the notation of Theorem 2.2 any vector $\tilde{v}$ can be expressed in the form $\gamma x + Xg$, where $\gamma = y^{\mathrm{H}}\tilde{v}$ and $g = Y^{\mathrm{H}}\tilde{v}$. There is an analogous expression for $\tilde{w}^{\mathrm{H}}$. These expansions allow us to state the following theorem.

Theorem 2.3. *In the notation of Theorem 2.2, let*

$$\tilde{v} = \gamma x + Xg \quad \text{and} \quad \tilde{w}^{\mathrm{H}} = \eta y^{\mathrm{H}} + h^{\mathrm{H}}Y^{\mathrm{H}}.$$

Then

$$\frac{\tilde{w}^{\mathrm{H}}A\tilde{v}}{\tilde{w}^{\mathrm{H}}\tilde{v}} = \frac{\gamma\eta\lambda + h^{\mathrm{H}}Lg}{\gamma\eta + h^{\mathrm{H}}g}. \tag{2.4}$$

Moreover if $\|\tilde{v}\| = \|\tilde{w}^{\mathrm{H}}\| = 1$, *then as* h *and* g *approach zero, we have*

$$|\gamma| \to 1 \quad \text{and} \quad |\eta| \to \kappa^{-1}. \tag{2.5}$$

It follows that

$$\frac{\tilde{w}^{\mathrm{H}}A\tilde{v}}{\tilde{w}^{\mathrm{H}}\tilde{v}} = \lambda + \eta^{-1}(h^{\mathrm{H}}Lg - \lambda h^{\mathrm{H}}g) + o(\|g\|\|h^{\mathrm{H}}\|). \tag{2.6}$$

Proof. The expression (2.4) follows immediately from the relations in Theorem 2.2. The limits in (2.5) follow from the fact that when g and h^{H} are zero we have $\tilde{v} = \gamma x$ and $\tilde{w}^{\mathrm{H}} = \eta y^{\mathrm{H}}$. Finally, (2.6) follows on factoring $\gamma\eta$ from numerator and denominator in (2.4) and expanding the denominator in a Neumann series. ∎

If we take norms in (2.6) and note that $|\lambda|, \|L\| \leq \|A\|$, we get an asymptotic error bound for the generalized Rayleigh quotient

$$\left|\frac{\tilde{w}^{\mathrm{H}}A\tilde{v}}{\tilde{w}^{\mathrm{H}}\tilde{v}} - \lambda\right| \leq 2\kappa\|A\|\|g\|\|h^{\mathrm{H}}\| + o(\|g\|\|h^{\mathrm{H}}\|). \tag{2.7}$$

Thus the accuracy in the generalized Rayleigh quotient is proportional to the product of the errors $\|g\|$ and $\|h^{\mathrm{H}}\|$ in the right and left vectors.

3. Convergence of the SVRQ iteration

In this section we will consider the convergence of the SVRQ iteration. A single step of algorithm (1.2) ideally consists of computing the right and left inferior singular vectors v and w^{H} of $A-\tau I$ and then computing the Rayleigh quotient $\hat{\tau} = w^{\mathrm{H}}Av/w^{\mathrm{H}}v$ to give a new shift. In practice, though, we do not compute the singular vectors exactly. Instead we obtain $\tilde{v} = v+\delta_v$ and $\tilde{w}^{\mathrm{H}} = w^{\mathrm{H}}+\delta_w^{\mathrm{H}}$, where v and w are the inferior singular vectors and δ_v and δ_w^{H} are the unknown errors. To study the convergence rate of algorithm (1.2), we study the relation between $|\hat{\tau}-\lambda|$ and $|\tau-\lambda|$. From (2.7) it is seen that the crux of the matter is to derive expressions for the vectors g and h^{H}.

We begin by writing the singular value decomposition of $A-\tau I$ in the form

$$\begin{pmatrix} W^{\mathrm{H}} \\ w^{\mathrm{H}} \end{pmatrix} (A-\tau I)(V\ \ v) = \begin{pmatrix} \Sigma & 0 \\ 0 & \sigma \end{pmatrix}.$$

Here $(V\ \ v)$ and $(W\ \ w)$ are unitary. The quantity σ is the inferior singular value of $A-\tau I$, and v and w^{H} are the right and left inferior singular vectors. Although we do not indicate it explicitly, the components of this decomposition are functions of τ.

We will need a lower bound on the smallest singular value of Σ. Since λ is simple, this singular value is nonzero when $\tau=\lambda$. Hence it is bounded below by a positive constant when τ is restricted to a sufficiently small neighborhood λ. Thus we can let

$\mu =$ a positive lower bound for the smallest singular value of Σ in some neighborhood of λ.

We now turn to bounding $g = Y^{\mathrm{H}}\tilde{v}$. We begin by expanding x in terms of the right singular vectors:

$$x = (v^{\mathrm{H}}x)v + VV^{\mathrm{H}}x. \tag{3.1}$$

Multiplying this relation by Y^{H} and using the relation $Y^{\mathrm{H}}x = 0$, we find after a little manipulation that

$$g = Y^{\mathrm{H}}\tilde{v} = Y^{\mathrm{H}}v + Y^{\mathrm{H}}\delta_v = -\frac{Y^{\mathrm{H}}VV^{\mathrm{H}}x}{v^{\mathrm{H}}x} + Y^{\mathrm{H}}\delta_v. \tag{3.2}$$

The next step is to derive an expression for $V^{\mathrm{H}}x$. To do this we first exploit the eigendecomposition of A and then the singular value decomposition, as in Theorem 2.1. Specifically, we have

$$(A-\tau I)x = (\lambda-\tau)x.$$

Multiplying this expression by W^{H} and using the relation $W^{\mathrm{H}}(A - \tau I) = \Sigma V^{\mathrm{H}}$ we get $\Sigma V^{\mathrm{H}} x = (\lambda - \tau) W^{\mathrm{H}} x$ or

$$V^{\mathrm{H}} x = (\lambda - \tau)\Sigma^{-1} W^{\mathrm{H}} x. \tag{3.3}$$

We can now derive a bound on g. Taking norms in (3.3), we get

$$\|V^{\mathrm{H}} x\| \leq \frac{\epsilon}{\mu}, \tag{3.4}$$

where we have set

$$\epsilon = |\lambda - \tau|.$$

Since (3.1) is a decomposition of x into orthogonal components and $\|x\| = 1$, it follows that

$$|v^{\mathrm{H}} x| \geq \sqrt{1 - (\epsilon/\mu)^2}.$$

Hence from (3.2)

$$\|g\| \leq \frac{\epsilon/\mu}{\sqrt{1-(\epsilon/\mu)^2}} + \|Y^{\mathrm{H}}\delta_v\| \leq \frac{\epsilon/\mu}{\sqrt{1-(\epsilon/\mu)^2}} + \|\delta_v\|.$$

The derivation of a bound for $h^{\mathrm{H}} = w^{\mathrm{H}} X$ is similar, and we only reproduce the result:

$$\|h^{\mathrm{H}}\| \leq \frac{\kappa\epsilon/\mu}{\sqrt{1-(\epsilon/\mu)^2}} + \kappa\|\delta_w^{\mathrm{H}}\|.$$

The additional factor κ comes from the fact that we work with the eigenvector y^{H} and the matrix X, whose norms are κ, instead of x and Y.

If we now substitute these bounds in (2.7) we obtain

$$|\lambda - \hat{\tau}| \leq 2\kappa^2\|A\| \left(\frac{\epsilon/\mu}{\sqrt{1-(\epsilon/\mu)^2}} + \|\delta_v\|\right)\left(\frac{\epsilon/\mu}{\sqrt{1-(\epsilon/\mu)^2}} + \|\delta_w^{\mathrm{H}}\|\right) + o(\|g\|\,\|h^{\mathrm{H}}\|).$$

If we gather higher order terms, we obtain the expression in the following theorem.

Theorem 3.1. *In the notation of algorithm* (1.2) *and Theorem* 2.2, *if* $\epsilon = |\lambda - \tau|$ *is sufficiently small, there is a a constant* μ *such that*

$$\hat{\epsilon} \equiv |\lambda - \hat{\tau}| \leq \rho\epsilon + o(\rho\epsilon), \tag{3.5}$$

where

$$\rho = 2\kappa^2\|A\|\left(\frac{\epsilon}{\mu^2} + \frac{\|\delta_v\| + \|\delta_w^{\mathrm{H}}\|}{\mu}\right). \tag{3.6}$$

It follows that if the starting value of τ is sufficiently close to λ, we can obtain a superlinear convergence rate from the SVRQ iteration:

- If the singular vectors v and w are determined exactly, then $\delta_v = 0$ and $\delta_w^{\mathrm{H}} = 0$ and the convergence rate is quadratic.
- If we compute the approximate singular vectors accurately enough that $\|\delta_v\| = O(\epsilon)$ and $\|\delta_w^{\mathrm{H}}\| = O(\epsilon)$, then the quadratic rate of convergence is preserved.
- If $\|\delta_v\| = O(\epsilon^\alpha)$ and $\|\delta_w^{\mathrm{H}}\| = O(\epsilon^\beta)$, where $\alpha + \beta > 1$, then the convergence rate is superlinear.

4. Discussion

We have shown that the convergence order for the eigenvalue is quadratic if the singular vectors are computed exactly. In this case, the bound (3.4) also establishes the quadratic convergence of the sequence of vectors $\tilde{v}$ to x. Specifically, the quantity $\|V^{\mathrm{H}}x\|$ is the sine of the angle between x and $\tilde{v}$ [7, §I.5], which therefore goes to zero as fast as ϵ. A similar result holds for the convergence of the vectors $\tilde{w}^{\mathrm{H}}$.

We have established the local superlinear convergence of the SVRQ iteration to a simple eigenvalue, as long as the approximate singular vectors are accurate enough. In this case, the vectors $\tilde{v}$ still converge to x, and we have a lower bound on the cosine of the angle between $\tilde{v}$ and x, namely

$$|\tilde{v}^{\mathrm{H}}x| \geq \sqrt{1 - (\epsilon/\mu)^2} - \|\delta_v\|,$$

with a similar expression for the convergence of $\tilde{w}^{\mathrm{H}}$.

The multiplier ρ in (3.5) depends on κ and μ. We have already seen that the quantity κ is the condition number of the eigenvalue λ. The quantity μ is related to the condition of the eigenvectors. For it can be shown that when $\tau = \lambda$

$$\mu^{-1} = \|\Sigma^{-1}\| \leq \|(L - \lambda I)^{-1}\|.$$

The quantity $\|(L - \lambda I)^{-1}\|^{-1}$ is written $\mathrm{sep}(\lambda, L)$, and its reciprocal governs the sensitivity of the eigenvectors corresponding to λ [7, §V.2].

If λ is a nondefective multiple eigenvalue of A, then $A - \lambda I$ has a zero singular value of multiplicity at least two. It this case, Σ must have a zero singular value, and our analysis fails because the required positive lower bound μ does not exist. The common sense of this situation is that perturbations of $A - \lambda I$ may cause the right and left singular vectors to move independently in subspaces of dimension at least two. This raises the possibility of generating orthogonal right and left inferior vectors, for which the Rayleigh quotient does not exist.[3] Fortunately, this problem should not affect our

[3]Except for the case of Hermitian A, the generalized Rayleigh quotient algorithm (1.1) has an analogous problem.

intended application to subspace methods for large eigenvalue problems, provided the subspaces $\mathcal{V}$ and $\mathcal{W}$ mentioned in the introduction are large enough to accommodate the multiplicity of the eigenvalue.

5. Conclusions

We have proved a result relating an eigenspace of a matrix to its inferior singular subspace, and we have analyzed a variant of a Raleigh quotient algorithm.

In our intended application to large eigenvalue problems where approximate singular vectors are computed by GKL bidiagonalization, the bound on the convergence ratio says that asymptotically we need only have $2\kappa^2\|A\|(\|\delta_v\| + \|\delta_w\|)/\mu < 1$ to improve the approximation to the target eigenvalue. This does not mean that we can dispense with the GKL bidiagonalization, since the shift changes after each refinement step. However, the bound does suggest that we can get adequate convergence with a constant number of bidiagonalization steps during the enhancement process. This will be the subject of a future investigation.

References

[1] D. R. Fokkema, G. L. G. Sleijpen, and H. A. Van der Vorst. Jacobi–Davidson style QR and QZ algorithms for the reduction of matrix pencils. Preprint 941, Department of Mathematics, Universiteit Utrech, 1996.

[2] G. H. Golub and W. Kahan. Calculating the singular values and pseudo-inverse of a matrix. *SIAM Journal on Numerical Analysis*, 2:205–224, 1965.

[3] R. B. Morgan. On restarting the Arnoldi method for large nonsymmetric eigenvalue problems. *Mathematics of Computation*, 65:1213–1230, 1996.

[4] A. M. Ostrowski. On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. III (generalizd Rayleigh quotient and characteristic roots with linear elementary divisors). *Arch. Rational Mech. Anal.*, 3:325–240, 1959.

[5] D. C. Sorensen. Implicit application of polynomial filters in a k-step arnoldi method. *SIAM Journal on Matrix Analysis and Applications*, 13:357–385, 1992.

[6] G. W. Stewart. Computable error bounds for aggregated Markov chains. *Journal of the ACM*, 30:271–285, 1983.

[7] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory.* Academic Press, Boston, 1990.